JOURNAL OF AGRICULTURAL SCIENCES

TARIM BİLİMLERİ DERGİSİ

ANKARA UNIVERSITY FACULTY OF AGRICULTURE

e-ISSN 2148-9297

JAS

Year 25

Volume 31

Issue 01

Ankara University
Faculty of Agriculture

JOURNAL OF AGRICULTURAL SCIENCES

TARIM BILIMLERI DERGISI

e-ISSN: 2148-9297

Ankara - TÜRKİYE

Year **2025** Volume **31** Issue **1**

JOURNAL OF AGRICULTURAL SCIENCES
TARIM BİLİMLERİ DERGİSİ
ANKARA UNIVERSITY FACULTY OF AGRICULTURE
e-ISSN 2148-9297

Product Information

Publisher	Ankara University, Faculty of Agriculture
Owner (On Behalf of Faculty)	Prof. Dr. Hasan Huseyin ATAR
Editor-in-Chief	Prof. Dr. Halit APAYDIN
Journal Administrator	Salih OZAYDIN
Library Coordinator	Dr. Can BESIMOGLU
IT Coordinator	Lecturer Murat KOSECAVUS
Graphic Design	Ismet KARAASLAN
Date of Online Publication	14.01.2025
Frequency	Published four times a year
Type of Publication	Double-blind peer-reviewed, widely distributed periodical
Aims and Scope	JAS publishes high quality original research articles that contain innovation or emerging technology in all fields of agricultural sciences for the development of agriculture.
Indexed and Abstracted in	Clarivate Science Citation Index Expanded (SCIE) Elsevier Scopus TUBITAK-ULAKBIM-TRDizin CAB International EBSCO FAO-AGRIS SOBIAD OpenAire BASE IFIS CNKI

Management Addres

Journal of Agricultural Sciences - Tarım Bilimleri Dergisi
Ankara University Faculty of Agriculture Publication Department 06110 Diskapi/Ankara-Türkiye
Telephone : +90 312 596 14 24 | Fax : +90 312 317 67 24
E-mail: tbdeditor@ankara.edu.tr | http://jas.ankara.edu.tr/

JOURNAL OF AGRICULTURAL SCIENCES
TARIM BİLİMLERİ DERGİSİ
ANKARA UNIVERSITY FACULTY OF AGRICULTURE
e-ISSN 2148-9297

JOURNAL OF AGRICULTURAL SCIENCES
TARIM BİLİMLERİ DERGİSİ
ANKARA UNIVERSITY FACULTY OF AGRICULTURE
e-ISSN 2148-9297

CONTENTS **2025, 31(1)**

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Nanotechnological Applications in Current Innovative Approaches in Dairy Technology- A review

Binnur Kaptan Günüç[a]

[a]*Department of Food Engineering, Faculty of Agriculture, Tekirdağ Namık Kemal University, Tekirdağ, TÜRKIYE*

ARTICLE INFO
Review Article
Corresponding Author: Binnur Kaptan Günüç, E-mail: bkaptan@nku.edu.tr
Received: 26 June 2024 / Revised: 23 September 2024 / Accepted: 25 September 2024 / Online: 14 January 2025

Cite this article
Kaptan Günüç B (2025). Nanotechnological Applications in Current Innovative Approaches in Dairy Technology- A review. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):1-11. DOI: 10.15832/ankutbd.1505367

ABSTRACT
Nanotechnology offers significant potential in the dairy industry, influencing a range of products such as milk, cheese, yogurt, butter, fermented milk, and buttermilk. The use of both bottom-up and top-down processing approaches yields extensive insights into the intrinsic and extrinsic characteristics of dairy products. A variety of nano-techniques including nanoemulsion, nanoencapsulation, nanoliposomes, nanotubes, nanofibers, and nanocapsules are employed within the dairy industry. These methods, in conjunction with nanosensors, nanolaminates, and nanocoatings, act as efficient packaging solutions, providing critical information on product safety, stability, and quality. Nanotechnology is also highly effective in detecting foodborne pathogens and excels in pathogen control. It plays an essential role in food fortification by encapsulating nutrients, ingredients, and compounds, thus enhancing the texture, taste, aroma, quality, and nutritional value of food. Notably, nanoparticles such as zinc oxide, silver, and titanium demonstrate potent mechanisms for disrupting bacterial cell walls, aiding in compound absorption, and improving physiological functions. This review presents the applications of nanotechnology in the dairy industry, along with pertinent studies and their outcomes regarding the utilization of these technologies in dairy products.

Keywords: Nanoparticles, Nanosensor, Nanoemulsion, Nanoencapsulation, Nanopackaging, Dairy Technology

1. Introduction

In recent years, the field of nanotechnology has seen significant growth, leading to the discovery of the various nano-sized materials such as nanoparticles, nanocomposites, and nanoemulsions worldwide. Nanotechnology is a multidisciplinary field that studies and manipulates nanostructures (0.1-100 nm) with various applications in different domains, including food science and dairy products (Delshadi et al. 2020; Omerović et al. 2021). Nanotechnology has introduced edible films and coatings for food packaging, preventing gas and moisture exchange and enhancing product safety and longevity (Nickols-Richardson & Piehowski, 2008; Sharma et al. 2017). Nanotechnology has also enabled the development of the functional dairy products fortified with minerals, vitamins, and biologically active ingredients that offer health benefits and address nutritional deficiencies (van der Hee et al. 2009). For example, some yogurts can increase calcium intake by 50% and provide vitamin D-3, improving bone health.

Nanotechnology enhances the sensory properties of dairy products by optimizing the functionality, solubility, stability, and bioavailability of milk components using nano-sized materials such as emulsifiers, stabilizers, nanostructured lipids, and proteins (Ghorbanzade et al. 2017; Braicu et al. 2019; Nile et al. 2020) Nanomaterials enhance the texture, consistency, emulsifying and foaming properties, taste, aroma, quality, and nutritional value of dairy products (Sharma et al. 2017). Nanotechnology opens up new possibilities for designing functional dairy products that align with consumers' needs and preferences (Nile et al. 2020). Nanotechnology has a significant impact on dairy product packaging, enhancing its safety, stability, and quality. Key features of this technology include nanosensors, nanolaminates, and nanocoatings, which aid in pathogen detection. In the dairy food industry, as with other sectors, nanotechnology plays a crucial role as a promising technology for solving problems via innovative solutions related to dairy food industry safety, processing, packaging (Figure 1), as well as functional milk products.

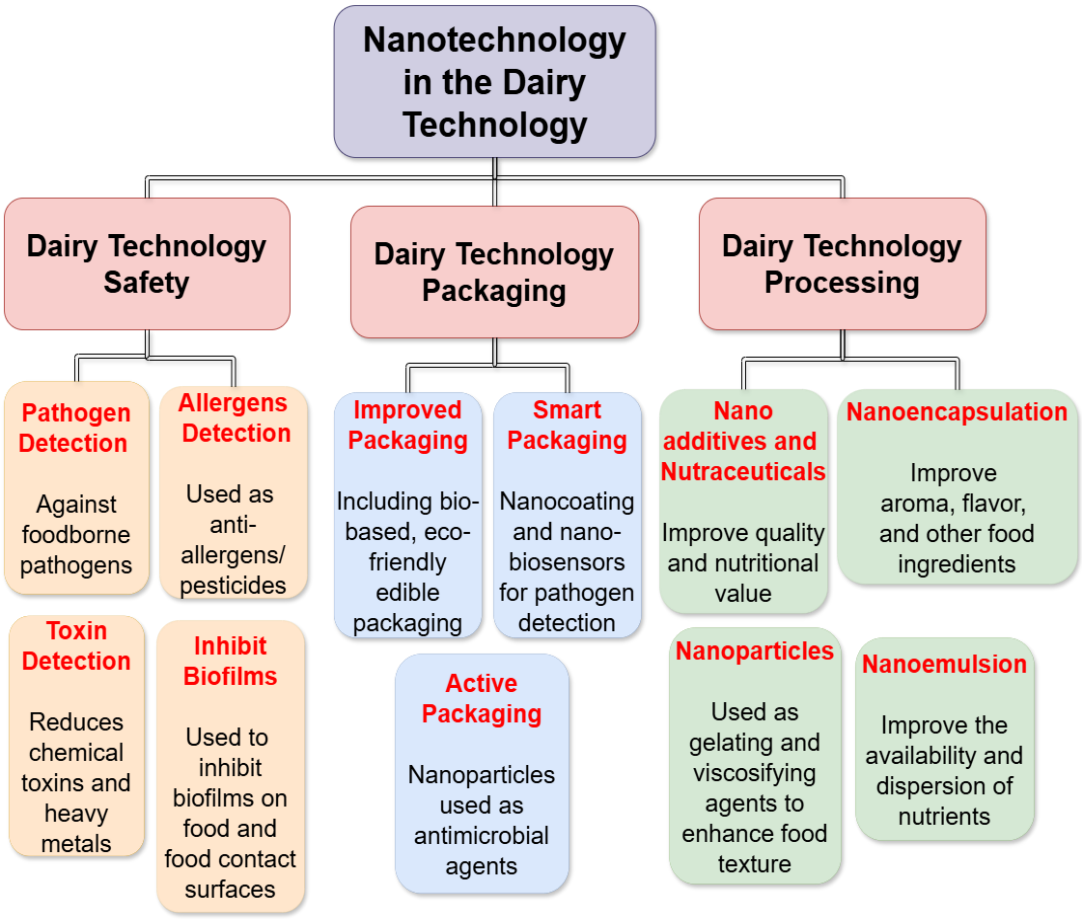


Figure 1 - The applications of nanotechnology in dairy technology safety, packaging and processing

Nanoparticles like zinc oxide, silver, and titanium disrupt the bacterial cell walls, improving bodily functions and enhancing compound absorption. These particles effectively detect the foodborne pathogens like *Listeria monocytogenes* (*L. monocytogenes*), *Salmonella typhimurium* (*S. typhimurium*), and *Pseudomonas aeruginosa* (*P. aeruginosa*), raising safety standards. Nanocapsules protect and release the bioactive ingredients, enhancing bioavailability and utilization of active components.

Nanopackaging extends the shelf life by inhibiting the microbial growth. Nanoadditives improve the sensory qualities like color, taste, and aroma. Nanosensors offer advanced traceability for the food quality and enhancing safety. Nanoemulsions can either reduce the fat content or increase the nutritional value in the dairy products, as shown by Khan et al. (2019). Nanotechnology applications in dairy products include nanoencapsulation, nanoemulsion, nanoliposomes, nanotubes, and nanofibers (Table 1). These applications improve texture, taste, and nutritional content, elevating the dairy product quality

Table 1 - Some studies on nanotechnological methods and their applications in dairy technology

Method	*Bioactive structures*	*Applications*	*Dairy Products*	*References*
Nanoliposome	Nisin	Antibacterial	Milk	da Silva Malheiros et al. 2010
Nanoliposomes	Catechin or green tea extract	Fortification	Cheese (Low-fat)	Rashidinejad et al. 2016
Nanoliposome	Bacteriocin CAMT6	Antibacterial	Whole and Skim milk	Li et al. 2023
Nanoliposomes	β-sitosterol	Fortification	Butter	Bagherpour et al. 2017
Nanoemulsion	Oregano essential oil	Coatings	Cheese (Low fat)	Artiga-Artigas et al. 2017
Nanoemulsion	Pectin/oregano oil	Coating	Cheese	Miljkovic et al. 2017
Nanoemulsions	Curcumin	Fortification	Ice cream	Borrin et al. 2018
Nanoemulsions	Zeaxanthin	Fortification	Yogurt	de Campo et al. 2019
Nanoemulsion	Carotenoid	Functional	Yogurt	Medeiros et al. 2019
Nanoemulsions	D Vitamin	Fortification	Cheddar Cheese	Banville et al. 2000
Nanoemulsions	Conjugated Linoleic acid	Fortification	Low in fat milk	Hashemi et al. 2020
Nanoemulsion	Nigella sativa oil	Fortification	Icecream	Mohammed et al. 2020
Nanoemulsion	Curcumin	Functional	Soft Cheese	Bagale et al. 2023
Nanoemulsions	*Eugenol and Cinnamaldehyde*	Antibacterial	Yogurt	Abdelhamid et al. 2023
Nanoemulsion	*Laurus nobilis L. extract*	Fortification	Processed cheeses	Hussein et al. 2023
Microencapsulation	Astaxanthin	Fortification	Yogurt	Feng et al. 2018
Nanoencapsulation	Fish oil (EPA-DHA)	Fortification	Yogurt	Ghorbanzade et al. 2017
Nanoencapsulation	Vitamin A, D_3	Fortification	Skim milk	Loewen et al. 2018
Nanoencapsulation	Vitamin D_3	Fortification	Yogurt based beverage (Lassi)	Maurya et al. 2019
Nanoencapsulation	Vitamin D_2	Fortification	Milk	Syama et al. 2019
Nanoencapsulation	Saffron extract	Functional	Ricotta cheese	Siyar et al. 2022
Nanoencapsulation	*Curcumin*	Functional	Dairy drink	Wang et al. 2022

2. Nanotechnological Methods Applied in Dairy Technology

2. 1. Nanoencapsulation

The use of nanotechnology in food technology has ushered in a new era, particularly through advancements in enhancing the bioavailability of milk components. Nanoencapsulation involves encapsulating bioactive compounds within nanoscale carriers, thereby shielding them from environmental factors and enabling controlled release during digestion (Rezaei et al. 2019; Pateiro et al. 2021; Maqsoudlou et al. 2022; Mohammad et al. 2022; Sharma et al. 2023).

This technology plays a crucial role in fortifying dairy products with essential nutrients like vitamins and minerals, ensuring their stability during processing and storage. This helps reduce food waste and extends the shelf life of dairy products (Chen et al. 2017; Dong & Zhong 2019). Incorporating probiotics into dairy products further enhances their health benefits. Nanoencapsulation acts as a protective measure, prolonging the shelf life of products and maintaining probiotic viability upon consumption. Research has demonstrated the efficacy of nanoencapsulation in enhancing the stability of probiotics, vitamins, and bioactive peptides in dairy products (Delshadi et al. 2020; Pudtikajorn et al. 2021).

The benefits of nanoencapsulation play a pivotal role in advancing the development of dairy products with improved nutritional quality and extended shelf life. Nanoencapsulation preserves the structural integrity of these compounds and facilitates controlled release after consumption. Nanoencapsulation also aligns with the consumer demand for functional foods, which deliver essential nutrition and additional health benefits. Nanoencapsulated dairy products enriched with probiotics, vitamins, and bioactive peptides address this demand by combining nutritional value with improved bioavailability. However, nanoencapsulation in the dairy industry faces challenges such as cost, regulatory approvals, and consumer safety concerns. Ongoing research and collaboration among scientists, industry stakeholders, and regulatory bodies are needed to overcome these challenges and fully realize the potential of nanoencapsulation in the dairy sector. In conclusion, nanoencapsulation technology introduces a ground-breaking paradigm in the dairy industry, contributing to the evolution of functional dairy products with enhanced health benefits. Sustained research efforts in this domain instill optimism for a future in the dairy industry characterized by innovative and health-promoting products.

2. 2. Nanoemulsions

International Nanoemulsions, colloidal systems with oil droplets of two immiscible liquids in an oil-in-water (O/W) or water-in-oil (W/O) configuration, are characterized by their small size, ranging from 10 to 100 nm (Saxena & Bhardwaj, 2017). These nanoscale structures (Figure 2) are often used as stabilizing agents for essential oils and offer several advantages over conventional emulsions.

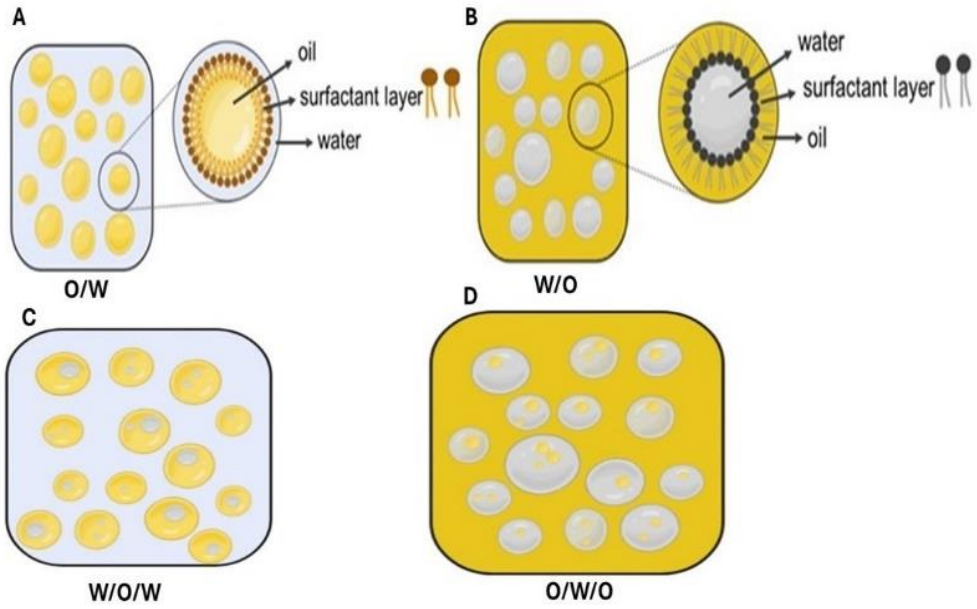


Figure 2 – Illustration of various types of nanoemulsions, depicting both two-phase (A and B) and multiple-phase (C and D) structures. A: oil-in-water (O/W) emulsion, B: water-in-oil (W/O) emulsion, C: water/oil/water (W/O/W) emulsion, D: oil/water/oil (O/W/O) emulsion

These include a higher surface area, lower light scattering, superior optical properties, enhanced bioavailability, and extended shelf life (Silva et al. 2012; Lane et al. 2013). In the dairy industry, nanoemulsions are used to encapsulate and transport various ingredients, such as antioxidants, antimicrobials, vitamins, minerals, flavors, and colorants. They also serve structural purposes by enhancing emulsion stability, reducing fat globule size, and improving creaminess and mouthfeel. The manipulation of droplet size in nanoemulsions significantly influences emulsion stability and properties such as color, taste, odor, nutritional value, and bioavailability, presenting an innovative approach to enhancing the physicochemical and sensory attributes of milk and dairy products. Nanoemulsion technology is instrumental in extending shelf life, increasing nutritional value, improving bioavailability, transporting functional ingredients, and developing novel dairy products (McClements 2010). Nanoemulsion technology, compatible with dairy processing equipment and methods, can be implemented through various techniques, including high-pressure homogenization, ultrasonic homogenization, or microfluidic devices (Lohith Kumar & Sarkar 2018).

This technology facilitates the controlled release of active ingredients, such as nutrients, proteins, and antioxidants, enhancing their effectiveness and bioavailability within the body. Traditional dairy formulations often encounter challenges such as phase separation, precipitation, and creaming. Nanoemulsion addresses these issues by decreasing droplet size, thereby increasing the surface area of dispersed phases, leading to heightened homogeneity and stability (Montes de Oca-Ávalos et al. 2017). This plays a pivotal role in enhancing the stability of dairy products by preventing phase separation, precipitation, and creaming. Critical sensory attributes like taste, texture, and appearance influence consumer acceptance of dairy products.

Nanoemulsion technology provides a platform to tailor these sensory properties by precisely controlling droplet size and distribution. Research suggests that nanoemulsions have a positive effect on creaminess, mouthfeel, and the overall sensory experience of dairy products, thereby influencing consumer preferences (Panghal et al. 2019). Successful applications of nano-sized emulsions with low-fat content have been demonstrated in ice cream production (Silva et al. 2012). The incorporation of nanoemulsion purple rice bran oil into frozen yogurts has improved the nutritional profile by increasing the natural antioxidant content (Sanabria 2012). Furthermore, stable curcumin nano-emulsions have exhibited effective oxygen scavenging in commercial milk samples, remaining stable at room temperature for up to a month (Joung et al. 2016). Moreover, nanoemulsion technology has been investigated for its potential in controlled lactose hydrolysis within milk. Freeze-dried capsules from S/O/W emulsions containing lactase have demonstrated the potential to gradually release lactase into dairy products, offering controlled lactose hydrolysis over three weeks of storage (Zhang and Zhong 2018). This suggests that S/O/W emulsions have the potential to enhance the quality and functionality of milk and dairy products by enabling the controlled release of lactase. In conclusion, nanoemulsion technology represents a substantial advancement in the dairy industry, providing a versatile and efficacious method for enhancing the quality, stability, and functionality of milk and dairy products.

This technology is instrumental in extending shelf life, increasing nutritional value, improving bioavailability, transporting functional ingredients, and developing novel dairy products. In conclusion, nanoemulsion technology represents a substantial advancement in the dairy industry, providing a versatile and effective method for enhancing the quality, stability, and functionality of milk and dairy products.

2. 3. Nanoliposome

Nanoliposome technology has been recognized for its potential to significantly improve the quality and functionality of dairy products. These nanocarriers, which are spherical lipid-based structures typically ranging from 20 to 200 nanometers, utilize a phospholipid bilayer to encapsulate an aqueous core. This allows for the efficient delivery of both hydrophobic and hydrophilic molecules (Khorasani et al. 2018; Brandelli et al. 2023).

With attributes such as a high surface area, biocompatibility, and customizable functionality, nanoliposomes enable a wide range of applications, particularly in the encapsulation of bioactive compounds like antioxidants, vitamins, and antimicrobial agents. They play a significant role in the development of functional dairy products by delivering nutrients to targeted cells or tissues (Zarrabi et al. 2020; Bondu & Yen 2022; Khafoor et al. 2023). The excellent encapsulation efficiency, stability, and controlled release properties of nanoliposomes ensure the protection of sensitive compounds during processing and storage, leading to optimal bioavailability in dairy products (Li et al. 2019). Nanoliposomes specifically designed to contain antioxidants are crucial in preventing oxidation reactions, thus preserving the nutritional quality and extending the shelf life of dairy products. When added to dairy product formulations, nanoliposomes effectively shield both hydrophilic and hydrophobic compounds, enhancing the bioavailability and stability of nutrients and aiding in the development of functional dairy products (Ghorbanzade et al. 2017; Srinivasan et al. 2019).

Antimicrobial nanoliposomes act as effective barriers against microorganisms, significantly prolonging the shelf life of dairy products. Research has shown that nanoliposomes, including those containing garlic extract, are effective in exerting antimicrobial effects against Listeria species in milk (Schmidt et al. 2009; Pinilla et al. 2017; Todorov et al. 2022). The combination of liposome-encapsulated nisin and garlic has been effective against pathogens such as *Escherichia coli* (*E. coli*), *Staphylococcus aureus* (*S. aureus*), Salmonella enteritidis, and *L. monocytogenes* in whole milk (Pinilla & Brandelli, 2016).

In cheese production, the use of nanoliposomes encapsulating antimicrobial peptides has been shown to extend the shelf life of cheese. This includes the use of partially purified soybean phosphatidylcholine (PC) and PC-cholesterol (7:3) nanoliposomes (Bagale et al. 2023). Bagale et al. (2023) also highlight the antimicrobial activity of curcumin encapsulated in liposomes against *S. aureus* and *E. coli* in milk-based soft cheese. Nanoliposomes have been effective in addressing related challenges to the incorporation of free antimicrobials into dairy products, demonstrating broad antimicrobial activity in white soft cheese (Mohammadi et al. 2015; Lopes & Brandelli 2018; El-Sayed & El-Sayed 2021.

Additionally, in fortifying milk matrices with omega-3 fatty acids, nanoliposomes have been effective in enhancing acceptance and improving taste, aroma, texture, and appearance compared to unencapsulated fish oil (Ghorbanzade et al. 2017; Rasti et al. 2017; Patel et al. 2020). Nanoliposomes play an important role in preserving bioactive compounds in milk that are sensitive to environmental factors, thus maintaining nutritional quality and extending the shelf life of dairy products. For products designed for lactose intolerance patients, nanoliposomes encapsulating the β-galactosidase enzyme have been determined to exhibit long-term activity at low temperatures, maintaining functional activity for 20 days to one month at +5 °C. In conclusion, nanoliposomes play a diverse role in the dairy industry, offering advantages in enhancing nutritional value, improving stability, and extending the shelf life of dairy products. The dynamically evolving field of nanoliposome technology in the dairy industry presents continued opportunities through advancements in nanotechnology

2. 4. Nanotubes

Nanotubes represent significant innovations in dairy technology, offering several advantages such as enhancing food texture and increasing colloidal stability (Katouzian & Jafari 2019). Nanotube-shaped nanocarriers, particularly those derived from milk protein α-lactalbumin, possess unique properties that can improve the quality of milk and extend its shelf life (Li et al. 2019). These nanotubes can be utilized for nanoencapsulating bioactive compounds, including vitamins and antioxidants, in fortified dairy products. This protective technique prevents the degradation of these compounds during processing and storage, ensuring consumers derive maximum nutritional benefits from milk.

Nanotubes are formed through the partial hydrolysis of α-lactalbumin by a protease from Bacillus licheniformis (BLP) under specific conditions, a process that has been studied extensively (Ipsen et al. 2001; Graveland-Bikker et al. 2006). Whey-based α-lactalbumin nanotubular gels, known for their stable structures and ability to carry active agents, show promise for various industrial applications (Tarhan et al. 2021). Protein nanotubes (PNTs) exhibit stability by forming a β-sheet structure, which arises from interactions between aromatic side chains and charged residues. This stability makes PNTs valuable nanocarriers in drug delivery applications (Wei et al. 2018). PNTs offer advantages such as prolonged intestinal retention time, enhanced cellular absorption, and increased membrane permeability, ultimately increasing the bioavailability of hydrophobic bioactive compounds (Bao et al. 2020).

In dairy applications, the addition of lycopene-loaded α-lactalbumin nanotubes to beverages has been shown to increase viscosity and enhance long-term stability, which is promising for the encapsulation and delivery of hydrophobic compounds (Chang et al. 2022). Furthermore, α-lactalbumin nanotubes have facilitated the survival of soy isoflavones through simulated in

vitro gastric digestion and promoted their release in the intestinal phase when added to soy milk (Liu et al. 2023). Nanotubes also offer a promising solution for challenges related to the water solubility, stability, and bioavailability of compounds such as curcumin. Encapsulation within α-lactalbumin nanocarriers enables curcumin to exhibit responsive release behavior in the intestines, suggesting potential for developing functional dairy beverages (Wang et al. 2022).

In conclusion, nanotubes present significant potential for various applications within the dairy industry. This technology holds promise for improving milk quality, extending shelf life, enhancing heat transfer efficiency, implementing milk monitoring through sensors, and facilitating the nanoencapsulation of bioactive compounds (Li et al. 2019).

2. 5. Nanofiber

Nanofiber technology investigates materials composed of fibers at the nanoscale. This includes biopolymer nanofibers produced via electrospinning, which are made up of cross-linked biopolymer molecules (Al-Abduljabbar & Farooq, 2023). Nanofibers, whether synthesized from synthetic or natural polymers, offer advantages such as lightweight, straightforward processing, and important loading capacity. In the dairy industry, nanofibers contribute to enhanced milk processing efficiency, improved product quality, and extended shelf life by facilitating the development of antimicrobial and active food packaging materials, as well as sensors that monitor pH or temperature changes during storage (Kumar et al. 2019; Senthil et al. 2019; Rostamabadi et al. 2020). A significant application of nanofibers in the dairy industry is filtration, where electrospun nanofiber membranes prove effective in separating impurities, bacteria, and contaminants in milk, surpassing traditional filtration methods (Li et al. 2019). Furthermore, nanofiber technology aids in developing packaging materials with improved barrier properties, creating an effective shield against oxygen, moisture, and external factors, leading to reduced microbial growth and product spoilage (Zhang et al. 2023).

Nanofibers containing milk proteins find applications in yogurt and cheese production, while those containing nano cellulose and nano chitin are proposed for use in the packaging of dairy products (Li et al. 2019; Senthil et al. 2019). Electrospun nanofibers of sodium alginate, used to encapsulate probiotic *L. brevis* cells, have been determined to significantly increase the survival of these cells compared to free bacteria when included in a functional yoghurt drink (Mohaisen et al. 2019). In addition, a smart promoter/polyvinyl alcohol film containing anthocyanin and limonene (Lui et al. 2017) has been reported to exhibit inhibitory activity against *S. aureus*, *B. subtilis*, and *A. niger* in pasteurized milk, demonstrating effectiveness in pH monitoring. Furthermore, an antimicrobial packaging material, created by separating allyl isothiocyanate in the vapor phase and incorporating it into gelatin electrospun fibers, was reported to extend the shelf life of cheese (Al-Moghazy et al. 2020). Additionally, nanofibers composed of moringa oil-loaded chitosan nanoparticles, utilized as preservatives for the biocontrol of *L. monocytogenes* and *S. aureus* in cheese, have been stated to exhibit high antibacterial activity without compromising sensory quality (Lin et al. 2019).

These systems encapsulate bioactive compounds within nanofiber structures during the production of milk and dairy products, thereby increasing stability, enabling controlled release, and providing protection against external factors such as heat and oxidation (Xue et al. 2019). In conclusion, research continues on the development of innovative nanofiber delivery systems in the dairy industry.

2. 6. Nanosensors

Advances in nanotechnology have accelerated the development of nanosensor technology, with recent progress in nanomaterials opening up potential applications in various fields, including the monitoring of milk and dairy products (Baranwal et al. 2022). Nanosensors, designed at the nanoscale (1 to 100 nm), offer advantages such as simple synthesis, easy surface chemistry, and cost-effectiveness. In the dairy industry, nanosensor technology serves purposes such as quality control, safety assurance, and real-time monitoring. The nanosensors, commonly used as labels and reporters, offer notable improvements in speed, selectivity, and sensitivity compared to traditional methods. Typically, nanosensors consist of a receptor that interacts with the analyte and a transducer that converts the response into measurable signals such as electrical, electrochemical, or optical signals. Constructed from nanomaterials like carbon nanotubes, quantum dots, graphene oxide, metallic gold, iron oxide particles, and electrospun polymeric nanofibers, nanosensors enhance the efficiency and sensitivity of analyses. The unique properties of nanomaterials, such as flexible shapes, enhanced optical, mechanical, electrical, and thermal properties, as well as high specific surface area and nanoconfinement, significantly enhance their signal transduction capabilities (Manoj et al. 2021). When evaluating nanosensors, key considerations include selectivity, which describes the system's ability to distinguish the analyte from other sample components.

Nanosensors are classified based on the transducer mechanism that generates output signals. These mechanisms encompass electrical, electrochemical, thermal, mechanical, optical, gravimetric, piezoelectric, and magnetic methods (Manoj et al. 2021). Electrochemical nanosensors measure electrochemical reactions with the analyte effectively (Stasyuk et al. 2022) using nanomaterials on carbon-based (Zhou et al. 2020) or glassy carbon electrodes (Jena & Raj 2008). Optical nanosensors, utilizing nanomaterials like metallic nanoparticles, upconversion nanoparticles, carbon nanotubes, and quantum dots, detect and analyze light, for measuring various physical, chemical, and biological parameters (Khansili et al. 2018; Pu et al. 2021). During the

processing, storage, and distribution stages, the dairy industry consistently encounters challenges in ensuring the safety and quality of milk and dairy products. Recent advances in nanotechnology have addressed these issues by developing nanosensors that offer real-time monitoring and precise detection capabilities (Wang et al. 2021; Kadam et al. 2022). Nanosensors contribute to improving product quality by enabling real-time monitoring of critical parameters such as temperature, humidity, and gas composition of milk and dairy products during processing, storage, and distribution (Kumar et al. 2021). Nanosensors have also been utilized to detect biological contaminants such as antibiotics, toxins, and pathogens, as well as chemical compounds such as pesticides and heavy metals in milk and dairy products (Liu et al. 2008; El-Ansary & Faddah 2010; Lopez & Merkoci 2011; Yotova et al. 2013; Berekaa 2015; Song et al. 2017; Bajpai et al. 2018). Moreover, nanosensors expedite the identification of substances that compromise milk purity, including urea, melamine (Kaneko et al. 2018), formaldehyde (Agharkar & Mane 2021; Hegde et al. 2021), detergents, ammonium sulfate, benzoic acid, salicylic acid (Bueno et al. 2014), and hydrogen peroxide (Nanda et al. 2016; Guinati et al. 2021; Vasconcelos et al. 2021), thus playing a pivotal role in preventing contaminated products from reaching consumers.

Nanosensors are instrumental in monitoring the shelf life of these products by evaluating freshness, spoilage, and microbial growth. This monitoring capability significantly enhances quality control within the dairy industry (Nascimento et al. 2017). Early detection is paramount for ensuring food safety and maintaining the integrity of dairy products. Consequently, nanosensor technology is becoming an indispensable tool for guaranteeing the safety and quality of milk and dairy products. Integrating nanosensors into the dairy industry's quality control processes is a vital step towards fulfilling the growing demands for precision and efficiency in dairy production

3. Challenges

In the dairy industry, the effective utilization of nanotechnology for innovative products and processes faces several challenges. A primary concern is developing safe and efficient distribution systems suitable for human consumption, requiring advanced yet cost-effective processing methods. Recognizing potential hazards, assessing toxicity, and addressing environmental concerns regarding nanoparticles are critical steps in this process. Specifically, attention should focus on the leaching and migration of nanoparticles from packaging materials into dairy products. Nanoparticles can bypass biological barriers, penetrating various tissues and organs, which raises significant safety concerns. Moreover, the synthesis of nanoparticles through different chemical processes poses a risk of harmful by-products, contributing to environmental pollution.

Therefore, a comprehensive understanding of the functions and toxicity of nanomaterials is essential for establishing practical usage and safety standards for nanotechnology in the dairy industry. A significant challenge lies in the fact that materials behave differently at the nanoscale, and our understanding of how to analyze these behaviors remains limited. Before the commercial use and production of antibacterial nanoparticles with environmentally friendly properties, more *in vitro* and in *vivo* studies involving the interaction of nanoparticles with living organisms are necessary. These studies are crucial for the responsible and safe integration of nanotechnology into the dairy industry, ensuring that all factors are carefully considered.

Consequently, a comprehensive risk assessment program, the establishment of regulatory frameworks, biosecurity considerations, and the resolution of public concerns are necessary prerequisites for the production, packaging, and consumption of nanotechnology-based dairy products.

4. Conclusions

Recent breakthroughs in nanotechnology for the dairy industry have shown promising advancements in enhancing the quality, safety, and efficiency of dairy products. One notable innovation is the nanoencapsulation technique, which improves the water solubility and antimicrobial activity of compounds. These encapsulated compounds can also be integrated into food-packaging films to act as antimicrobial agents, thereby enhancing food safety. Nanosensors are another significant development, being developed for smart packaging that monitors the quality of stored foods and detects contaminants, ensuring higher standards of food safety. Nano-based delivery systems have been successfully employed in dairy products to encapsulate vitamins, antioxidants, flavors, and minerals. This approach not only enhances the bioactive properties but also boosts the nutritional value of dairy products.

Nanostructured materials are being explored for their role in preventing spoilage and oxidation, thereby extending the shelf life of dairy products. These advancements reflect a broader trend where nanotechnology is leveraged to develop new functional ingredients, optimize processing at micro- and nanoscales, and enhance overall food safety and quality. However, integrating nanotechnology into the dairy industry poses challenges. Establishing a safe distribution system is complex and costly, while understanding the behaviors of nanomaterials and ensuring regulatory compliance are crucial for managing their potential toxicity and optimizing their functions. Despite these challenges, nanotechnology holds promise for revolutionizing functional dairy formulation and production, offering novel opportunities for innovation in the industry.

References

Abdelhamid S M, Edris A E & Sadek Z (2023). Novel approach for the inhibition of *Helicobacter pylori* contamination in yogurt using selected probiotics combined with eugenol and cinnamaldehyde nanoemulsions. *Food chemistry* 417: 135877. https://doi.org/10.1016/j.foodchem.2023.135877

Agharkar M & Mane S (2021). Utilization of gold nanoparticles to detect formalin adulteration in milk. Materials Today: Proceedings 45: 4421-4423. https://doi.org/10.1016/j.matpr.2020.12.233

Al-Abduljabbar A & Farooq I (2023). Electrospun polymer nanofibers: Processing, properties, and applications. *Polymers* 15: 65. https://doi.org/10.3390/polym15010065

Al-Moghazy M, Mahmoud M & Nada A A (2020). Fabrication of cellulose-based adhesive composite as an active packaging material to extend the shelf life of cheese. *International Journal of Biological Macromolecules* 160*:* 264-275. https://doi.org/10.1016/j.ijbiomac.2020.05.217

Artiga-Artigas M, Acevedo-Fani A & Martín-Belloso O (2017). Improving the shelf life of low-fat cut cheese using nanoemulsion-based edible coatings containing oregano essential oil and mandarin fiber. *Food Control* 76*:* 1-12. https://doi.org/10.1016/j.foodcont.2017.01.001

Bagale U, Kadi A, Abotaleb M, Potoroko I & Sonawane S H (2023). Prospect of bioactive curcumin nanoemulsion as effective agency to improve milk based soft cheese by using ultrasound encapsulation approach. *International Journal of Molecular Sciences* 24: 2663. https://doi.org/10.3390/ijms24032663

Bagherpour S, Alizadeh A, Ghanbarzadeh S, Mohammadi M & Hamishehkar H (2017). Preparation and characterization of beta-sitosterol-loaded nanostructured lipid carriers for butter enrichment. *Food Bioscience* 20: 51-55. https://doi.org/10.1016/j.fbio.2017.07.010

Bajpai V K, Kamle M, Shukla S, Mahato D K, Chandra P, Hwang S K, Kumar P, Huh Y S & Han Y K (2018). Prospects of using nanotechnology for food preservation, safety, and security. *Journal of Food and Drug Analysis*, 26: 1201–1214. https://doi.org/10.1016/j.jfda.2018.06.011

Banville C, Vuillemard J C & Lacroix C (2000). Comparison of different methods for fortifying Cheddar cheese with vitamin D. *International Dairy Journal* 10: 375-382. https://doi.org/10.1016/S0958-6946(00)00054-6

Baranwal J, Barse B, Gatto G, Broncova G & Kumar A (2022). Electrochemical sensors and their applications: A review. *Chemosensors* 10: 363. https://doi.org/10.3390/chemosensors10090363

Bao C, Liu B, Li B, Chai J, Zhang L, Jiao L, Li D, Yu Z, Ren E, Shi X & Li Y (2020). Enhanced transport of shape and rigidity-tuned?-lactalbumin nanotubes across intestinal mucus and cellular barriers. *Nano Letters* 20(2). https://doi.org/10.1021/acs.nanolett.9b04841

Berekaa M M (2015). Nanotechnology in food industry; advances in food processing, packaging and food safety. *International Journal of Current Microbiology and Applied Sciences* 4: 345-357. ISSN: 2319-7706. http://www.ijcmas.com

Bondu C & Yen F T (2022). Nanoliposomes, from food industry to nutraceuticals: Interests and uses. *Innovative Food Science and Emerging Technologies* 81. https://doi.org/10.1016/j.ifset.2022.103140

Borrin T R, Georges E L, Brito-Oliveira T C, Moraes I C F & Pinho S C (2018). Technological and sensory evaluation of pineapple ice creams incorporating curcumin-loaded nanoemulsions obtained by the emulsion inversion point method. *International Journal of Dairy Technology* 71(2). https://doi.org/10.1111/1471-0307.12451

Braicu C, Gulei D, Raduly L, Harangus A, Rusu A & Berindan-Neagoe I (2019). Altered expression of miR-181 affects cell fate and targets drug resistance-related mechanisms. *Molecular Aspects of Medicine* 70: 90-105. https://doi.org/10.1016/j.mam.2019.10.007

Brandelli A, Lopes N A & Pinilla C M B (2023). Nanostructured antimicrobials for quality and safety improvement in dairy products. Foods, 12: 2549. https://doi.org/10.3390/foods12132549

Bueno L, de Araujo W R, Salles M O, Kussuda M Y & Paixão T R L C (2014). Voltammetric electronic tongue for discrimination of milk adulterated with urea, formaldehyde and melamine. *Chemosensors* 2: 251-266. https://doi.org/10.3390/chemosensors2040251

Chang R, Liu B, Wang Q, Zhang J, Yuan F, Zhang H, Chen S, Liang S & Li Y (2022). The encapsulation of lycopene with α-lactalbumin nanotubes to enhance their anti-oxidant activity, viscosity and colloidal stability in dairy drink. *Food Hydrocolloids* 131: https://doi.org/10.1016/j.foodhyd.2022.107792

Chen F, Fan G Q, Zhang Z, Zhang R, Deng Z Y & McClements D J (2017). Encapsulation of omega-3 fatty acids in nanoemulsions and microgels: Impact of delivery system type and protein addition on gastrointestinal fate. *Food Research International* 100: 387–395. https://doi.org/10.1016/j.foodres.2017.07.039

da Silva Malheiros P, Daroit D J & Brandelli A (2010). Food applications of liposome-encapsulated antimicrobial peptides. *Trends in Food Science & Technology* 21*:* 284–292. https://doi.org/10.1016/j.tifs.2010.03.003

de Campo C, Queiroz Assis R, Marques da Silva M, Haas Costa T M, Paese K, Stanisçuaski Guterres S, de Oliveira Rios A & Hickmann Flôres S (2019). Incorporation of zeaxanthin nanoparticles in yogurt: Influence on physicochemical properties, carotenoid stability and sensory analysis. *Food Chemistry* 301: 125230. https://doi.org/10.1016/j.foodchem.2019.125230

Delshadi R, Bahrami A, Tafti A G, Barba F J & Williams L L (2020). Micro and nano-encapsulation of vegetable and essential oils to develop functional food products with improved nutritional profiles. *Trends in Food Science & Technology* 104: 72–834. https://doi.org/10.1016/j.tifs.2020.07.004

Dong L & Zhong Q (2019). Dispersible biopolymer particles loaded with lactase as a potential delivery system to control lactose hydrolysis in milk. *Journal of Agricultural and Food Chemistry* 67: 6559–6568. https://doi.org/10.1021/acs.jafc.9b01546

El-Ansary A & Faddah L M (2010). Nanoparticles as biochemical sensors. *Nanotechnology, Science and Applications* 3: 65-76. https://doi.org/10.2147/NSA.S8199

El-Sayed H S & El-Sayed S M (2021). A modern trend to preserve white soft cheese using nano-emulsified solutions containing cumin essential oil. *Environmental Nanotechnology, Monitoring & Management* 100499. https://doi.org/10.1016/j.enmm.2021.100499

Feng Z Z, Li M Y, Wang Y T & Zhu M J (2018). Astaxanthin from Phaffia rhodozyma: Microencapsulation with carboxymethyl cellulose sodium and microcrystalline cellulose and effects of microencapsulated astaxanthin on yogurt properties. *LWT* 96: 234-241. https://doi.org/10.1016/j.lwt.2018.04.084

Ghorbanzade T, Jafari S M, Akhavan S & Hadavi R (2017). Nano-encapsulation of fish oil in nano-liposomes and its application in fortification of yogurt. *Food Chemistry* 216: 146–152. https://doi.org/10.1016/j.foodchem.2016.08.022

Graveland-Bikker J F, Fritz G, Glatter O & de Kruif C G (2006). Growth and structure of α-lactalbumin nanotubes. *Journal of Applied Crystallography* 39: 180-184. https://doi.org/10.1107/S0021889805043244

Guinati B G S, Sousa L R, Oliveira K A & Coltro W K (2021). Simultaneous analysis of multiple adulterants in milk using microfluidic paper-based analytical devices. *Analytical Methods: Advancing Methods and Applications* 13: 5383-5390. https://doi.org/10.1039/D1AY01339D

Hashemi F S, Farzadnia F, Aghajani A, NobariAzar F A & Pezeshki A (2020). Conjugated linoleic acid loaded nanostructured lipid carrier as a potential antioxidant nanocarrier for food applications. *Food Science & Nutrition* 8: 4185–4195. https://doi.org/10.1002/fsn3.1712

Hegde H R, Chidangil S & Sinha R K (2021). Refractive index and formaldehyde sensing with silver nanocubes. *Royal Society of Chemistry Advances* 11: 8042-8050. https://doi.org/10.1039/D0RA10161C

Hussein J, El-Bana M, Abdel Latif Y, El-Sayed S, Youssef A, Elnaggar M & Medhat D (2023). Processed cheeses fortified by Laurus nobilis L. extract nanoemulsion ameliorate hyperhomocysteinemia in Ehrlich ascites carcinoma model. *Egyptian Journal of Chemistry* 66(2): 199-211. https://doi.org/10.21608/ejchem.2022.135198.5944

Ipsen R, Otte J & Qvist K B (2001). Molecular self-assembly of partially hydrolysed α-lactalbumin resulting in strong gels with a novel microstructure. *Journal of Dairy Research* 68(2). https://doi.org/10.1017/S0022029901004769

Jena B K & Raj C R (2008). Optical sensing of biomedically important polyionic drugs using nano-sized gold particles. *Biosensors and Bioelectronics* 23: 1285–1290. https://doi.org/10.1016/j.bios.2007.11.014

Joung H J, Choi M J, Kim J T, Park S H, Park H J & Shin G H (2016). Development of food-grade curcumin nanoemulsion and its potential application to food beverage system: Antioxidant property and in vitro digestion. *Journal of Food Science* 81(3): N745–N753. https://doi.org/10.1111/1750-3841.13224

Kadam A A, Saratale G D, Ghodake G S, Saratale R G, Shahzad A, Magotra V K, Kumar M, Palem R R & Sung J S (2022). Recent advances in the development of laccase-based biosensors via nano-immobilization techniques. *Chemosensors*, 10(58): 58. https://doi.org/10.3390/chemosensors10020058

Kaneko N, Horii K, Akitomi J, Kato S, Shiratori I & Waga I (2018). An aptamer-based biosensor for direct, label-free detection of melamine in raw milk. *Sensors (Switzerland)* 18: 3227. https://doi.org/10.3390/s18103227

Katouzian I & Jafari S M (2019). Nanotubes of α-lactalbumin for encapsulation of food ingredients. In S M Jafari (Ed.), *Biopolymer Nanostructures for Food Encapsulation Purposes. Nanoencapsulation in the Food Industry* (pp. 101–124). Academic Press. https://doi.org/10.1016/B978-0-12-815663-6.00004-5

Khafoor A A, Karim A S & Sajadi S M (2023). Recent progress in synthesis of nano-based liposomal drug delivery systems: A glance to their medicinal applications. *Results in Surfaces and Interfaces* 11*:* 100124. https://doi.org/10.1016/j.rsurfi.2023.100124

Khan I, Saeed K & Khan I (2019). Nanoparticles: Properties, applications and toxicities. *Arabian Journal of Chemistry* 12: 908–931. https://doi.org/10.1016/j.arabjc.2017.05.011

Khansili N, Rattu G & Krishna P M (2018). Label-free optical biosensors for food and biological sensor applications. *Sensors and Actuators B: Chemical* 65: 35-49. https://doi.org/10.1016/j.snb.2018.03.004

Khorasani S, Danaei M & Mozafari M R (2018). Nanoliposome technology for the food and nutraceutical industries. *Trends in Food Science and Technology* 79. https://doi.org/10.1016/j.tifs.2018.07.009

Kumar T S M, Kumar K S, Rajini N, Siengchin S, Ayrilmis N & Rajulu A V (2019). A comprehensive review of electrospun nanofibers: Food and packaging perspective. *Composites Part B: Engineering* 175*:* 107074. https://doi.org/10.1016/j.compositesb.2019.107074

Kumar D, Farrukh M & Faisal N (2021). Nanocomposites in the food packaging industry: Recent trends and applications. In I Management Association (Ed.), *Research Anthology on Food Waste Reduction and Alternative Diets for Food and Nutrition Security* (pp. 122-146). IGI Global. https://doi.org/10.4018/978-1-7998-5354-1.ch006

Lane K, Derbyshire E, Smith C, Mahadevan K & Li W (2013). Sensory evaluation of a yogurt drink containing an omega-3 nanoemulsion with enhanced bioavailability. *Proceedings of the Nutrition Society*, *72*(OCE2), E99. https://doi.org/10.1017/S0029665113001109

Li Y, Zhang X & Zhang R (2019). Preparation and characterization of electrospun casein/pectin nanofibers for potential application in yogurt. *Journal of Food Science and Technology* 56: 676-682

Li Q, Lv L, Liu Y, Fang Z, Deng Q, Liang W, Wu Y & Chen Z (2023). Preparation, characterization and application of bacteriocin CAMT6 nanoliposomes using resveratrol as a novel stabilizer. *Food Chemistry*, *403*, 134293. https://doi.org/10.1016/j.foodchem.2022.134293

Lin L, Gu Y & Cui H (2019). Moringa oil/chitosan nanoparticles embedded gelatin nanofibers for food packaging against *Listeria monocytogenes* and *Staphylococcus aureus* on cheese. *Food Packaging and Shelf Life* 19: 86-93 https://doi.org/10.1016/j.fpsl.2018.12.005

Liu S, Yuan L, Yue X, Zheng Z & Tang Z (2008). Recent advances in nanosensors for organophosphate pesticide detection. *Advanced Powder Technology* 19: 419–441. https://doi.org/10.1016/S0921-8831(08)60910-3

Liu B, Xu H, Zhao H Y, Liu W, Zhao L Y & Li Y (2017). Preparation and characterization of intelligent starch/PVA films for simultaneous colorimetric indication and antimicrobial activity for food packaging applications. *Carbohydrate Polymers* 157: 842-849.

Liu B, Thum C, Wang Q, Feng C, Li T, Damiani Victorelli F, Li X, Chang R, Chen S, Gong Y & Li Y (2023). The fortification of encapsulated soy isoflavones and texture modification of soy milk by α-lactalbumin nanotubes. *Food Chemistry* 419 pp. https://doi.org/10.1016/j.foodchem.2023.135979

Loewen A, Chan B & Li-Chan E C Y (2018). Optimization of vitamins A and D3 loading in re-assembled casein micelles and effect of loading on stability of vitamin D3 during storage. *Food Chemistry* 240: 472–481. https://doi.org/10.1016/j.foodchem.2017.07.126

Lopez B P & Merkoci A (2011). Nanomaterials based biosensors for food analysis applications. *Trends in Food Science and Technology* 22: 625–639. https://doi.org/10.1016/j.tifs.2011.04.001

Lopes N A & Brandelli A (2018). Nanostructures for delivery of natural antimicrobials in food. *Critical Reviews in Food Science and Nutrition*, 58: 2202–2212. https://doi.org/10.1080/10408398.2017.1308915

Lohith Kumar D H & Sarkar P (2018). Encapsulation of bioactive compounds using nanoemulsions. *Environmental Chemistry Letters, 16*, 59-70. https://doi.org/10.1007/s10311-017-0663-x

Maurya V K & Aggarwal M (2019). Fabrication of nano-structured lipid carrier for encapsulation of vitamin D3 for fortification of 'Lassi'; A milk based beverage. *The Journal of Steroid Biochemistry and Molecular Biology* 193: 105429. https://doi.org/10.1016/j.jsbmb.2019.105429

McClements D J (2010). Design of nano-laminated coatings to control bioavailability of lipophilic food components. *Journal of Food Science* 75(1). https://doi.org/10.1111/j.1750-3841.2009.01452.x

Medeiros A K O C, Gomes C C, Amaral M L Q A, Medeiros L D G, Medeiros I, Porto D L, Aragão C F S, Maciel B L L, Morais A H A & Passos T S (2019). Nanoencapsulation improved water solubility and color stability of carotenoids extracted from Cantaloupe melon (*Cucumis melo L.*). *Food Chemistry* 270: 562–572. https://doi.org/10.1016/j.foodchem.2018.07.099

Miljkovic M G, Nesic A R, Davidovic S Z, Radovanovic N R & Dimitrijevic S I (2017). The use of nanoemulsion-based edible coatings to prolong the shelf-life of cheese. *Journal of International Scientific Publications* 5: 131-138

Mohaisen M J M, Yildirim R M, Yilmaz M T & Durak M Z (2019). Production of functional yogurt drink, apple and orange juice using nano-encapsulated *L. brevis* within sodium alginate-based biopolymers. *Science of Advanced Materials* 11: 1788–1797. https://doi.org/10.1166/sam.2019.3708

Mohammad Z H, Ahmad F, Ibrahim S A & Zaidi S (2022). Application of nanotechnology in different aspects of the food industry. *Discover. Food* 2: 12. https://doi.org/10.1007/s44187-022-00013-9

Mohammed N K, Muhialdin B J & Hussin A S M (2020). Characterization of nanoemulsion of Nigella sativa oil and its application in ice cream. *Food Science and Nutrition* 8(6). https://doi.org/10.1002/fsn3.1500

Mohammadi R, Mahmoudzadeh M, Atefi M, Khosravi-Darani K & Mozafari M R (2015). Applications of nanoliposomes in cheese technology. *International Journal of Dairy Technology* 68: 11–23. https://doi.org/10.1111/1471-0307.12174

Manoj D, Shanmugasundaram S & Anandharamakrishnan C (2021). Nanosensing and nanobiosensing: Concepts, methods, and applications for quality evaluation of liquid foods. *Food Control* 126: 108017. https://doi.org/10.1016/j.foodcont.2021.108017

Montes de Oca-Ávalos J M, Candal R J & Herrera M L (2017). Nanoemulsions: Stability and physical properties. *Current Opinion in Food Science* 16: 1–19. https://doi.org/10.1016/j.cofs.2017.06.003

Maqsoudlou A, Assadpour E, Mohebodini H & Jafari S M (2022). The influence of nanodelivery systems on the antioxidant activity of natural bioactive compounds. *Critical Reviews in Food Science and Nutrition* 62: 3208–3231. https://doi.org/10.1080/10408398.2020.1863907

Nanda S S, Yi D K & Kim K (2016). Graphene oxide based fluorometric detection of hydrogen peroxide in milk. *Journal of Nanoscience and Nanotechnology* 16: 1181–1185. https://doi.org/10.1166/jnn.2016.10646

Nascimento C F, Santos P M, Pereira-Filho E R & Rocha F R P (2017). Recent advances on determination of milk adulterants. *Food Chemistry*, 221: 1232-1244. https://doi.org/10.1016/j.foodchem.2016.11.034

Nickols-Richardson S M & Piehowski K E (2008). Nanotechnology in nutritional sciences. *Minerva Biotecnologica* 20: 117–126.

Nile S H, Baskar V, Selvaraj D, Nile A, Xiao J & Kai G (2020). Nanotechnologies in food science: Applications, recent trends, and future perspectives. *Nano-Micro Letters, 12*, 1-34. https://doi.org/10.1007/s40820-020-0383-9

Omerović N, Djisalov M, Živojević K, Mladenović M, Vunduk J, Milenković I, Knežević N Ž, Gadjanski I & Vidić J (2021). Antimicrobial nanoparticles and biodegradable polymer composites for active food packaging applications. *Comprehensive Reviews in Food Science and Food Safety* 20: 2428–2454. https://doi.org/10.1111/1541-4337.12727

Panghal A, Chhikara N, Anshid V, Sai Charan M V, Surendran V, Malik A & Dhull S B (2019). Nanoemulsions: A promising tool for dairy sector. In R Prasad, V Kumar, M Kumar & D Choudhary (Eds.), *Nanobiotechnology in Bioformulations* (pp. 67-92). Nanotechnology in the Life Sciences. Springer. https://doi.org/10.1007/978-3-030-17061-5_4

Pateiro M, Gómez B, Munekata P E S, Barba F J, Putnik P, Kovačević D B & Lorenzo J M (2021). Nanoencapsulation of promising bioactive compounds to improve their absorption, stability, functionality and the appearance of the final food products. *Molecules (Basel, Switzerland)* 26(6), 1547. https://doi.org/10.3390/molecules26061547

Patel D K, Kim H B, Dutta S D, Ganguly K & Lim K T (2020). Carbon nanotubes-based nanomaterials and their agricultural and biotechnological applications. *Materials* 13(7): 1679. https://doi.org/10.3390/ma13071679

Pinilla C M & Brandelli A (2016). Antimicrobial activity of nanoliposomes co-encapsulating nisin and garlic extract against Gram-positive and Gram-negative bacteria in milk. *Innovative Food Science and Emerging Technologies*, 36: 287–293. https://doi.org/10.1016/j.ifset.2016.07.017

Pinilla C M, Noreña C P & Brandelli A (2017). Development and characterization of phosphatidylcholine nanovesicles, containing garlic extract, with antilisterial activity in milk. *Food Chemistry* 220: 470–476. https://doi.org/10.1016/j.foodchem.2016.10.027

Pu H, Xu Y, Sun D W, Wei Q & Li X (2021). Optical nanosensors for biofilm detection in the food industry: Principles, applications and challenges. *Critical Reviews in Food Science and Nutrition*, 61(13): 2107–2124. https://doi.org/10.1080/10408398.2020.1808877

Pudtikajorn K, Sae-leaw T & Benjakul S (2021). Characterization of fortified pasteurized cow milk with nanoliposome loaded with skipjack tuna eyeball oil. *International Journal of Food Science and Technology* 56(12): 5893–5903. https://doi.org/10.1111/ijfs.15196

Rashidinejad A, Birch E J, Sun-Waterhouse D & Everett D W (2016). Effect of liposomal encapsulation on the recovery and antioxidant properties of green tea catechins incorporated into a hard low-fat cheese following in vitro simulated gastrointestinal digestion. *Food and Bioproducts Processing* 100: 238–245. https://doi.org/10.1016/j.fbp.2016.07.005

Rasti B, Erfanian A & Selamat J (2017). Novel nanoliposomal encapsulated omega-3 fatty acids and their applications in food. *Food Chemistry* 230: 690–696. https://doi.org/10.1016/j.foodchem.2017.03.089

Rezaei A, Fathi M & Jafari S M (2019). Nanoencapsulation of hydrophobic and low-soluble food bioactive compounds within different nanocarriers. *Food Hydrocolloids* 88: 146–162. https://doi.org/10.1016/j.foodhyd.2018.10.003

Rostamabadi H, Assadpour E, Tabarestani H S, Falsafi S R & Jafari S M (2020). Electrospinning approach for nanoencapsulation of bioactive compounds; recent advances and innovations. *Trends in Food Science & Technology* 100: 190–209. https://doi.org/10.1016/j.tifs.2020.04.012

Sanabria L A A (2012). Development of a frozen yogurt fortified with a nano-emulsion containing purple rice bran oil. Louisiana State University, Master's Theses 206. Retrieved from https://repository.lsu.edu/gradschool_theses/206

Saxena D C & Bhardwaj M (2017). Development of organic and inorganic nanoparticles and their subsequent application in nanocomposites for food and non-food packaging systems. *Journal of Nanomedicine and Nanotechnology*. https://doi.org/10.4172/2157-7439-C1-052

Schmidt S E, Holub G, Sturino J M & Taylor T M (2009). Suppression of Listeria monocytogenes scott a in fluid milk by free and liposome-entrapped nisin. *Probiotics and Antimicrobial Proteins* 1(2): 152–158. https://doi.org/10.1007/s12602-009-9022-y

Senthil M K, Kumar K S, Rajini N, Siengchin S, Ayrilmis N & Varada Rajulu A (2019). A comprehensive review of electrospun nanofibers: Food and packaging perspective. *Composites Part B: Engineering* 175*:* 107074. https://doi.org/10.1016/j.compositesb.2019.107074

Sharma C, Dhiman R, Rokana N & Panwar H (2017). Nanotechnology: An untapped resource for food packaging. *Frontiers in Microbiology*, 8: 1735. https://doi.org/10.3389/fmicb.2017.01735

Sharma A, Nagarajan J, Gopalakrishnan K, Bodana V, Singh A, Prabhakar P K, Suhag R & Kumar R (2023). Nanotechnology applications and implications in food industry. In R Pudake, N Chauhan & C Kole (Eds.), *Nanotechnology Applications for Food Safety and Quality Monitoring* (pp. 171–182). https://doi.org/10.1016/B978-0-323-85791-8.00016-1

Silva H D, Cerqueira M Â & Vicente A A (2012). Nanoemulsions for food applications: Development and characterization. *Food Bioprocess Technology* 5(3): 854–867. https://doi.org/10.1007/s11947-011-0683-7

Siyar Z, Motamedzadegan A, Milani J M & Rashidinejad A (2022). The effect of the liposomal encapsulated saffron extract on the physicochemical properties of a functional ricotta cheese. *Molecules (Basel, Switzerland)*, 27(1): 120. https://doi.org/10.3390/molecules27010120

Song L, Zhang L, Huang Y, Chen L, Zhang G, Shen Z, Zhang J, Xiao Z & Chen T (2017). Amplifying the signal of localized surface plasmon resonance sensing for the sensitive detection of *Escherichia coli O157*. *Scientific Reports*, 7: 3288. https://doi.org/10.1038/s41598-017-03495-1

Srinivasan V, Chavan S, Jain U & Tarwadi K (2019). Liposomes for nanodelivery systems in food products. In R Pudake, N Chauhan & C Kole (Eds.), *Nanoscience for Sustainable Agriculture*. Springer. https://doi.org/10.1007/978-3-319-97852-9_24

Stasyuk N Y, Gayda G Z, Zakalskiy A E, Fayura L R, Zakalska O M, Sibirny A A, Nisnevitch M & Gonchar M V (2022). Amperometric biosensors for L-arginine and creatinine assay based on recombinant deiminases and ammonium-sensitive Cu/Zn (Hg)S nanoparticles. *Talanta*, 238(Pt 1), 122996. https://doi.org/10.1016/j.talanta.2021.122996

Syama M A, Arora S, Gupta C, Sharma A & Sharma V (2019). Enhancement of vitamin D2 stability in fortified milk during light exposure and commercial heat treatments by complexation with milk proteins. *Food Bioscience* 29: 17–23. https://doi.org/10.1016/j.fbio.2019.03.005

Tarhan Ö, Hamaker B R & Campanella O H (2021). Structure and binding ability of self-assembled α-lactalbumin protein nanotubular gels. *Biotechnology Progress*, 37(3), e3127. https://doi.org/10.1002/btpr.3127

Todorov S D, Popov I, Weeks R & Chikindas M L (2022). Use of bacteriocins and bacteriocinogenic beneficial organisms in food products: Benefits, challenges, concerns. *Foods (Basel, Switzerland)* 11(9): 3145. https://doi.org/10.3390/foods11193145

Wang J, Zhou Y & Jiang L (2021). Bio-inspired track-etched polymeric nanochannels: Steady-StateBiosensors for Detection of Analytes. *ACS Nano* 15: 18974-19013. https://doi.org/10.1021/acsnano.1c08582

Wang Q, Yu W, Li Z, Liu B, Hu Y, Chen S, de Vries R, Yuan Y, Erazo Quintero L E, Hou G, Hu C & Li Y (2022). The stability and bioavailability of curcumin loaded α-lactalbumin nanocarriers formulated in functional dairy drink. *Food Hydrocolloids*, 131: 107807. https://doi.org/10.1016/j.foodhyd.2022.107807

Wei J N, Zeng X A, Tang T, Jiang Z & Liu Y Y (2018). Unfolding and nanotube formation of ovalbumin induced by pulsed electric field. *Innovative Food Science and Emerging Technologies* 45*:* 249–254. https://doi.org/10.1016/j.ifset.2017.10.011

Xue J, Wu T, Dai Y & Xia Y (2019). Electrospinning and electrospun nanofibers: Methods, materials, and applications. *Chemical Reviews*, 119(9): 5298–5415. https://doi.org/10.1021/acs.chemrev.8b00593

Vasconcelos H, Matias A, Jorge P, Saraiva C, Mendes J, Araújo J, Dias B, Santos P, Almeida J M M M & Coelho L C C (2021). Optical biosensor for the detection of hydrogen peroxide in milk. *Chemistry Proceedings* 5: 55. https://doi.org/10.3390/CSAC2021-10466

van der Hee R M, Miret S, Slettenaar M, Duchateau G S, Rietveld A G, Wilkinson J E, Quail P J, Berry M J, Dainty J R, Teucher B & Fairweather-Tait S J (2009). Calcium absorption from fortified ice cream formulations compared with calcium absorption from milk. *Journal of the American Dietetic Association*, 109(5): 830–835. https://doi.org/10.1016/j.jada.2009.02.017

Yotova L, Yaneva S & Marinkova D (2013). Biomimetic nanosensors for determination of toxic compounds in food and agricultural products. *Journal Chemistry Technology and Metallurgy* 48: 215–227

Zarrabi A, Alipoor Amro Abadi M, Khorasani S, Mohammadabadi M R, Jamshidi A, Torkaman S, Taghavi E, Mozafari M R & Rasti B (2020). Nanoliposomes and tocosomes as multifunctional nanocarriers for the encapsulation of nutraceutical and dietary molecules. *Molecules (Basel, Switzerland)* 25(3): 638. https://doi.org/10.3390/molecules25030638

Zhang Y & Zhong Q (2018). Freeze-dried capsules prepared from emulsions with encapsulated lactase as a potential delivery system to control lactose hydrolysis in milk. *Food Chemistry* 241: 397–402. https://doi.org/10.1016/j.foodchem.2017.09.004

Zhang M, Ahmed A & Xu L (2023). Electrospun Nanofibers for Functional Food Packaging Application. *Materials*, 16(17): 5937. https://doi.org/10.3390/ma16175937

Zhou Y & Kubota L T (2020). Trends in electrochemical sensing. *Chem Electro Chem* 7: 3684–3685. https://doi.org/10.1002/celc.202001025

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Assessment of Mineral Contents and Technological Properties of Dry Bean Genotypes Grown Under Organic Farming Conditions with Multivariate Analysis

Hamdi Ozaktan[a*], Oğuz Erol[b], Satı Uzun[a], Oğuzhan Uzun[c]

[a]*Erciyes University Agricultural Faculty Field Crops Department, Kayseri, TÜRKIYE*

[b]*Yozgat Bozok University, Cannabis Research Institute, Department of Agriculture and Food, Yozgat, TÜRKIYE*

[c]*Erciyes University, Faculty of Agriculture, Department of Soil Science, Kayseri, TÜRKIYE*

ARTICLE INFO
Research Article
Corresponding Author: Hamdi Ozaktan, E-mail: ozaktan_03@hotmail.com
Received: 28 February 2024 / Revised: 26 June 2024 / Accepted: 02 July 2024 / Online: 14 January 2025

Cite this article
Ozaktan H, Erol O, Uzun S, Uzun O (2025). Assessment of Mineral Contents and Technological Properties of Dry Bean Genotypes Grown Under Organic Farming Conditions with Multivariate Analysis. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):12-21. DOI: 10.15832/ankutbd.1444522

ABSTRACT
Beans are an important source of essential minerals such as iron, zinc, calcium, and magnesium, which are crucial in various physiological functions. The mineral contents of beans are vital in ensuring a balanced and healthy diet, as these minerals are involved in bone health and, immune system function. Additionally, the technological properties of beans, including cooking time, water absorption capacity, and swelling capacity, are important in determining their culinary applications and consumer acceptance. The cooking quality and number of seeds destructed after cooking of beans significantly influence their palatability and overall consumer satisfaction. Assessing the technological properties of different bean genotypes grown under organic farming conditions allows researchers to identify genotypes with desirable cooking characteristics and texture, leading to improved consumer acceptance and culinary applications. Organic farming practices aim to produce food without synthetic chemicals, promoting environmental sustainability and ensuring the production of high-quality and nutritious crops. In this research, 20 bean genotypes were grown under organic conditions for two years. To analyze the complex data obtained from the assessment of mineral contents and technological properties of beans, multivariate analysis techniques (correlation, cluster, scatter plot, biplots etc.) are employed. There was a positive relationship between cooking time and Ca mineral. There was a negative relationship between the coefficient of hydration and water absorption capacity and Zn mineral. Positive correlation between Fe, Mn, Cu, Mg, K, P and S elements was observed. Likewise, examining one of the dry weight, dry volume, wet weight, wet volume, water absorption capacity and swelling index values, which are clustered in the same region and have approximately the same axis length, can save time and be consumable. The zinc, sodium, iron, and copper contents in the beans grown under organic conditions were found to be higher than the data reported in the literature.

Keywords: Cooking time, Dry bean, Macro-micro element, Multivariate analysis, Nutrition, Organic farming

1. Introduction

Legumes are an excellent protein source while rich in carbohydrates, resistant starch fiber, potassium, copper, phosphorus, manganese, iron, magnesium and B vitamins (Mullins & Arjmandi 2021). The consumption of legumes such as dry beans and peas increases the intake of dietary fiber, protein, folate, zinc, iron, and magnesium in human nutrition, while reducing saturated fat and total fat intake (Mitchell et al. 2009). Dry beans, due to their high fiber and resistant starch content, elicit a lower glycemic response compared to other high-carbohydrate foods; this may be a factor in the prevention or treatment of diabetes (Ludwig 2002) and colon cancer (Mathers 2002). Additionally, the protein, fiber, and folate in dry beans have been shown to play potential roles in preventing heart disease (Cobiac et al. 1990) and possibly certain types of cancer (Michels et al. 2006). Arunasalam et al. (2004) have also reported that non-nutritive phytochemicals such as saponins, oligosaccharides, and phytate derivatives in dry beans may have a role in cancer prevention. Legumes are an inexpensive and long-term source of protein, micronutrients, essential phytochemicals and complex carbohydrates for a healthier lifestyle. Their composition is suitable for people with diabetes and celiac disease, as well as for consumers concerned about satiety (Singh et al. 2022).

Since deficiencies of potassium (K), phosphorus (P), magnesium (Mg), iron (Fe), zinc (Zn), or copper (Cu) in the human body contribute to the development of various diseases, mineral deficiency is a public health problem affecting thousands of people worldwide (Riberio et al. 2022). Magnesium is an important biological element found in bound form in cells and has many important functions in regulating cellular functions (Jahnen-Dechent & Ketteler 2012). While zinc is crucial for many physiological processes in humans, it also plays essential roles as a regulator or coenzyme of more than 300 enzymes (Schubert et al. 2015). The amount of copper in the body is so perfectly balanced that while a small amount is necessary for many

physiological activities, too much can be life-threatening. As a cofactor, for example, it is important to transfer of electrons to oxygen in the respiratory chain (Husain & Mahmood 2019). Iron is an important nutritional mineral for erythropoiesis, cellular energy metabolism, and immune system development and function. Iron-deficiency anemia is the world's most common nutritional disorder (McLean et al. 2009). In light of this information, according to the World Health Organization (WHO) records, 60% of an individual's daily protein intake should be from plant-based proteins and 40% from animal-based proteins for a quality and balanced diet (Ozaktan et al. 2023b).

Organic farming is often practiced to eliminate the dangers of intensive use of chemical fertilizers and offers satisfactory solutions to such problems (Ozaktan & Doymaz 2022). On the other hand, consumers perceive organic foods as healthier and safer than non-organic ones, which is reflected in the increasing number of organic producers worldwide (Barreto et al. 2021). According to Willer et al. (2024) World Organic Agriculture Statistics and Trends report, 72.3 million hectares of land worldwide are managed organically. The organic food market was estimated at 106.4 billion Euros in 2019, while the European market covers 39% of the global market (Malissiova et al. 2022). While P, S, V, Mn, Co, Ni, As, B, K (2% more), Ca (83% more), Fe and, Zn minerals were higher in dry beans grown under organic farming conditions, Mg (8% more) and Cu (16% more) were determined at higher levels in crops grown under conventional conditions (Rodríguez Madrera et al. 2024). Similarly, Akbaba et al. (2012) reported that Ca, Fe, Mn, P, Zn, K, Mg levels in dry bean grains grown under organic conditions were higher than those grown under conventional conditions, while no difference was detected for Cu and S elements. While significant differences in K, Cu, Mn, Ca, Na and Zn levels were determined between organic and non-organic crops, variations in some elements occur depending on the growing region (Barreto et al. 2021).

Determining the technological properties of beans is extremely important. Especially, cooking times for beans are very important for consumers due to two main reasons, the adequacy of time and the scarcity of cooking fuel (Cichy et al. 2019). The age of the seed, environmental conditions during production, cooking method, and genetics are factors that affect the cooking time of kidney beans (Stanley 1992). Therefore, dry beans, known as the meat of the poor, one of the most widely grown edible legumes worldwide, are not only a food source used in the diets of more than 300 million people, but also a good source of protein, vitamins and mineral (Ozaktan et al. 2023a). The mineral profile of beans can vary significantly among genotypes (Pinheiro et al. 2010). The aim of this study was to determine the mineral contents and technological characteristics of bean genotypes grown under organic conditions and to interpret them by multivariate statistical analysis methods.

2. Material and Methods

2.1 Material

In the experiment, local genotypes named Çomaklı, Gömeç, Kızık and Güzelöz which were intensively grown and consumed in Kayseri province between 1990-2017, together with certified varieties that Akman 98, Adabeyazı, Alberto, Altın, Aras 98, Arslan, Batallı, Berrak, Cihan, Dermason, Göynük 98, Noyanbey 98, Önceler 98, Özdemir, Özmen & Şahin 90 were used. And the term genotype was used for completeness of meaning in the article.

2.2 Methods

The experiment was conducted at Erciyes University research field in Kayseri province, located at an altitude of 1094 m above sea level, between the east longitudes of 34° 56'–36° 59' and north latitudes of 37° 45'–38° 18', without the use of any chemical inputs under organic conditions in 2018 and 2019, following a randomized complete block desing with three replications. The trial consisted of plots with a length of 3 m, row spacing of 45 cm, within-row spacing of 10 cm, and 6 rows. A 1 m gap between blocks and a 0.5 m gap between plots were maintained. After hand sowing, the water requirement of the plants was met through sprinkler irrigation until emergence and the first weed control. Subsequently, a drip irrigation system was installed, with one drip line per row.

2.3 Climate of the study area

Throughout the trial vegetation period (May-September), the average temperature values recorded during the first year were higher than the long-term average, while the monthly average relative humidity values were partially lower than the long-term average. Upon examining the data for the second year, monthly average temperature values were recorded as 17.4-22.3 °C, and monthly average relative humidity values were recorded as 49.1-55.8%. The total precipitation during the vegetation period was 134.2 mm in the first year, 137.5 mm in the second year, with a long-term average of 125.5 mm. In the second year, particularly the rainfall events in June, July, and August extended the vegetation period for beans.

Table 1- The monthly temperature (°C), relative humidity (%), and precipitation (mm) values for the vegetation period in the years 2018, 2019, and 1931-2019*

	Monthly average temperature (°C)			*Monthly average relative humidity (%)*			*Monthly total precipitation (mm)*		
Months	*2018*	*2019*	*1931-2019*	*2018*	*2019*	*1931-2019*	*2018*	*2019*	*1931-2019*
May	*16.7*	17.4	*15.0*	*61.2*	50.2	*61.0*	*51.9*	*23.7*	*51.5*
June	*20.4*	21.3	*19.0*	*56.7*	55.8	*55.8*	*78.8*	*55.2*	*40.2*
July	*24.1*	21.6	*22.2*	*45*	49.1	*49.3*	*0.6*	*35.9*	*10.6*
August	*22.9*	22.3	*22.0*	*42.3*	50.3	*49.1*	-	*12.1*	*8.7*
September	*19.2*	17.4	*17.4*	*45.5*	51.2	*53.7*	*2.9*	*10.6*	*14.5*

*: Kayseri Meteorology Provincial Directorate data

2.4 Soil characteristics of the study area

The experiment was carried out in the area where soil preparation was made after soil cultivation and stone collection in the area where agricultural activities had been carried out for many years (approximately 6-8 years). To determine the initial condition before establishing the trial to represent the trial area, soil samples were collected from different points, representing the trial area, at depths of 0-30 cm. Upon examining the results of the soil analysis for the trial area, it was determined that the trial area soils belong to the sandy loam soil class. The trial was conducted in the same area and with the same trial design in both years. The average values for the soil parameters were determined as follows: available phosphorus 6.785-7.80 kg/ha, pH 7.91-8.0, organic matter 0.30-0.67%, lime content 1.59-2.38%, and electrical conductivity (EC) 0.109-0.117 mmhos/cm.

2.5 Parameters examined and analyzes

Technological traits: Fresh weight, dry weight, water absorption capacity, water absorption index, fresh volume, dry volume, swelling capacity, swelling index and cooking time were determined in accordance with the methods specified in Gulumser et al. (2008); Ozaktan & Doymaz (2022); hydration coefficient was determined in accordance with the method specified in Savage et al. (2001); Ozaktan & Doymaz (2022); bulk density was determined in accordance with the method specified in Singh et al. (2010); Ozaktan & Doymaz (2022).

Mineral contents: Macro (Ca, K, Mg, Na, P, S) and micro element (Cu, Zn, Mn, Fe) contents were determined with the use of Agilent 5800 VDV model ICP-OES device. About 0.5 g ground sample was supplemented with 10 mL Nitric (Merck)+ Perchloric acid (Merck) mixture. Samples were subjected to acid-digestion and mineral composition readings were performed in an ICP-OES spectrometer (Mertens 2005); Ozaktan & Doymaz (2022).

2.6 Statistical analysis

The statistics of the data obtained from the examined parameters were separately conducted for each year. To determine the relationships among the examined characteristics, multivariate analysis methods such as correlation, principal component analysis (PCA) - biplot, and cluster analysis were applied to the average values of the examined parameters over the two years. The specified statistical analyses were performed using the JMP Pro 17 software package.

3. Results and Discussion

The results of the variance analysis for the first year, second year, and combined years of the examined features are presented in Figure 1. The obtained average values and Tukey groups are provided in Figures 2 and 3. The values derived from the averages of the years were evaluated in a biplot. The impact of the examined technological features and mineral content on bean genotypes was found to be statistically significant in both years (Figure 1). However, when the combined years were examined, only the effects of Cu and Zn contents on genotypes were found to be statistically insignificant (Figure 1).

Figure 1- Variance analysis data for mineral content and technological traits

Technological traits												
Years	*Source of variation*	*Dry weight (g)*	*Dry volume (ml)*	*Fresh weight (g)*	*Fresh volume (ml)*	*Water absorption capacity (g/seed)*	*Hydration coefficient (%)*	*Swelling capacity (ml/seed)*	*Swelling index (%)*	*Bulk density (g/ml)*	*Cooking time (min.)*	*Number of seeds destructed after cooking*
2019	*Block*	2.223	4.789*	1.832	0.211	2.166	0.119	0.6145	4.846*	6.463*	15.655**	0.935
	Genotypes	45.515**	52.596**	32.205**	19.797**	34.343**	3.403*	7.3177**	3.232*	6.113**	29.035**	13.171**
2020	*Block*	0.626	0.993	0.861	2.774*	0.645	1.156	3.4796*	3.873*	1.0975	3.376*	5.636*
	Genotypes	23.058**	46.680**	23.703**	12.196**	38.130**	5.386**	2.6030*	7.611**	9.7264**	9.113**	3.9185*
Average of years	*Year (Y)*	10.311*	23.453*	3.783*	3.293*	0.542	7.411	2.981*	5.827**	10.772*	0.163	11.422*
	Genotypes (G)	63.545**	95.547**	53.982**	30.230**	67.778**	6.290**	7.325**	8.093**	13.992**	18.178**	6.010**
	Y x G	2.804*	2.412*	2.958*	1.739*	3.554**	1.732*	3.060*	2.064*	2.164*	12.595**	2.452*
Mineral contents												
2019	*Block*	*Ca*	*Cu*	*Fe*	*K*	*Mg*	*Mn*	*Na*	*P*	*S*	*Zn*	
	Genotypes	0.058	0.128	1.001	0.214	0.158	0.428	0.740	0.210	0.329	0.613	
2020	*Block*	16.257**	7.095**	7.051**	178.145**	12.379**	6.245**	14.533**	125.410**	84.080**	23.328**	
	Genotypes	2.576	0.022	2.914	0.626	1.013	2.054	1.250	0.879	0.872	3.235	
Average of years	*Year (Y)*	202.23**	95397.260**	0.072	1751.301**	553.702**	211.666**	21.564*	2546.175	1813.345**	12.146*	
	Genotypes (G)	18.97**	3.923**	5.210**	110.275**	9.646**	9.148**	10.891**	71.768	42.687**	17.967**	
	Y x G	13.28**	2.386*	6.632**	94.979**	10.335**	4.152**	12.956**	71.180	41.074**	17.967**	

*$P \leq 0.05$, **$P \leq 0.01$

When Figure 2 is examined for technological features, the Güzelöz genotype has the lowest values for the swelling index and unit volume weight parameters, while it has the highest values for parameters such as dry weight, dry volume, fresh weight, fresh volume, water absorption capacity, hydration capacity, and post-cooking disintegrated grain count.

Upon analyzing the average data for the first and second years, the following observations were made: the highest average dry weight of 68.61 g was obtained from the Güzelöz genotype, while the lowest value of 24.40 g was observed in the Alberto genotype; the highest average dry volume of 77.33 mL was obtained from the Güzelöz genotype, and the lowest dry volume of 19.67 mL was observed in the Berrak genotype; the highest average fresh weight of 151.96 g was observed in the Güzelöz genotype, and the lowest value of 52.33 g was found in the Alberto genotype; the highest average fresh volume of 137.33 mL was obtained from the Güzelöz genotype, while the lowest value of 46.67 mL was observed in the Alberto genotype; the highest average water absorption capacity of 0.904 g/grain was found in the Güzelöz genotype, and the lowest value of 0.229 g/grain was observed in the Alberto genotype; the highest average hydration index of 1.318% was observed in the Güzelöz genotype, and the lowest value of 0.951% was found in the Özdemir genotype; the highest average hydration coefficient of 121.2% was obtained from the Güzelöz genotype, while the lowest value of 84.4% was observed in the Özdemir genotype; the highest average swelling capacity of 0.792 mL/grain was observed in the Güzelöz genotype, and the lowest value of 0.345 mL/grain was found in the Özmen genotype; the highest average swelling index of 3.11% was observed in the Noyanbey 98 genotype, while the lowest value of 1.60% was found in the Güzelöz genotype; the highest average unit volume weight of 1.566 g/mL was observed in the Berrak genotype, and the lowest value of 0.817 g/mL was found in the Güzelöz genotype; the highest average cooking time of 47.33 minutes was observed in the Alberto genotype, while the lowest value of 27.33 minutes was found in the Aras 98 genotype; the highest average post-cooking disintegrated grain count of 26.667 was observed in the Noyanbey genotype, while the Önceler 98 variety showed no disintegration in both years.

In bean research, dry volume values have been reported by Yeken et al. (2019) as 60-100 mL, Çalışkan et al. (2018) as 43.8-56 mL, Ercan et al. (1994) as 13.0-36.0 mL, and Elkoca & Çınar (2015) as 101.3-17.3 mL. Çalışkan et al. (2018) reported water absorption capacity in beans as 0.51-0.25 g/grain. The values reported by Yalçın et al. (2018) and Kaya et al. (2016) are consistent with the results of this study. Swelling capacity in beans has been reported by Wani et al. (2017) as 0.09-0.28 mL/seed, Yeken et al. (2019) as 0.204-0.850 mL/grain, Çalışkan et al. (2018) as 0.87-0.45 mL/grain, Ercan et al. (1994) as 0.165-0.478 mL/grain, and Nadeem et al. (2020) as 0.10-1.445 mL/grain. Swelling index values in beans, as stated by Yeken et al. (2019), ranging from %1.692-9.00, show similarity. The values obtained in this research are higher than early stduies reported by Nadeem et al. (2020), Wani et al. (2017), and Çalışkan et al. (2018). The differences in technological traits could be attributed to using of different materials and climatic conditions of experimental area. Cooking times in beans have been reported by Wani et al. (2017) as 38.67-86.67 minutes, Yeken et al. (2019) as 26-100 minutes, Çalışkan et al. (2018) as 44-67.5 minutes, Ercan et al. (1994) as 22-36 minutes, and Nadeem et al. (2020) as 46.469-218.222 minutes. Özpekmez (2015) reported post-cooking disintegration degrees in beans as %0.33–12.00. Furthermore, the results obtained from the examined parameters align with the findings of many researchers (Yeken et al. (2019), Çalışkan et al. (2018), Biçer et al. (2017), Kaya et al. (2016), Ceyhan (2010), Öztaş et al. (2007), Ercan et al. (1994).

When examining the average data for the mineral content of bean genotypes grown under organic conditions for the first and second years (Figure 3), the calcium content reaches its highest value in the first year at 1914.48 ppm, obtained from the Noyanbey 98 variety. Following closely in the same statistical group are the Berrak with 1833.6 ppm, Dermason with 1809.6 ppm, Çomaklı with 1782.3 ppm, Batallı with 1613.0 ppm, and Aras 98 with 1592.0 ppm. The lowest value is observed in the

second year at 811.9 ppm, obtained from the Özdemir variety. Studies on calcium values in beans by various researchers report values such as 3210-3771 ppm (Al-Numair et al. 2009), 970-1600 ppm (Pedrosa et al. 2015), 3800-9100 ppm (Kajiwara et al. 2021), 610-960 ppm (de Oliveira et al. 2018), 881-1162 ppm (Ramírez-Ojeda et al. 2018), 1320-1870 ppm (Ribeiro & Kläsener 2020), 826-1650 ppm (Wang et al. 2010), and 610-960 ppm (Ferreira et al. 2014). The results obtained in this study show similarity with the literature.

The highest copper content, at 22.74 ppm, is obtained from the Arslan variety in the second year, while the lowest value, at 3.74 ppm, is obtained from the Şahin 90 variety in the first year. The Noyanbey 98 variety with 3.92 ppm, Göynük 98 with 4.08 ppm, and Çomaklı with 4.41 ppm follow closely and are statistically in the same group. Studies on copper values in beans by various researchers report values such as 7.24 ppm (Akinyele & Shokunbi 2015), 6.8-5.2 ppm (Al-Numair et al. 2009), 3.1-14 ppm (de Oliveira et al. 2018), 5.5-9.9 ppm (Ramírez-Ojeda et al. 2018), 6.35-8.87 ppm (Ribeiro & Kläsener 2020), 2.9-9.9 ppm (Wang et al. 2010), and 3.1-14 ppm (Ferreira et al. 2014). While the first-year results from the experiment are in line with the literature, some second-year averages appear slightly higher compared to the mentioned sources.

When examining the averages for iron content, the highest value, at 111.99 ppm, is obtained from the Akman 98 variety in the first year, while the lowest value, at 24.55 ppm, is obtained from the Şahin 90 variety in the first year. Studies on iron values in beans by various researchers report values such as 48.75 ppm (Akinyele & Shokunbi 2015), 106.7-86.0 ppm (Al-Numair et al. 2009), 60.32-60.25 ppm (Pedrosa et al. 2015), 56-101 ppm (de Oliveira et al. 2018), 42-48 ppm (Ramírez-Ojeda et al. 2018), 53.52-65.80 ppm (Ribeiro & Kläsener 2020), 54.1-67.3 ppm (Wang et al. 2010), and 56-94 ppm (Ferreira et al. 2014). The results obtained are consistent with the literature.

The highest potassium content, at 13258 ppm, is obtained from the Akman 98 variety in the first year, followed closely by the Gömeç variety with 13137 ppm, and statistically, they are in the same group. The lowest value, at 7206 ppm, is obtained from the Adabeyazı variety in the second year. Studies on potassium values in beans by various researchers report values such as 13190-11510 ppm (Al-Numair et al. 2009), 13500-16750 ppm (Kajiwara et al. 2021), 10030-11340 ppm (Ribeiro & Kläsener 2020), 15050-12250 ppm (Wang et al. 2010), and 7400-9700 ppm (Ferreira et al. 2014). Our results are consistent with these studies.

The highest magnesium content, at 1700.8 ppm, is obtained from the Gömeç variety in the first year, followed by the Akman 98 variety with 1665.9 ppm. The lowest value, at 748.7 ppm, is obtained from the Aras 98 variety in the second year. Studies on magnesium values in beans by various researchers report values such as 3110-2890 ppm (Al-Numair et al. 2009), 1460-1040 ppm (Pedrosa et al. 2015), 1190-1770 ppm (Kajiwara et al. 2021), 1092-1289 ppm (Ramírez-Ojeda et al. 2018), 2200-2420 ppm (Ribeiro & Kläsener 2020), 1430-1995 ppm (Wang et al. 2010), and 530-700 ppm (Ferreira et al. 2014). The findings are in line with the literature.

The highest manganese content, at 17.71 ppm, is obtained from the Akman 98 variety in the first year, while the Dermason & Gömeç varieties have statistically higher average manganese content. The lowest value, at 5.80 ppm, is obtained from the Cihan variety in the second year. Studies on manganese values in beans by various researchers report values such as 13.58 ppm (Akinyele & Shokunbi 2015), 22.2-28.8 ppm (Al-Numair et al. 2009), 11-15 ppm (Ramírez-Ojeda et al. 2018), and 13.6-19.2 ppm (Wang et al. 2010). Our findings are in line with the literature.

When looking at the sodium content, the highest value, at 686.4 ppm, is observed in the Kızık variety in the second year, while the lowest value, at 189.8 ppm, is found in the Gömeç variety in the first year. Studies on sodium values in beans by various researchers report values such as 230-240 ppm (Al-Numair et al. 2009) and 200 ppm (Pedrosa et al. 2015).

The highest phosphorus content, at 4833 ppm, is obtained from the Gömeç variety in the first year, while the lowest value, at 1978 ppm, is obtained from the Aras 98 variety in the second year. Studies on phosphorus values in beans by various researchers report values such as 1850-2100 ppm (Al-Numair et al. 2009), 1320-1870 ppm (Ribeiro & Kläsener 2020), 3623-5020 ppm (Wang et al. 2010), and 3100-3700 ppm (Ferreira et al. 2014). Our results are consistent with the literature.

When examining the average values for sulfur content, the highest value, at 2673.7 ppm, is obtained from the Akman 98 variety in the first year, while the lowest value, at 993.8 ppm, is obtained from the Aras 98 variety in the second year. Studies on sulfur values in beans by de Oliveira et al. (2018) report values ranging from 1000 to 2300 ppm, and Ferreira et al. (2014) report values ranging from 1700 to 2000 ppm. Our results show similarity with these findings.

When examining the average values for zinc content in bean varieties, the highest value, at 60.63 ppm, is obtained from the Özdemir variety in the first year, while the lowest value, at 20.74 ppm, is obtained from the Aras 98 variety in the second year. Studies on zinc values in beans by Akinyele and Shokunbi (2015) report a value of 36.14 ppm, Al-Numair et al. (2009) report a range of 24.6-27.5 ppm, Pedrosa et al. (2015) report a range of 21.24-21.60 ppm, de Oliveira et al. (2018) report a range of 33-58 ppm, Ramírez-Ojeda et al. (2018) report a range of 21-37 ppm, Wang et al. (2010) report a range of 25.3-30.4 ppm, and Ferreira et al. (2014) report a range of 33-58 ppm. Our obtained results show similarity with the literature.

Figure 2- Average data and Tukey groups for the technological traits of dry bean genotypes

		Genotypes																			
Traits	Years	Adabeyazı	Akman 98	Alberto	Altın	Aras 98	Arslan	Batallı	Berrak	Cihan	Çomaklı	Dermason	Gömeç	Göynük 98	Güzelöz	Kızık	Noyanbey 98	Önceler 98	Özdemir	Özmen	Şahin
Dry weight (g)	2019	33.03 e-h	35.12 efg	24.40 ı	35.81 d-g	36.77 def	35.29 efg	32.87 e-h	26.96 hı	43.39 cd	28.40 ghı	45.01 bc	30.91 f-ı	36.79 def	68.61 a	30.89 f-ı	40.01 cde	33.68 e-h	52.43 b	28.77 ghı	32.49 e-h
	2020	39.33 c-h	36.39 c-ı	34.83 d-ı	36.35 c-ı	38.05 c-h	35.55 c-ı	34.17 e-ı	31.40 ghı	41.53 b-e	30.72 hı	43.56 bcd	32.08 ghı	43.93 bc	63.25 a	32.73 f-ı	41.20 c-f	34.72 e-ı	50.29 b	28.68 ı	39.61 c-g
Dry volume (ml)	2019	24.0 efg	24.0 efg	16.0 h	26.0 def	25.3 ef	24.7 efg	24.0 efg	17.3 gh	33.3 cd	18.7 fgh	36.0 bc	20.7 e-h	25.3 ef	64.0 a	22.0 e-h	27.3 de	26.0 def	41.3 b	20.0 e-h	23.3 e-h
	2020	28.67 c-f	27.33 def	26.67 def	33.33 cd	28.00 c-f	27.33 def	22.67 ef	22.00 ef	33.33 cd	22.67 ef	37.33 bc	21.33 f	32.67 cd	77.33 a	24.00 def	31.33 cde	26.67 def	46.67 b	20.67 f	29.33 c-f
Fresh weight (g)	2019	66.60 e-ı	72.29 d-ı	52.33 ı	73.91 c-h	77.04 b-g	69.97 e-ı	67.08 e-ı	56.33 ghı	93.97 bc	56.12 ghı	92.63 bcd	59.57 f-ı	77.51 b-f	151.96 a	60.69 f-ı	87.16 b-e	69.67 e-ı	96.85 b	55.56 hı	69.68 e-ı
	2020	79.85 b-f	72.73 c-h	74.83 c-h	76.63 c-g	79.81 b-f	68.93 d-h	68.51 d-h	58.53 gh	85.72 bcd	61.95 fgh	82.99 b-e	63.65 fgh	90.84 bc	132.83 a	64.91 e-h	86.04 bcd	73.81 c-h	98.17 b	56.53 h	83.91 bcd
Fresh volume (ml)	2019	57.33 ef	66.67b-f	46.67 f	73.33 b-e	70.67 b-f	68.00 b-f	64.00 c-f	50.67 ef	85.33 bcd	51.33 ef	88.00 bc	57.33 ef	72.67 b-e	137.33 a	62.67 def	84.67 bcd	68.00 b-f	89.33 b	54.67 ef	66.67 b-f
	2020	76.00 b-f	68.00 b-f	70.00 b-f	73.33 b-f	76.67 b-e	68.00 b-f	64.67 c-f	52.67 ef	85.33 bc	60.67 def	78.67 bcd	65.33 c-f	85.33 bc	124.00 a	62.67 c-f	83.33 bcd	69.33 b-f	92.00 b	52.00 f	70.67 b-f
Water absorption capacity (g/seed)	2019	0.417 c-h	0.399 e-h	0.299 h	0.426 c-g	0.438 c-f	0.413 d-h	0.382 e-h	0.306 gh	0.536 bc	0.335 fgh	0.526 bcd	0.346 fgh	0.418 c-h	0.904 a	0.380 e-h	0.478 b-e	0.422 c-g	0.584 b	0.329 fgh	0.393 e-h
	2020	0.445 b-f	0.389 d-h	0.410 c-g	0.414 c-g	0.447 b-f	0.379 e-h	0.384 d-h	0.351 gh	0.442 b-f	0.361 fgh	0.527 b	0.344 gh	0.469 bcd	0.804 a	0.367 fgh	0.462 b-e	0.391 c-h	0.479 bc	0.307 h	0.443 b-f
Hydration coefficient (%)	2019	101.3 abc	106.2 abc	114.4 ab	106.2 abc	109.5 abc	98.3 abc	103.9 abc	108.8 abc	117.1 ab	97.5 abc	106.1 abc	92.7 bc	110.6 abc	121.2 a	96.4 abc	117.9 ab	106.9 abc	84.4 c	93.1 bc	114.5 ab
	2020	102.7 a-e	99.9 a-e	114.7 a	110.8 abc	109.5 abc	93.7 cde	100.2 a-e	86.3 e	106.6 a-d	101.6 a-e	90.3 de	98.4 a-e	106.8 a-d	110.2 abc	98.8 a-e	109.3 abc	112.6 ab	95.1 b-d	97.1 a-e	112.2 abc
Swelling capacity (ml/seed)	2019	0.416 cde	0.457 b-e	0.328 e	0.531 b-e	0.493 b-e	0.517 b-e	0.447 b-e	0.347 de	0.548 bcd	0.397 cde	0.576 bc	0.443 b-e	0.486 b-e	0.792 a	0.521 b-e	0.582 bc	0.492 b-e	0.631 ab	0.427 b-e	0.458 b-e
	2020	0.521 ab	0.435 ab	0.445 ab	0.411 ab	0.521 ab	0.464 ab	0.469 ab	0.397 ab	0.520 ab	0.439 ab	0.551 a	0.485 ab	0.527 ab	0.538 a	0.441 ab	0.535 a	0.427 ab	0.453 ab	0.345 b	0.413 ab
Swelling index (%)	2019	2.41 ab	2.83 ab	2.92 a	2.84 ab	2.79 ab	2.76 ab	2.67 ab	2.93 a	2.57 ab	2.77 ab	2.44 ab	2.78 ab	2.88 ab	2.15 b	2.85 ab	3.11 a	2.62 ab	2.16 b	2.74 ab	2.86 ab
	2020	2.65 a-d	2.49 a-e	2.65 a-d	2.20 c-f	2.75 abc	2.50 a-e	2.86 ab	2.40 b-e	2.56 a-e	2.69 a-d	2.10 def	3.06 a	2.61 a-d	1.60 f	2.65 a-d	2.68 a-d	2.61 a-e	1.97 ef	2.55 a-e	2.36 b-e
Bulk density (g/ml)	2019	1.377 abc	1.483 abc	1.525 ab	1.389 abc	1.457 abc	1.431 abc	1.369 abc	1.566 a	1.302 bcd	1.533 ab	1.250 cd	1.497 abc	1.457 abc	1.072 d	1.410 abc	1.469 abc	1.297 bcd	1.269cd	1.446 abc	1.394 abc
	2020	1.373 ab	1.333 a-d	1.320 a-d	1.093 cde	1.370 abc	1.307 a-d	1.513 a	1.437 ab	1.247 a-d	1.367 abc	1.167 bcd	1.510 a	1.343 a-d	0.817 e	1.380 ab	1.320 a-d	1.307 a-d	1.080 de	1.407 ab	1.367 abc
Cooking time (min.)	2019	33.00 abc	42.67 g-j	47.33 j	35.33 b-ı	34.67 b-ı	34.33 b-ı	45.67 ıj	43.67 hıj	38.33 d-g	43.67 hıj	42.00 f-ı	33.67 bcd	39.00 e-h	36.67 cde	38.33 d-g	39.00 e-h	28.00 a	37.00 c-f	33.00 abc	30.33 ab
	2020	44.00 e	39.67 bcd	36.33 bcd	40.33 de	27.33 a	39.00 de	42.33 de	40.00 de	30.67 abc	39.00 de	42.00 de	35.67 bcd	29.00 ab	38.67 de	40.00 de	36.67 cde	36.00 bcd	36.00 bcd	37.00 cde	37.67 cde
Number of seeds destructed after cooking	2019	6.333 de	2.333 abc	0.000 a	4.333 bcd	0.000 a	4.667 cde	1.000 ab	0.333 a	4.667 cde	0.000 a	0.000 a	1.333 abc	0.333 a	1.667 abc	0.667 a	4.667 cde	0.000 a	0.333 a	2.000 abc	8.000 e
	2020	12.000 abc	2.667 ab	2.667 ab	9.333 abc	16.000 abc	14.667 abc	6.667 ab	1.333 a	12.000 abc	6.667 ab	4.000 ab	12.000 abc	4.000 ab	4.000 ab	9.333 abc	26.667 c	0.000 a	0.000 a	6.667 ab	21.333 bc

Figure 3- Average data and Tukey groups for the mineral contents of dry bean genotypes

		Genotypes																			
Element	Years	Adabeyazı	Akman 98	Alberto	Altın	Aras 98	Arslan	Batallı	Berrak	Cihan	Çomaklı	Dermason	Gömeç	Göynük 98	Güzelöz	Kızık	Noyanbey 98	Önceler 98	Özdemir	Özmen	Şahin
Ca (ppm)	2019	1413.8 cde	1474.9 b-e	1505.6 b-e	1328.6 de	1592.0 a-e	1290.1 de	1613.0 a-d	1833.6 ab	1392.3 de	1782.3 abc	1809.6 ab	1312.2 de	1358.3 de	1276.5 de	1242.3 de	1914.5 a	847.1 fg	811.9 g	1505.6 b-e	1215.0 ef
	2020	1157.9 d-h	1247.4 b-f	1161.8 d-h	1299.6 a-e	1002.8 ghı	1509.8 a	1227.6 c-g	1193.3 c-g	1003.3 ghı	1048.5 f-ı	1158.1 d-h	1240.1 b-g	952.4 hı	1416.2 abc	1136.3 e-h	1384.4 a-d	840.8 ı	1114.2 e-h	1470.6 ab	1313.2 a-e
Cu (ppm)	2019	5.63 bc	8.59 a	5.76 bc	5.43 bc	4.60 bc	5.32 bc	6.79 ab	4.85 bc	5.00 bc	4.41 c	5.35 bc	6.75 ab	4.08 c	5.27 bc	5.26 bc	3.92 c	5.17 bc	5.06 bc	4.64 bc	3.74 c
	2020	18.00 ab	21.30 ab	19.41 ab	19.58 ab	17.14 b	22.74 a	19.61 ab	22.11 ab	19.87 ab	19.59 ab	17.93 ab	19.52 ab	17.79 ab	21.77 ab	21.69 ab	19.04 ab	20.99 ab	21.23 ab	18.50 ab	19.49 ab
Fe (ppm)	2019	27.38 cd	111.9 a	42.97 bcd	30.91 cd	85.48 ab	30.50 cd	61.36 bcd	58.62 bcd	46.95 bcd	43.32 bcd	38.56 cd	67.30 bc	34.19 cd	44.71 bcd	39.53 cd	39.05 cd	40.07 cd	37.98 cd	45.99 bcd	24.55 d
	2020	34.41 b	45.52 b	47.42 b	38.52 b	38.97 b	83.34 a	45.61 b	50.37 b	36.25 b	40.02 b	41.04 b	43.49 b	39.22 b	54.79 ab	48.03 b	64.59 ab	48.52 b	37.08 b	49.30 b	49.73 b
K (ppm)	2019	8912 fgh	13258 a	8745 ghı	8772 ghı	9036 efg	8350 hıj	8941 fg	9027 efg	9571 cde	9409 def	11494 b	13137 a	8327 hıj	9785 cd	9683 cd	9282 d-g	10063 c	8038 j	9665 cd	8365 ıj
	2020	7206 g	8488 b-e	8603 bcd	8533 b-e	7439 fg	9074 b	8795 bcd	8287 de	8275 de	8389 cde	8827 bcd	8369 cde	8538 b-e	10285 a	8694 bcd	8700 bcd	9929 a	8852 bcd	8930 bc	7964 ef
Mg (ppm)	2019	1094.8 bc	1665.9 a	1143.3 bc	1031.7 bc	1061.1 bc	1010.9 bc	1191.2 bc	1029.0 bc	1007.8 bc	974.2 bc	1302.0 b	1700.8 a	936.6 c	1140.9 bc	1107.5 bc	1022.5 bc	1169.9 bc	896.9 c	1018.8 bc	869.9 c
	2020	797.7 bc	836.2 abc	899.3 abc	959.6 ab	748.7 c	1030.1 a	933.5 abc	787.7 bc	821.1 bc	861.7 abc	866.4 abc	846.0 abc	786.0 bc	926.1 abc	836.3 abc	866.6 abc	1035.5 a	915.7 abc	890.2 abc	874.0 abc
Mn (ppm)	2019	9.21 bcd	17.71 a	11.32 bcd	9.27 bcd	9.55 bcd	9.00 cd	11.69 bcd	11.80 bcd	9.89 bcd	10.72 bcd	12.91 abc	14.21 ab	8.74 cd	11.44 bcd	11.82 bcd	11.25 bcd	9.14 bcd	6.93 d	11.81 bcd	8.43 cd
	2020	7.55 c-g	8.64 a-e	8.73 a-e	7.63 b-g	6.21 fg	9.88 a	8.90 a-e	9.33 abc	5.80 g	9.11 a-d	8.27 a-e	9.10 a-d	6.90 efg	9.23 abc	8.65 a-e	7.86 a-f	7.08 d-g	6.96 efg	9.61 ab	7.67 b-g
Na (ppm)	2019	308.0 bc	555.1 a	350.9 bc	193.1 c	303.3 bc	313.3 bc	320.2 bc	228.9 c	266.2 c	234.7 c	229.5 c	189.8 c	262.0 c	196.9 c	227.6 c	203.8 c	495.0 ab	588.3 a	574.6 a	584.4 a
	2020	364.5 bc	453.9 bc	369.2 bc	302.8 c	348.6 bc	388.2 bc	314.8 bc	362.1 bc	403.5 bc	310.5 bc	408.1 bc	325.1 bc	317.6 bc	466.9 bc	686.4 a	354.0 bc	479.0 b	314.8 bc	329.8 bc	317.9 bc
P (ppm)	2019	3146 c-f	4447 b	2756 g-j	3147 c-f	2966 efg	3057 def	2737 g-j	2898 f-ı	3127 c-f	2930 e-h	3349 c	4833 a	2313 k	3317 cd	2970 efg	2516 jk	3118 c-f	2617 ıj	3187 cde	2661 hıj
	2020	2211 ef	2179 ef	2548 bcd	2687 abc	1978 f	2844 a	2433 cde	2507 bcd	2275 de	2358 de	2365 de	2464 cde	2322 de	2981 a	2338 de	2259 def	2785 ab	2543 bcd	2796 ab	2384 de
S (ppm)	2019	1416.9 h-k	2673.7 a	1388.3 h-k	1634.7 fgh	1622.4 fgh	1364.7 ıjk	2354.9 b	2239.6 bc	1964.5 de	1018.61	2030.2 cd	2289.1 bc	1186.4 kl	1734.6 ef	1451.0 g-j	1447.4 g-k	1239.6 jkl	1460.1 g-j	1680.1 fg	1229.5 jkl
	2020	1026.9 bc	1162.7 abc	1161.2 abc	1149.6 abc	993.8 c	1248.4 abc	1159.6 abc	1182.2 abc	1051.0 abc	1147.9 abc	1174.3 abc	1134.3 abc	1069.4 abc	1236.5 abc	1059.5 abc	1206.8 abc	1272.4 ab	1033.4	1300.5 a	1153.1 abc
Zn (ppm)	2019	27.12 efg	40.98 bcd	26.08 efg	24.88 efg	27.51 ef	24.05 efg	42.48 bc	26.34 efg	31.64 cde	16.02 g	30.43 def	46.79 b	19.05 fg	25.61 efg	28.37 ef	23.10 efg	26.08 efg	60.63 a	27.37 efg	28.44 ef
	2020	23.84 bc	26.76 abc	27.47 abc	26.37 abc	20.74 c	31.71 ab	27.98 abc	30.25 ab	25.48 bc	23.70 bc	29.64 ab	23.83 bc	23.29 bc	34.45 a	29.80 ab	26.71 abc	34.23 a	26.53 abc	30.75 ab	28.30 abc

The scatterplot and correlation analysis outputs for the examined parameters are presented in Figure 4. Red circles indicate a positive relationship between parameters, while blue circles represent a negative relationship. The size of the circle reflects the strength of the relationship. Additionally, the scatterplot analysis output shows five different-colored scatterings, representing five different groups obtained from cluster analysis. The fitted line in the analysis also illustrates the relationship between two parameters. Among the technological parameters, both positive and negative high-level relationships were observed. Similar relationships were also observed among mineral content parameters. When examining the relationships between technological and mineral content parameters, especially for the cooking time, positive correlations were found with Ca content (r= 0.538), positive relationships with Cu, Fe, Mn, and S elements, and a negative relationship with Na element. Negative relationships were observed between the hydration coefficient (r= -0.352) and water absorption index (r= -0.310) with Zn content. Positive

relationships were observed between bulk density and Ca, Fe, and Mn contents. Swelling index showed positive relationships with Ca and Mn contents, and a negative relationship with Zn. Negative relationships were also observed between dry weight and dry volume with Ca content. The findings obtained are consistent with the results found by many researchers (Rodríguez Madrera et al. 2024; Ozaktan 2021; Ozaktan & Doymaz 2022).

Principal component analysis (PCA) is a versatile statistical method used to reduce the state-by-variables data table to its basic characteristics called principal components, which are several linear combinations of the original variables that maximally explain the variance of all variables (Greenacre et al. 2022). The biplot analysis output for the technological characteristics and mineral content of 20 different bean genotypes is presented in Figure 5. The colored circles in Figure 5 represent the positions of genotypes on the biplot resulting from the cluster analysis of the examined parameters. In PCA analysis, the lengths of the axes, their positions, and the angles with other axes express relationships. In the biplot analysis obtained from the two-year averages, 22 independent principal components were identified, with PC1 and PC2 values being 32.1% and 24.5%, respectively. In PC1, descriptors with high impact values were dry weight, dry volume, fresh weight, fresh volume, water absorption capacity, and swelling capacity; in PC2, Mg, K, S, P, Cu, Mn, and Fe were the descriptors with the highest impact values (Table 2). The axes of fresh weight, fresh volume, dry weight, dry volume, water absorption capacity, and swelling capacity were positioned in the same region, forming low-degree angles among them, and the sufficiently long axes indicated a high positive relationship between these parameters. Findings obtained by is compatible with the literature (Ozaktan 2021; Ozaktan & Doymaz 2022). Kibar (2019) reported that the first two PCs explained 58.45% of the total variance in principal component analysis for biochemical properties (mineral, ash, protein, pH) of wheat. He reported that the first principal component (PC1), accounted for 35.72% of the variance and was associated with K, P, Ca, Mn, Co, As, Pb, Ni, Cr, protein and pH contents and the second principal component (PC2) accounted for 22.73% of the mineral composition consisting of Na, Fe, Sn and Cd contents. Additionally, the genotype Güzelöz, located within the pink-lined circle, had the highest values for these parameters, followed by the Özdemir genotype. When examining the position, length, and angles between the axes formed by Mg, K, S, P, Cu, Mn, and Fe parameters, a high positive relationship between them was observed. Furthermore, the genotypes Gömeç and Akman were identified as leading genotypes with the highest values for these mineral contents. Positive relationships were observed between the Ca mineral and cooking time, with the Batallı genotype having the longest cooking time among the genotypes located in the same region and approximately the same axis length. Analyzing the biplot output of the cluster analysis on the examined parameters to determine the relationship between genotypes, it is observed that Güzelöz, Akman 98, and Gömeç varieties are distinct from the other varieties.

Table 2- Eigen values, percent and cum percent, for investigated characteristics

Number	*Eigenvalue*	*Percent*	*Cum Percent*
1	7.058311	32.083	32.083
2	5.394635	24.521	56.604
3	2.535145	11.523	68.128
4	1.668333	7.583	75.711
5	1.056192	4.801	80.512
6	0.931535	4.234	84.746
7	0.841928	3.827	88.573
8	0.641969	2.918	91.491
9	0.423336	1.924	93.415
10	0.364414	1.656	95.072
11	0.287439	1.307	96.378
12	0.239557	1.089	97.467
13	0.211780	0.963	98.430
14	0.165151	0.751	99.181
15	0.079363	0.361	99.541
16	0.052090	0.237	99.778
17	0.034666	0.158	99.936
18	0.009674	0.044	99.980
19	0.002581	0.012	99.991
20	0.001115	0.005	99.996
21	0.000724	0.003	100.000
22	0.000062	0.000	100.000

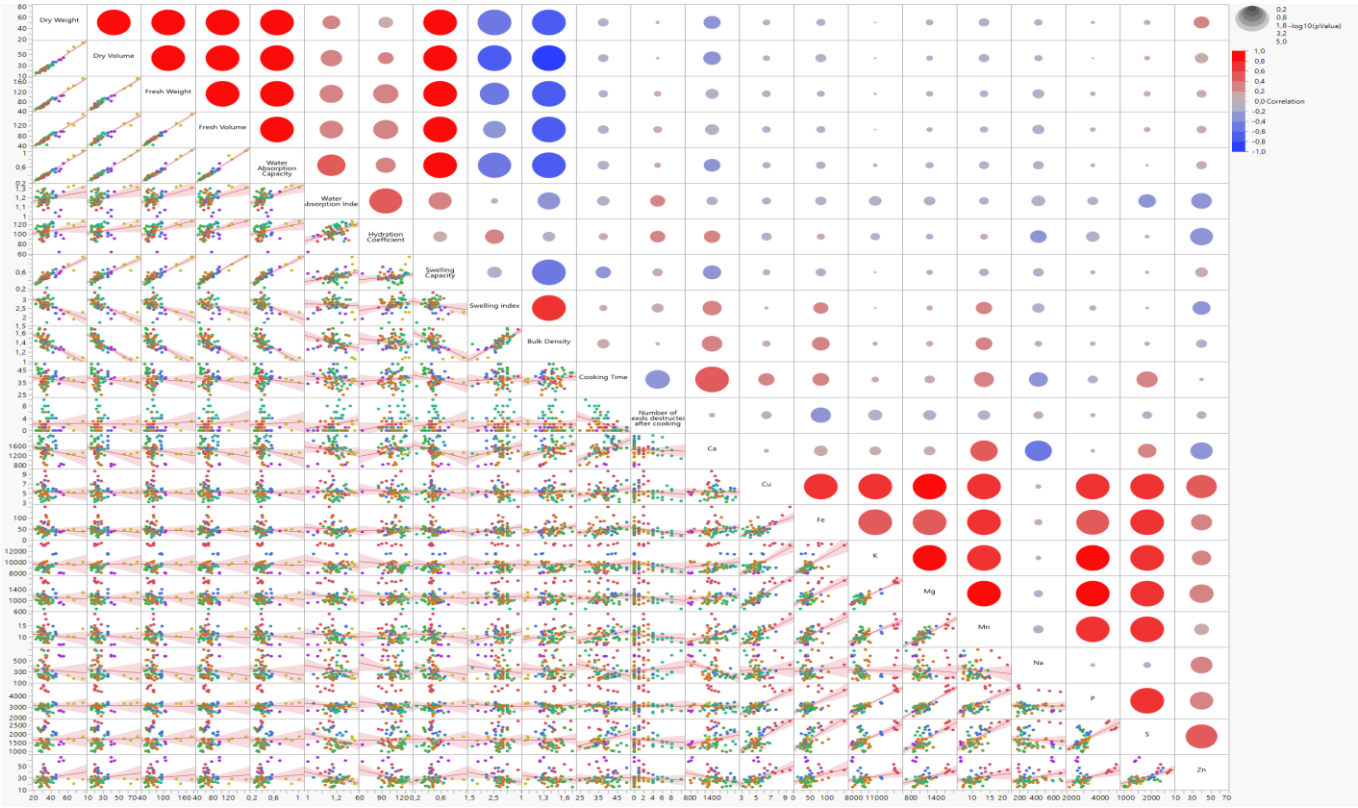


Figure 4- Scatterplot matrix for overview of correlations and fit lines

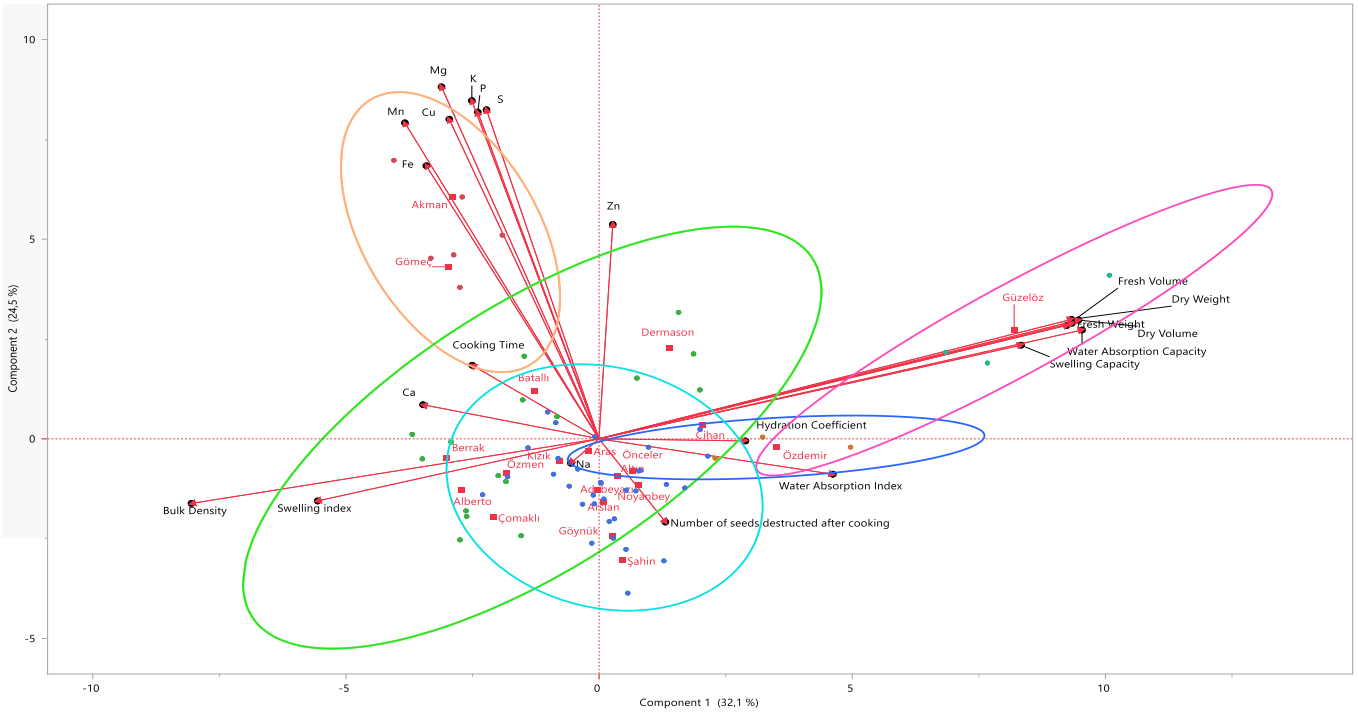


Figure 5- The groups formed by biplot analysis on dry bean genotypes and the examined parameters

4. Conclusions

In this study, which evaluated the mineral composition and technological characteristics of 20 bean genotypes grown under organic conditions, genotypes were clustered into 5 different groups. Among the genotypes, Güzelöz differed from the other genotypes in terms of the analyzed characteristics. There was a positive relationship between cooking time and Ca mineral. There was a negative relationship between the coefficient of hydration and water absorption capacity and Zn mineral. A positive

correlation was found between each of the elements Fe, Mn, Cu, Mg, K, P and S with all other elements. Especially when examining mineral matter in beans, reading one of the elements K, P and S can give us information about the other two elements. Likewise, examining one of the dry weight, dry volume, wet weight, wet volume, water absorption capacity and swelling index values, which are clustered in the same region and have approximately the same axis length, can save time and consumable. The zinc, sodium, iron, and copper contents in the beans grown under organic conditions were found to be higher than the data reported in the literature.

Acknowledgements

This study was supported by Turkish Scientific Research Council (TUBİTAK) with the project number of 119O226.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Akbaba U, Sahin Y & Turkez H (2012). Comparison of element contents in haricot beans grown under organic and conventional farming regime for human nutrition and health. Acta Scientiarum Polonorum. Hortorum Cultus 11(2). doi: 10.3906/fiz-1107-12

Akinyele I O & Shokunbi O S (2015). Concentrations of Mn, Fe, Cu, Zn, Cr, Cd, Pb, Ni in selected Nigerian tubers, legumes and cereals and estimates of the adult daily intakes. Food Chemistry 173: 702-708. doi: 10.1016/j.foodchem.2014.10.098

Al-Numair K S, Ahmed S E B, Al-Assaf A H & Alamri M S (2009). Hydrochloric acid extractable minerals and phytate and polyphenols contents of sprouted faba and white bean cultivars. Food Chemistry 113(4): 997-1002. doi: 10.1016/j.foodchem.2008.08.051

Barreto N M, Pimenta N G, Braz B F, Freire A S, Santelli R E, Oliveira A C & Perrone D (2021). Organic black beans (*Phaseolus vulgaris* L.) from Rio de Janeiro state, Brazil, present more phenolic compounds and better nutritional profile than nonorganic. Foods 10(4). 900. doi: 10.3390/foods10040900

Biçer B T, Albayrak Ö & Akıncı C (2017). The effect of different sowing times on yield and yield components in chickpea. *Adnan Menderes University Journal of Agricultural Sciences* 14(1): 51-57. doi:10.25308/aduziraat.295351 (in Turkish)

Ceyhan E, (2010). Variations in grain properties of dry bean (*Phaseolus vulgaris* L.). *International Journal of Agricultural Research* 5(7): 476-481. doi: 10.3923/ijar.2010.476.481

Cichy K A, Wiesinger J A, Berry M, Nchimbi-Msolla S, Fourie D, Porch T G & Miklas P N (2019). The role of genotype and production environment in determining the cooking time of dry beans (*Phaseolus vulgaris* L.). Legume Science 1(1). e13. doi: 10.1002/leg3.13

Cobiac L, McArthur R & Nestel P J (1990). Can eating baked beans lower plasma cholesterol. *European Journal of Clinical Nutrition* 44(11): 819-822

Çalışkan S, Aytekin R İ, Yağız A K & Yavuz C (2018). Effect of full and limited irrigation treatments on seed quality of some common bean cultivars (*Phaseolus vulgaris* L.). *Turkish Journal of Agriculture-Food Science and Technology* 6(12): 1853-1859. doi:10.24925/turjaf.v6i12.1853-1859.2351

de Oliveira A P, Mateó B D S O, Fioroto A M, de Oliveira P V & Naozuka J (2018). Effect of cooking on the bioaccessibility of essential elements in different varieties of beans (*Phaseolus vulgaris L.*). *Journal of Food Composition and Analysis* 67: 135-140. doi: 10.1016/j.jfca.2018.01.012

Elkoca E & Cinar T (2015). Adaptation of some dry bean (Phaseolus vulgaris L.) varieties and lines to Erzurum ecological conditions and their agronomic and quality characteristics. *Anatolian Journal of Agricultural Sciences* 30(2): 141-153. doi:10.7161/anajas.2015.30.2.141-153 (in Turkish)

Ercan R, Ayhan A T L I, Köksel H & Ayşen D A Ğ (1994). Cooking Quality and Composition of Dried Beans Produced in Turkey (English). Food 19(5). (in Turkish)

Ferreira A S, Naozuka J, Kelmer G A & Oliveira P V (2014). Effects of the domestic cooking on elemental chemical composition of beans species (*Phaseolus vulgaris* L.). *Journal of Food Processing* 2014. doi:10.1155/2014/972508

Greenacre M, Groenen P J F & Hastie T (2022). Principal component analysis. Nat Rev Methods Primers 2. 100 (2022). doi: 10.1038/s43586-022-00184-w

Gulumser A, Bozoglu H & Pesken E (2008). Edible legumes (Application Book), 2nd edition, OMU Faculty of Agriculture, Samsun

Husain N & Mahmood R (2019). Copper (II) generates ROS and RNS, impairs antioxidant system and damages membrane and DNA in human blood cells. Environmental Science and Pollution Research 26: 20654-20668. doi:10.1007/s11356-019-05345-

Shi J, Arunasalam K, Yeung D, Kakuda Y, Mittal G & Jiang Y (2004). Saponins from edible legumes: chemistry, processing, and health benefits. *Journal of medicinal food* 7(1): 67-78. doi: 10.1089/109662004322984734

Jahnen-Dechent W & Ketteler M (2012). Magnesium basics. Clinical kidney journal 5(Suppl_1). i3-i14. doi: 10.1093/ndtplus/sfr163

Kajiwara V, Moda-Cirino V & dos Santos Scholz M B (2021). The influence of chemical composition diversity in the cooking quality of Andean bean genotypes. Food Chemistry 339: 127917. doi: 10.1016/j.foodchem.2020.127917

Kaya M, Karaman R, & Çapar M (2016). Evaluation of chickpea genotypes grown in Göller region provinces in terms of some quality and technological characteristics. *Journal of the Central Research Institute of Field Crops* 25(1): 184-190. doi:10.21566/tarbitderg.280377 (in Turkish)

Ludwig D S (2002). The glycemic index: physiological mechanisms relating to obesity, diabetes, and cardiovascular disease. Jama 287(18): 2414-2423. doi: 10.1001/jama.287.18.2414

Malissiova E, Tsokana K, Soultani G, Alexandraki M, Katsioulis A & Manouras A (2022). Organic food: A Study of consumer perception and preferences in Greece. Applied Food Research 2(1). 100129. doi:10.1016/j.afres.2022.100129

Mathers J C (2002). Pulses and carcinogenesis: potential for the prevention of colon, breast and other cancers. British Journal of Nutrition 88(S3): 273-279

Mertens D (2005). AOAC official method 975.03. (2005). Metal in plants and pet foods. Official Methods of Analysis, 18th edn. Horwitz, W., and G.W., Latimer, (Eds), 3–4

McLean E, Cogswell M, Egli I, Wojdyla D & De Benoist B (2009). Worldwide prevalence of anaemia, WHO vitamin and mineral nutrition information system, 1993–2005. Public health nutrition 12(4): 444-454. doi:10.1017/S1368980008002401

Michels K B, Giovannucci E, Chan A T, Singhania R, Fuchs C S & Willett W C (2006). Fruit and vegetable consumption and colorectal adenomas in the Nurses' Health Study. Cancer Research 66(7): 3942-3953. doi: 10.1158/0008-5472.CAN-05-3637

Mitchell D C, Lawrence F R, Hartman T J & Curran J M (2009). Consumption of dry beans, peas, and lentils could improve diet quality in the US population. Journal of the American dietetic association 109(5): 909-913. doi: 10.1016/j.jada.2009.02.029

Mullins A P & Arjmandi B H (2021). Health benefits of plant-based nutrition: focus on beans in cardiometabolic diseases. Nutrients 13(2): 519. doi: 10.3390/nu13020519

Nadeem M A, Çilesiz Y, Ali F, Baloch F S & Karakoy T (2020). Investigation of quality and cooking traits diversity in a global common bean germplasm. *Global Journal of Botanical Science* 8: 21-29. doi: 10.12974/2311-858X.2020.08.3

Ozaktan H & Doymaz A (2022). Mineral composition and technological and morphological performance of beans as influenced by organic seaweed-extracted fertilizers applied in different growth stages. *Journal of Food Composition and Analysis* 114: 104741. doi:10.1016/j.jfca.2022.104741

Ozaktan H, Uzun S, Uzun O & Yasar Ciftci C (2023a). Assessment of agro-morphological traits of common bean genotypes grown under organic farming conditions with multi-variate analyses and applications. Gesunde Pflanzen 75(3): 515-523. doi: 10.1007/s10343-022-00713-3

Ozaktan H, Uzun S, Uzun O & Yasar Ciftci C (2023b). Change in chemical composition and morphological traits of chickpea (*Cicer arietinum* L.) genotypes grown under natural conditions. Gesunde Pflanzen 75(4): 1385-1400. doi:10.1007/s10343-022-00741-z

Öztaş E, Bucak B, Al V & Kahraman A (2007). Determination of winter hardiness, yield and other traits of different chickpea (Cicer arietinum L.) cultivars under Harran Plain conditions. *Harran University Faculty of Agriculture Journal* 11(3-4): 81-85. (in Turkish)

Pedrosa M M, Cuadrado C, Burbano C, Muzquiz M, Cabellos B, Olmedilla-Alonso B & Asensio-Vegas C (2015). Effects of industrial canning on the proximate composition, bioactive compounds contents and nutritional profile of two Spanish common dry beans (*Phaseolus vulgaris* L.). Food Chemistry 166: 68-75. doi: 10.1016/j.foodchem.2014.05.158

Pinheiro C, Baeta J P, Pereira A M, Domingues H & Ricardo C P (2010). Diversity of seed mineral composition of *Phaseolus vulgaris* L. germplasm. Journal of Food Composition and Analysis 23(4): 319-325. doi: 10.1016/j.jfca.2010.01.005

Ramírez-Ojeda A M, Moreno-Rojas R & Cámara-Martos F (2018). Mineral and trace element content in legumes (lentils, chickpeas and beans): Bioaccesibility and probabilistic assessment of the dietary intake. *Journal of Food Composition and Analysis* 73: 17-28. doi: 10.1016/j.jfca.2018.07.007

Ribeiro N D & Kläsener G R (2020). Physical quality and mineral composition of new Mesoamerican bean lines developed for cultivation in Brazil. *Journal of Food Composition and Analysis* 89: 103479. doi: 10.1016/j.jfca.2020.103479

Ribeiro N D, Kläsener G R, Argenta H D S & Andrade F F D (2022). Selection of common bean genotypes with higher macro-and micromineral concentrations in the grains. Pesquisa Agropecuária Brasileira 57. doi: 10.1590/S1678-3921.pab2022.v57.02757

Rodríguez Madrera R, Campa Negrillo A & Ferreira Fernández J J (2024). Modulation of the nutritional and functional values of common bean by farming system: organic vs. conventional. Frontiers in Sustainable Food Systems 7: 1282427. doi: 10.3389/fsufs.2023.1282427

Savage G P, Savage G E, Russell A C & Koolaard J P (2001). Search for predictors of cooking quality of marrowfat pea (*Pisum sativum* L) cultivars. Journal of the Science of Food and Agriculture 81(8): 701-705. doi: 10.1002/jsfa.860.

Schubert C, Guttek K, Reinhold A, Grüngreiff K & Reinhold D (2015). The influence of the trace element zinc on the immune system. LaboratoriumsMedizin 39(s1). 000010151520150060 doi: 10.1515/labmed-2015-0060

Singh N, Kaur N, Rana J C & Sharma S K (2010). Diversity in seed and flour properties in field pea (*Pisum sativum*) germplasm. Food chemistry 122(3): 518-525. doi: 10.1016/j.foodchem.2010.02.064.

Singh N, Jain P, Ujinwal M & Langyan S (2022). Escalate protein plates from legumes for sustainable human nutrition. Frontiers in Nutrition 9: 977-986. doi: 10.3389/fnut.2022.977986

Stanley D W (1992). Hard beans—a problem for growers, processors, and consumers. HortTechnology 2(3): 370-378. doi:10.21273/HORTTECH.2.3.370

Wang N, Hatcher D W, Tyler R T, Toews R & Gawalko E J (2010). Effect of cooking on the composition of beans (Phaseolus vulgaris L.) and chickpeas (*Cicer arietinum* L.). Food Research International 43(2): 589-594. doi: 10.1016/j.foodres.2009.07.012

Wani I A, Sogi D S, Wani A A & Gill B S (2017). Physical and cooking characteristics of some Indian kidney bean (*Phaseolus vulgaris* L.) cultivars. *Journal of the Saudi Society of Agricultural Sciences* 16(1): 7-15. doi: 10.1016/j.jssas.2014.12.002

Willer H, Trávníček J & Schlatter S (2024). The World of Organic Agriculture. Statistics and Emerging Trends 2024.

Yalçın F, Zeki M U T & Köse Ö D E (2018). Determination of suitable chickpea (Cicer arietinum L.) varieties that will provide high yield under Afyonkarahisar and Yozgat conditions. *Journal of Agricultural Faculty of Gaziosmanpaşa University* (JAFAG). 35(1): 46-59. doi: 10.13002/jafag4367 (in Turkish)

Yeken M Z, Çancı H, Kantar F, Karacaören B, Özer G & Çiftçi V (2019). Variation in cooking quality traits in Phaseolus bean germplasm from Western Anatolia. *Banats Journal of Biotechnology*. doı: 10.7904/2068-4738-x(20)-37

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Characterization of the Volatile Profile of Bee Venom from Different Regions in Türkiye Using Gas Chromatography-Mass Spectrometry

Buket Aydeniz-Guneser[a*], Onur Guneser[a], Meral Kekeçoğlu[b,d], Sevgi Kolayli[c]

[a]*Department of Food Engineering, Faculty of Engineering and Natural Sciences, Uşak University, Uşak, TÜRKİYE*

[b]*Department of Biology, Faculty of Arts and Science, Düzce University, Düzce, TÜRKİYE*

[c]*Department of Chemistry, Faculty of Sciences, Karadeniz Technical University, Trabzon, TÜRKİYE*

[d]*Düzce University, Beekeeping Research Development and Application Centre (DAGEM), Düzce, TÜRKİYE*

ARTICLE INFO

Research Article

Corresponding Author: Buket Aydeniz-Guneser, E-mail: buket.guneser@usak.edu.tr

Received: 01 February 2024 / Revised: 30 June 2024 / Accepted: 29 July 2024 / Online: 14 January 2025

Cite this article

Aydeniz-Guneser B, Guneser O, Kekeçoğlu M, Kolayli S (2025). Characterization of the Volatile Profile of Bee Venom from Different Regions in Türkiye Using Gas Chromatography-Mass Spectrometry. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):22-32. DOI: 10.15832/ankutbd.1430185

ABSTRACT

The volatile organic compounds of bee venoms from four different populations of *Apis mellifera anatoliaca*, came from different regions in Türkiye, were analyzed using solid phase microextraction technique combined with gas chromatography-mass spectrometry. A total of 144 volatile compounds were identified in the bee venom samples. The identified volatile compounds included esters, terpenoids, alcohols, acid esters, aldehydes, ketones, and hydrocarbons. It was determined that ester-type volatile compounds characterized the bee venom obtained from the Central Anatolia Region, while bee venom from the Western Black Sea Region had a higher amount of volatile terpenes with spicy and woody aromas. Further studies are required to understand the volatile profile of bee venom, which consists of plant and animal secondary metabolites.

Keywords: Bee Venom, *Apis Mellifera Anatoliaca*, Volatile Compounds, Gas Chromatography-Mass Spectrometry

1. Introduction

Bee venom (Apitoxin) is a complex mixture of natural compounds, characterized by a yellowish-brownish colour, a sharp odour, and a crystalline form. It is secreted by worker bees of *Apis mellifera anatoliaca* (Bogdanov 2016; Flanjak et al. 2021; Uzuner et al. 2021). Honeybees have two separate glands in their sting as the acid venom gland and the alkaline gland, which collectively form Dufour's gland. The venom gland, which plays a crucial role in the production and storage of bee venom, is the most important defence mechanism for the bee. The secretion of bee venom begins in the venom gland with the emergence of a new adult bee (Ali 2014; Özkök 2018). Although immature young bees have some venom, they cannot produce it effectively because their stings have not yet hardened. Bees reach their full venom production potential within 16-19 days. The capacity for venom production varies according to several factors, including season, race, and nutritional status. Depending on the honeybee race and the collection period, a honeybee produces an average of 0.05 mL to 0.3 mL of venom per day (Özkök 2018).

Bee venom consists of many biologically active molecules. Among these, peptides like melittin and apamin, as well as histamine, epinephrine, phospholipase A2, and hyaluronidase have been identified as major components in bee venom (Melda et al. 2021). The quantity and quality of bee venom are influenced by various environmental and genetic factors, including the bees' nutrition sources, the season and period of production, the collecting techniques used, and the genetic origin of bee races (Ramos et al. 2018; Somwongin et al. 2018; Kekeçoğlu et al. 2022). Different characteristics of bee venom from various honeybee races have been determined by numerous studies. For instance, Africanized bee venom has been found to be less active than European bee venom in different species of mice (Kumar et al. 2014; Zidan et al. 2018; Hussein et al. 2019).

The popularity of bee venom has increased for therapeutic use in apitherapy, showing promise in the treatment of various autoimmune diseases and conditions affecting the nervous system, as well as rheumatoid arthritis, osteoarthritis, scleroderma, psoriasis, and skin cancer (Chung et al. 2012; Hwang et al. 2015; Tanugur-Samanci & Kekeçoğlu 2021; Uzuner et al. 2021). Bee venom has traditionally been used for treating skin diseases, and inflammation or pain in muscles, joints, or fibrous tissues. It has also been used in traditional medicine as a preventive agent for various acute and chronic diseases. Moreover, that recent

studies have shown bee venom has anti-carcinogenic activity against some types of cancer, such as breast, liver, and prostate cancer (Park et al. 2011; El Sharkawi et al. 2015; Jung et al. 2018; Ağan & Kekeçoğlu 2020; El-Didamony et al. 2022).

The adaptation of *A. mellifera* races is affected by floral morphologies and climatic conditions. The morphological and physiological structures of honeybees, especially their gland size and secretion, can be differentiated according to environmental adaptation. As a result of this differentiation of morphological and physiological structures of honeybees, the physico-chemical characteristics and bioactivities of bee-collected and brewed products such as honey and propolis, as well as bee secretion products such as bee venom, beeswax were varied (Castro-Vázquez et al. 2008; Miguel & Antunes 2011; Kekecoglu et al. 2021; Kaziur-Cegla et al. 2022). For instance, Ali et al. (2019) observed that the accumulation of lipofuscin and the size of acini in the hypopharyngeal glands of honeybees, responsible for the synthesis of the main royal jelly proteins, were different in foragers and nurses' bees compared to exotic bee races of *A. mellifera carnica* Pollmann and *A. mellifera ligustica* Spinola during the summer and winter seasons. Al-Ghamdi et al. (2011) investigated the differences in hypopharyngeal glands of *Apis mellifera jemenitica* and Carniolan hybrid bees. The researchers observed that the number of secretory cells was higher in the Carniolan hybrid than in *Apis mellifera jemenitica*, while the cytoplasm of both strains was found to have similar characteristics based on the results of haematoxylin and eosin staining analysis.

Studies on the composition of bee venom has predominantly focused on identifying its protein components, while studies on the volatile compound profile of bee venom is scarce, indicating a need for further researches on this subject (Abd El-Wahed et al. 2021; Çaprazlı & Kekeçoğlu 2021; Kaziur-Cegla et al. 2022). The identification and characterization of the volatile compounds of bee venom are essential for revealing its bioactivity and therapeutic properties in apitherapy. In this context, different classes of volatile compounds in bee venom were identified in a most recent study by Isidorov et al. (2023). One hundred forty-nine volatile compounds were identified in bee venom samples from Poland. The researchers found that fresh bee venom samples had higher amounts of lactones and terpenoids than dried bee venom samples. 2-nonanone, 2-nonanol, acetic acid, isoamyl acetate, toluene, α-pinene, limonene, dihydromyrcenol, and γ-heptalactone were found to be major volatiles in fresh bee venom, while dried bee venom samples contained 2-heptanone, isopentanol, acetic acid, octanoic acid, isoamyl-3-methy-2-butenoate, ethyloctanoate, and methylsalicylate at high levels.

The present study aims to reveal the volatile profiles of the bee venom samples from the Anatolian race in various regions of Türkiye using the solid phase microextraction (SPME) technique with gas chromatography-mass spectrometry (GC-MS), which is a powerful and efficient tool for volatile metabolite studies on certain agricultural products.

2. Material and Methods

2.1. Materials

Bee venom (BV) samples were collected using an electroshock-based device (Almışlar Bee Venom Collecting Machine, Product No. 1568, İzmir, Türkiye), following the method described by Samancı & Kekeçoğlu (2019) with minor modifications. The specifications of the bee venom collector are as follows: input voltage: 12 VDC, timer on: 3 sec, timer off: 7 sec, collector frames: 40 cm x 50 cm, and maximum operating time: 15-20 min. Sampling was conducted from inside the hives between 02:00 pm and 05:00 pm, once every 15 days throughout 2022 (May-August). All bee venom samples were stored at -18 °C prior to analysis.

2.2. Analysis of the volatile profiles of bee venom samples

The volatile compounds of bee venom samples were extracted using the solid phase microextraction (SPME) technique (Pawliszyn 1999). SPME is solvent-free, simple, and fast sample preparation technique that easily couples with chromatographic instruments. The SPME technique effectively reduces the sample preparation, extraction and concentration steps to a single step in the analysis of a broad range of volatile compounds from various biological matrices, such as foodstuffs. Despite the advantages of SPME, it has limitations, including fiber lifetime, and limited capacity in high analyte concentrations. The SPME technique was chosen in this study for the extraction and concentration of the volatile compounds in bee venom due to low sample amount and low volatile concentrations (Pawliszyn 1999; Souza-Silva et al. 2015).

Approximately 0.5 g of bee venom was weighed into 40 mL SPME vials, to which 2.5 µL of internal standard was added. Then, the vials were placed in a 40 °C water bath for 20 minutes. Afterwards, an SPME needle (2 cm-50/30 µm divinylbenzene/carboxene/polydimethylsiloxane stable flex, Bellefonte, USA) equipped with a septum cap (tan PTFE/silicone septa, Supelco, USA) was inserted into the vial headspace to extract volatile compounds for 20 minutes at a depth of 2 cm. Following extraction, the SPME needle was immediately injected into the GC-MS system (HP 6890, 7895C MS, Agilent, USA) operating in splitless mode (Güneşer & Yüceer 2017).

Chromatographic separation of volatile compounds was achieved using a polar HP-INNOWax column (Polyethylene Glycol column, 60 m × 0.25 mm ID, 0.25 µm film thickness; J&W Scientific, Folsom, CA, USA). The GC-MS conditions were programmed as follows: helium carrier gas at a constant flow rate of 1.0 mL/min, initial oven temperature of 40 °C held for 10 minutes, ramped to 250 °C at 5 °C/min, final temperature of 250 °C held for 10 minutes. Electron impact ionization at 70 eV was

used for mass spectra acquisition in the range of 35–350 amu. The capillary direct interface temperature was set at 280 °C with a scan rate of 4.45 scans/s.

The volatile organic compounds in the bee venom samples was tentatively identified by comparing obtained mass spectra with those in the National Institute of Standards and Technology (NIST) and Wiley Registry of Mass Spectral database (7th Edition, Wiley) with a mass spectral match threshold of >50. The quantification of volatile organic compounds was expressed as relative abundance (µg/kg) using semi-quantitative techniques (Equation 1) (Güneşer & Yüceer 2017; Avsar et al. 2004). An internal standard comprising 2-methyl-3-heptanone (1 µL/L, for neutral/basic volatiles) and 2-methylpentanoic acid (6 µL/L, for acidic volatiles) in diethyl ether was prepared.

$$\text{Concentration of volatile compounds } \left(\frac{\mu g}{kg}\right) = \frac{\text{Amount of of internal standart x peak area of the compound}}{\text{peak area of internal standart}} \quad \text{(Eq. 1)}$$

3. Results and Discussion

A total of 144 volatile compounds were identified in the bee venom samples obtained from different regions of Türkiye within this study (**BV1**: Central Anatolia Region, **BV2**: Aegean Region; **BV3**: South Marmara Region; **BV4**: Western Black Sea Region of Türkiye. The volatile profile of the bee venom samples is shown in Table 1. The volatile compounds consists of acid esters, alcohols, aldehydes, esters, hydrocarbons, ketones, and terpenoids, The most abundant volatile compounds in the bee venoms were alcohol, esters and terpenes, which are associated with fruity, floral, and plant nuances (Jeleń & Gracka 2016).

As can be seen in Table 1, BV1, BV2, BV3, and BV4 consisted of 67, 67, 68, and 59 volatile compounds, respectively. Isoamyl acetate, octyl acetate, isoamyl alcohol, phenyl ethyl alcohol, (E) 2-decen-1-ol, 2-ethyl-1 hexanol, 1 octanol, 2-nonanol, cyclooctyl alcohol, d-limonene, α-pinene, and phenol were identified in all bee venom samples. The variability in the amounts of alcohol and terpene volatiles in the bee venom samples can be attributed to the diversity of floral resources in each region and regional differences. Thus, it was reported by Ouradi et al. (2021) that different botanical origins, unifloral or multifloral localities have strong influences in the formation of volatiles such as alcohols, aldehydes, esters, ketones and terpenes in honey samples.

Bee venom samples had a similar number of compounds in terms of acidic and alcohol-derived volatile compounds, while aldehydes, ketones, and especially esters volatiles were observed to be quite different. The volatiles including ethyl pentanoate, ethyl senecioate, 2-heptanol, ethyl heptanoate, 2,5-dimethyl pyrazine, dimethyl trisulfide, 2-nonanone, 2-ethyl-5-methyl-pyrazine, trimethyl pyrazine, camphor, bornyl acetate, isobornyl acetate, methyl decanoate, menthol, ethyl benzoate, 1-decanol, 2-acetylpyrrole, and 2-formyl pyrrole were found only in sample BV1. Moreover, it was determined that BV1 had the highest amount of esters volatile compounds compared to other venom samples (Table 1), Ester volatiles, including ethyl acetate, ethyl propionate, ethyl butanoate, are associated with floral and fruity nuances (Jeleń & Gracka 2016). Therefore, BV1 sample was considered to have a fresher and floral flavour than the other samples in terms of sensory characteristics. It is thought that the higher esters content in BV1 sample could be used chromatographic fingerprint to monitor authentication for bee venom obtained from Central Anatolia Region. Hence, it is well known that certain ester type volatiles are the sting alarm pheromones for honeybees, contributing to defensive behaviours. All sting alarm pheromones components are still being investigated for their roles (recruitment, flight, orientation etc.) in bees (Wager & Breed 2000). In this context, the findings obtained in BV1 could be used to reveal defensive behaviour in honeybees from Anatolian Region in Türkiye.

A further point to consider in the volatile profile of BV1 is that the only sulphur-based volatile present in this sample. Dimethyl trisulfide has been reported to be capable of converting highly toxic cyanide into less toxic thiocyanate in rats. Therefore, it exhibits impressive activity in preventing cyanide poisoning, and this pharmacokinetic behavior of the compound has attracted much attention as an encouraging cyanide antidote (Rockwood et al. 2016; Bhadra et al. 2019).

In addition to sulphur-based volatiles, pyrazine-derived volatiles (trimethyl pyrazine, 2,5-dimethyl pyrazine, 2-ethyl-5-methyl-pyrazine, and 2-formylpyrrole) with their pleasant nutty, roasted, and coffee-like aroma (Maga & Sizer 1973; Mortzfeld et al. 2020) were detected only in BV1. Pyrazines and pyrroles can be used as heat markers in food processing and occur in pyrolysis reactions during the destruction of nitrogen sources. Pyrazine-derived volatiles in honeys have been reported to be stored at 10 °C, 20 °C and at 40 °C (Castro-Vázquez et al. 2008). Electrical stimulation can be used to induce bees to sting during the venom extraction. The duration of electrical stimulation can vary from 15 minutes to 8 hours, the temperature of the device from -5 °C to 40 °C (El-Saeady et al. 2016). These electrical impulses may produce pyrazines. Similar to pyrazines, thiazoles are also secondary metabolites formed from amino acids (Kumar & Aggarwal 2019). The occurrence of 2-mercapto-4-phenyl thiazole may be related to the heat treatment or electrical impulse applied to BV1. Pyrazines and thiazoles were also detected only in BV1. A possible explanation for these findings might be that BV1 is exposed to higher thermal conditions compared to other bee venoms.

Table 1- Volatile profile of bee venom samples

No	*Volatile compounds*	*Class*	*Odor description*	*Concentration of Volatile Compounds [mg/kg bee venom]*[a]			
				BV1	***BV2***	***BV3***	***BV4***
1.	Acetic acid	Acid	sharp, pungent, sour vinegar	-	2.54±1.10 [a]	14.94±0.14	6.52±1.57
2.	Propionic acid	Acid	pungent, acidic, cheesy	0.87±.24	-[b]	-	-
3.	Butanoic acid	Acid	sharp, acetic, cheesy	-	0.76±0.10	1.36±0.14	1.64±0.39
4.	Pentanoic acid	Acid	acidic, sharp, cheesy	1.23±0.44	0.14±0.01	1.31±0.12	-
5.	Heptanoic acid	Acid	rancid, sour, cheesy	1.29±0.58	0.21±0.04	1.29±0.06	-
6.	Octanoic acid	Acid	goaty, lamb, mutton fatty	105.08±13.52	45.85±26.06	-	-
7.	Nonanoic acid	Acid	waxy, dirty, cheesy dairy	0.67±0.30	1.60±0.22	0.87±0.01	-
8.	3-methyl butanoic acid	Acid	sour sweaty, cheesy, tropical	7.24±0.40	0.64±0.17	2.73±0.37	2.92±0.85
9.	3-methyl-2-butenoic acid	Acid	green, phenolic, dairy	187.67±13.37	27.67±3.27	73.66±0.46	67.76±18.07
10.	Hexanoic acid	Acid	sour, fatty, sweaty cheesy	2.35±0.60	0.41±0.04	-	0.69±0.16
11.	2-ethyl hexanoic acid	Acid	herbal, earthy	-	-	0.22±0.01	0.18±0.04
12.	Ethanol	Alcohol	solvent	6836.56±158.22	147.36±19.73	2737.85±479.30	1040.30±70.83
13.	Isoamyl alcohol	Alcohol	fusel, alcoholic, whiskey	1644.40±113.15	669.96±47.57	6999.36±1821.54	18208.12±1949.98
14.	3-methyl-1-pentanol	Alcohol	fusel, cognac, winey	-	-	574.27±27.34	-
15.	2-heptanol	Alcohol	fresh, lemongrass, herbal	132.99±17.19	-	-	-
16.	1-hexanol	Alcohol	pungent, ethereal, fusel oily	-	113.30±6.02	944.57±192.32	1202.53±105.85
17.	2-octanol	Alcohol	fresh, spicy, green	-	-	-	584.79±75.11
18.	1-heptanol	Alcohol	musty, leafy, violet	-	90.09±3.54	-	-
19.	2-ethyl-1 hexanol	Alcohol	citrus, fresh, floral,	1094.19±120.47	119.18±1.00	8426.57±1652.59	541.33±52.93
20.	2-nonanol	Alcohol	waxy, green, creamy	1798.54±278.59	207.04±3.70	1429.23±290.29	1642.60±163.76
21.	2-formyl-1-methyl pyrrole	Alcohol	musty, beefy, coffee	347.24±41.39	-	422.92±180.64	-
22.	1-nonanol	Alcohol	fresh, fatty, floral	-	161.78±27.39	-	-
23.	Endo borneol	Alcohol	earthy, minty, camphoreous	42.97±4.82	24.59±2.35	133.16±48.46	-
24.	2-tetra decanol	Alcohol	fruity, waxy, coconut	-	3.93±0.51	-	-
25.	1-butanol	Alcohol	oily, sweet, balsamic	-	-	167.97±84.79	226.51±3.79
26.	1-decanol	Alcohol	fatty, waxy, floral, orange	144.97±53.40	-	-	-
27.	(E) 2-decen-1-ol	Alcohol	waxy, ozone, citrus rose	379.27±100.21	171.54±8.40	864.88±156.53	1874.01±146.98
28.	*p*-propenylanisole	Alcohol	sweet, anise, licorice	-	-	-	258.15±1.56
29.	Benzyl alcohol	Alcohol	floral, rose, phenolic	-	241.61±1.43	1385.79±254.65	7166.39±765.81
30.	Phenyl ethyl alcohol	Alcohol	floral, rose, rose	202.23±21.17	147.44±6.35	768.40±153.13	1320.78±132.16
31.	2-acetylpyrrole	Alcohol	musty, nutty, coumarinic	55.03±7.58	-	-	-
32.	2-phenoxy ethanol	Alcohol	floral, rose, dried rose	67.55±15.63	98.65±0.12	-	-
33.	4-methyl decane	Hydrocarbon	-	-	-	-	143.75±13.24
34.	3,6-dimethyl decane	Hydrocarbon	-	344.03±12.26	150.29±43.12	-	175.24±18.51
35.	5- ethyl-2-methyl octane	Hydrocarbon	-	-	106.07±28.09	-	-
36.	3,7 dimethyl undecane	Hydrocarbon	-	-	104.53±14.33	-	-

37.	1-octene	Hydrocarbon	-	-	-	-	622.30±23.72
38.	2-pentene	Hydrocarbon	-	-	-	-	503.63±121.12
39.	Dichloromethane	Hydrocarbon	-	938.04±288.95	-	-	-
40.	Hexanal	Hydrocarbon	green, fatty, leafy	-	-	-	197.37±29.87
41.	Benzaldehyde	Aldehyde	almond, fruity, powdery	-	224.01±33.02	13469.91±2218.85	24434.86±2228.58
42.	4- methoxy benzaldehyde	Aldehyde	sweet, powdery, anise	-	15.01±0.32	-	-
43.	Decanal	Aldehyde	sweet, aldehydic, orange	-	657.07±67.93	386.28±282.72	1189.32±177.00
44.	2-methyl-3-phenyl propanal	Aldehyde	fresh, marine, ocean		59.93±11.51	-	-
45.	3-methyl butanal	Aldehyde	ethereal, aldehydic, chocolate	-	-	-	364.92±24.73
46.	Octanal	Aldehyde	aldehydic, waxy, citrus	-	-	3267.49±1269.27	5544.78±692.17
47.	Nonanal	Aldehyde	waxy, aldehydic, rose	-	2364.41±124.79	7389.97±692.31	7421.09±1025.25
48.	E-2- octenal	Aldehyde	fresh, cucumber, fatty	-	-	-	3873.49±356.43
49.	Acetone	Ketone	ethereal, apple, pear	-	-	2129.01±410.47	1499.98±75.28
50.	2-pentanone	Ketone	fruity, ethereal, winey	-	-	134.75±18.02	-
51.	6-methyl-5-hepten-2-one	Ketone	citrus, green, musty	-	-	662.26±137.33	536.02±68.29
52.	2-nonanone	Ketone	sweet, waxy, soapy	285.76±33.75	-	-	-
53.	2-decanone	Ketone	orange, floral, fatty	129.42±48.53	11.80±0.30	-	-
54.	Butyrolactone	Ketone	creamy, oily, caramellic	229.95±29.44	60.80±4.74	-	-
55.	Geranyl acetone	Ketone	rose, leafy, floral,		-	123.90±26.05	-
56.	α-isomethyl ionone	Ketone	orris, woody, floral	76.86±47.52	294.18±34.11	-	-
57.	Camphor	Ketone	camphoreous	152.41±16.54	-	-	-
58.	α-amorphene	Terpene	-	-	-	-	381.81±25.02
59.	Epizonarene	Terpene	-	-	-	155.21±45.77	-
60.	L-verbenone	Terpene	verbenone	-	45.90±2.86	95.17±2.79	-
61.	α-muurolene	Terpene	woody	-	-	269.75±97.77	426.58±44.47
62.	Piperitone	Terpene	herbal, minty, camphoreous	-	382.00±9.21	-	-
63.	L-calamenene	Terpene	herbal, spice	-	-	462.87±137.46	405.34±4.10
64.	α-calacorene	Terpene	woody	-	-	42.10±18.39	-
65.	Δ-selinene	Terpene	woody	-	-	84.95±21.86	
66.	α-terpineol	Terpene	lilac, floral, terpenic	100.51±10.97	53.41±7.15	-	-
67.	Geraniol	Terpene	sweet, floral, fruity, rose	-	97.26±13.47	-	-
68.	Cedrol	Terpene	cedarwood, woody, dry sweet	-	-	186.43±66.52	-
69.	Eugenol	Terpene	sweet spicy, clove, woody	-	170.46±21.48	-	-
70.	Thymol	Terpene	herbal, thyme, phenolic	36.98±10.00	21.20±2.87	-	597.75±102.60
71.	Bulnesol	Terpene	spicy	-	-	43.08±12.88	-
72.	α-bisabolol	Terpene	floral, peppery, balsamic	-	-	28.48±7.86	-
73.	Beta eudesmol	Terpene	woody, woody, green	-	-	197.01±59.93	222.31±34.31
74.	β-linalool	Terpene	citrus, orange, floral	121.10±6.56	1751.04±78.83	-	819.13±60.24
75.	Guaiol	Terpene	guaiacwood, rose tea, woody	51.22±5.70	-	126.59±28.74	-
76.	1 octanol	Terpene	waxy, green, orange, rose	2558.04±203.41	1007.09±50.70	5629.25±1207.17	4727.97±456.01
77.	Z-3-octen-1-ol	Terpene	fresh, fatty, greasy, melon	-	-	-	268.49±31.07

78.	Cyclooctyl alcohol	Terpene	-	402.80±51.78	508.55±16.17	1016.44±676.45	3700.31±395.74
79.	Ethyl-2-formyl pyrrole	Terpene	burnt, roasted, smoky	551.24±67.20	-	-	-
80.	Menthol	Terpene	cooling, mentholic, minty	598.94±383.98	-	-	-
81.	Trans-pinocarveol	Terpene	woody, balsamic, fennel	-	60.38±8.17	320.43±93.34	-
82.	β-gurjunene	Terpene	woody, balsamic	-	-	46.13±12.05	-
83.	β-myrecene	Terpene	peppery, terpenic, spicy	48.23±4.42	40.42±3.89	105.54±26.14	-
84.	D-limonene	Terpene	citrus, orange fresh, sweet	2679.12±33.88	15.34±5.11	1101.11±139.08	393.31±29.39
85.	β-pinene	Terpene	woody-green, pine-like	311.09±50.23	50.31±2.61	553.31±132.49	-
86.	β-cymene	Terpene	fresh, citrus, terpenic	152.13±8.05	53.22±22.59	397.12±128.79	-
87.	Styrene	Terpene	sweet, balsamic, floral	114.14±18.32	-	-	318.36±7.19
88.	Copaene	Terpene	woody, spicy, honey	-	-	-	282.28±16.83
89.	Caryophyllene	Terpene	sweet, woody, spicy,	-	317.98±19.65	896.63±180.11	-
90.	Alloaromadedrene	Terpene	woody	-	-	-	337.69±10.90
91.	Δ-cadinene	Terpene	thyme, herbal, woody	-	-	-	91.27±2.96
92.	α-pinene	Terpene	woody, pine, terpenic	578.06±177.93	32.39±2.08	2068.96±756.30	2167.01±285.01
93.	α-terpinen	Terpene	woody, terpenic, lemon	-	-	299.97±55.67	-
94.	Camphene	Terpene	camphoreous, pine, woody	-	-	-	74.61±6.06
95.	β-cyclocitral	Terpene	tropical saffron, herbal	-	-	-	277.22±4.99
96.	*p*-xylene	Terpene	sweet	-	-	127.62±20.75	257.28±35.55
97.	α-N-Methyl ionone	Terpene	orris, violet, powdery	-	34.89±4.38	-	-
98.	β-ionone	Terpene	floral, woody, fruity	45.51±8.56	33.83±4.47	-	-
99.	Lilial	Terpene	muguet	-	145.99±36.87	-	-
100.	Piperonal	Terpene	cherry, vanilla, sweet cherry	-	12.03±	-	-
101.	E,E 2,5 heptadiene	Terpene	-	-	83.92±8.75	-	-
102.	Methyl acetate	Ester	ethereal, sweet, fruity	1265.64±107.75	-	-	-
103.	Ethyl acetate	Ester	ethereal, fruity, sweet	4466.35±900.32	161.24±17.23	665.52±242.00	
104.	3-octen-1-ol acetate (Z)	Ester	fresh, fatty, greasy	-	115.74±8.29	-	-
105.	Isoamyl acetate	Ester	sweet, banana, ripe estery	596.67±108.86	268.85±13.23	3261.12±640.52	2274.55±237.14
106.	Ethyl pentanoate	Ester	fruity, apple, pineapple	244.24±8.35	-	-	-
107.	Hexyl acetate	Ester	fruity, green, apple	-	-	971.76±372.48	188.70±21.69
108.	Octyl acetate	Ester	green, earthy, mushroom,	423.96.84±113.63	252.22±15.08	3725.35±599.55	1651.74±265.54
109.	Bornyl acetate	Ester	woody, camphoreous, menthol	70.38±11.37	-	-	-
110.	Isobornyl acetate	Ester	balsamic, camphoreous, herbal	117.20±29.13	-	-	-
111.	Pentyl propionate	Ester	-	-	-	-	558.53±47.56
112.	Ethyl butanoate	Ester	fruity, juicy, pineapple	529.31±91.12	-	294.12±70.42	94.01±8.16
113.	Ethyl senecioate	Ester	-	19.01±4.89	-	-	-
114.	Ethyl hexanoate	Ester	-	940.19±97.11	-	-	249.68±35.34
115.	Geraniol acetate	Ester	floral, rose, lavender	-	214.09±27.11	-	-
116.	Ethyl propionate	Ester	sweet, fruity, rummy	1056.48±260.51	-	-	-
117.	Ethyl heptanoate	Ester	fruity, pineapple, cognac	620.85±156.73	-	-	-

118.	Ethyl octanoate	Ester	-	5964.08±451.27	-	204.12±149.16	-
119.	Methyl octanoate	Ester	waxy, green, sweet orange	-	133.32±15.13	826.16±205.16	-
120.	Ethyl nonanoate	Ester	waxy, cognac, fruity	2429.75±129.84	94.14±15.08	1459.07±353.03	-
121.	Methyl decanoate	Ester	oily, winey, fruity	194.95±71.82	-	-	-
122.	Ethyl decanoate		Sweet, waxy, fruity	950.71±73.22	-	126.51±57.13	-
123.	Ethyl benzoate	Ester	minty, sweet, wintergreen	182.76±57.00	-	-	-
124.	2-ethyl-3-hydroxy hexyl-2-methylpropanoate	Ester	-	41.87±10.62	-	-	-
125.	Isoamyl butanoate	Ester	fruity, green, apricot	-	-	1459.30±232.78	1765.97±88.56
126.	A-phenyl ethyl acetate	Ester	sweet, honey, floral	-	-	252.72±74.48	-
127.	4- tert-butylcyclohexyl acetate	Ester	woody, cedar, floral	82.12±29.53	27.58±2.60	-	-
128.	Benzyl acetate	Ester	sweet, floral, jasmin	-	12.13±4.20	267.83±69.92	1693.93±147.79
129.	Phenylmethyl acetate	Ester	sweet, floral, honey	-	-	-	169.80±30.41
130.	E-2-decenyl acetate	Ester	waxy, fatty	-	-	-	851.71±104.79
131.	Trimethyl pyrazine	Pyrazine	nutty, nut skin, earthy	101.88±40.74	-	-	-
132.	2,5 dimethyl pyrazine	Pyrazine	cocoa, roasted, nutty beefy	984.66±127.46	-	-	-
133.	2-ethyl-5-methyl-pyrazine	Pyrazine	nutty, coffee, hazelnut	393.55±59.12	-	-	-
134.	2-formylpyrrole	Pyrazine	musty, beefy, coffee	43.78±2.41	-	-	-
135.	2 methoxy phenol	Phenol	phenolic, smoky, spicy	92.90±39.73	-	207.69±52.64	-
136.	2 methoxy-4-methyl phenol	Phenol	spicy, clove, vanilla	-	-	178.72±29.39	-
137.	2-methoxy-4-ethyl phenol	Phenol	coffea, coffea arabica,	-	-	66.54±20.85	-
138.	Phenol	Phenol	phenolic, plastic, rubbery	76.96±5.76	13.69±0.72	89.17±18.25	540.79±51.55
139.	2,4,bis (1.1-dimethyl-ethyl) phenol	Phenol	phenolic	-	128.35± 3.14	-	-
140.	Dimethyl trisulfide	Organic trisulfide	sulfurous, garlic, onion	77.87±20.52	-	-	-
141.	2-mercapto-4-phenyl thiazole	Thiazole	-	187.74±7.10	-	-	-
142.	Elemicin	Alkenylbenzene	spicy, floral	-	32.24± 7.05	-	-
143.	Myristicin	Alkenylbenzene	spicy, warm, balsamic	-	13.38±2.91	-	-
144.	Butylated hydroxytoluene (BHT)	Synthetic anti-oxidant	phenolic, camphoreous	189.41±154.32	16.09±3.93	148.16±28.39	-

Abbreviations: BV1: Central Anatolia Region, BV2: Aegean Region; BV3: South Marmara Region; BV4: Western Black Sea Region.
[a]Mean relative abundance = [concentration of internal standard × peak area of compound]/[peak area of the internal standard], [b]:not detected

The pharmacological activities of pyrazines and thiazoles have been confirmed by numerous clinical studies. Reported biological activities range extensively from anti-microbial, anti-helminthic, anti-inflammatory, and anti-convulsant properties to anti-proliferative and anti-tumor effects. The potential therapeutic effects of pyrazines and thiazoles may contribute to the medical use of BV1. In general, the findings obtained from bee venoms encourage further research into the development of drug formulations (Ali & Sayed 2021; Huigens et al. 2022).

In the case of BV2 sample, 5-ethyl-2-methyl octane, 3,7-dimethyl undecane, 3-octen-1-ol acetate (Z), 1-nonanol, 2-tetra decanol, piperitone, 2-methyl-3-pheynyl propanal, 4-methoxy benzaldehyde, α-N-methyl ionone, piperonal, geraniol acetate, geraniol, lilial and eugenol were detected only in this sample. Geraniol, geraniol acetate, lilial, and eugenol have characteristic fruity-floral, lavender-like, lily-like, clove-like, and spicy odour notes, respectively (Caputi & Aprea 2011; Jeleń & Gracka 2016).

Elemicin and myristicin belong to the class of alkenylbenzene and are secondary metabolites synthesized by various plants (nutmeg, sweet basil, sweet bay, tarragon etc.) (Götz et al. 2022). It is noteworthy that elemicin and myristicin were found only in BV2 among the different bee venom samples. Although these alkenyl benzenes are considered as potential toxicants in food products, a recent review on these compounds emphasized that elemicin and myristicin have several pharmacological and therapeutic activities such as anti-inflammatory, anti-proliferative, and antioxidant properties based on *in vitro* and *in vivo* studies. The researchers emphasized that, considering the ethnopharmacology of myristicin, further clinical studies are urgently needed to confirm its use as a therapeutic agent. Bee venom is a natural source of myristicin with promising potential for its future use in medicine (Seneme et al. 2021; Götz et al. 2022).

In the case of BV3 sample, 3-methyl-1-pentanol, α-phenyl ethyl acetate, α-terpinene, epizonarene, geranyl acetone, 2-methoxy-4-methyl phenol, 2-methoxy-4-ethyl phenol, cedrol, α-calacorene, delta selinene, beta gurjunene, bulnesol, and alpha bisabolol are unique volatiles of this bee venom. 3-methyl-1-pentanol and α-terpinene were the major volatile compounds among these identified aromatic compounds. While 3-methyl-1-pentanol is associated with green, apple with an alcoholic nuance, α-terpinene gives woody and terpenic aromas (Anonymous 2024). BV3 was noteworthy among bee venoms in this study due to its higher terpene and phenol content. In fact, this finding confirms that bee venom can be used for medical purpose. Paulino et al. (2002) highlighted the potential applications of monoterpenes for medical purpose extensively. The researchers pointed out that terpenes exhibit remarkable biological activities including antidiabetic, anti-inflammatory, anti-tumor, cardioprotective, hepatoprotective, and neuroprotective.

Although sample BV4 had the lowest volatile compound profile, the major volatile compounds detected in it, such as, 2-pentene, 2-octanol, copaene, β-cyclocitral, alloaromadedrene, δ-cadinene, α-amorphene, and p-propenyl anisole were not found in other samples. 2-octanol, phenyl methyl acetate, E-2-decenyl acetate, copaene, alloaromadedrene, δ-cadinene and α-amorphene contribute significantly to the unique spicy and woody aroma of BV4 (Anonymous 2024; Jeleń & Gracka 2016). Considering its overall chemical composition, BV4 appears to have lower levels of acidic volatiles, ketones and phenolic contents compared to other bee venoms.

Butylated hydroxytoluene (BHT) is a synthetic antioxidant widely used in many formulations from the food industry to the pharmaceutical, petroleum, and rubber industries (Yehye et al. 2015). The detection of BHT in the bee venom samples except BV4 suggests that synthetic antioxidants contained in oils and bee cakes used to feed honeybees may be transferred to bee products. In an interview with the producers of BV4 (M, Almışlar, September 2022), the manufacturers stated that bee cake is not used to feed bees. The fact that BHT was not detected in BV4 also supports this finding.

α-pinene (Salehi et al. 2019), d-limonene (Anandakumar et al. 2021), isoamyl acetate, and isoamyl alcohol (Ando et al. 2015) are known to exhibit different bioactive effects and are closely related to bee venom's anti-tumor, anti-inflammatory, anti-allergic, analgesic, and anti-microbial activities in the treatment of both acute and chronic conditions. In a recent study by Abd El-Wahed et al. (2021), volatile compounds were identified in beehive air containing a mixture of bee bread, beeswax, honey, royal jelly, propolis, larvae and bee venom. The authors determined a total of 56 volatile compounds, mainly short-chain fatty acids. Similar to our findings, benzyl_alcohol, octanal, (E)-2-octenal, isoamyl_acetate, *n*-octyl acetate, 2-nonanol, benzyl_acetate were detected in the beehive air. Another recently study reported the detection acids, esters, alcohols, and terpenoids in volatile compound analyses carried out on bee products such as propolis. It is noteworthy that terpenes represented a higher proportion (40%) than other volatile compounds (Ding et al. 2021).

4. Conclusions

Bee venom produced by honeybees for self-defence, is also used as a complementary and alternative medicine, which has attracted particular attention due to its various pharmaceutical applications. While the pharmaceutical activity of bee venom is thought to be derived from the venom proteins, such as melittin and apamin, the volatile compounds of bee venom have been relatively neglected. However, bee venom is a holistic entity, and its efficacy should be considered together with its matrix. The findings of the present study regarding bee venoms produced by *Apis mellifera anatoliaca* populations from four different regions in Türkiye revealed that bee venom consists mostly of terpenes and esters. The result of the GC-MS analysis showed that bee

venom samples contain a variety of volatiles, including esters, terpenes, alcohols, organic acids, aldehydes, ketones, and hydrocarbons. It is thought that the differences detected are due to diet, production method and different harvesting times. The volatile compounds of bee venom should be considered when establishing the quality parameters of bee venom. In conclusion, further research is now needed to determine exactly what the essential compounds of bee venom are and what affects them.

Acknowledgements

The authors are greatly indebted to Prof. Dr. Yonca Karagül YÜCEER for her support and permitting our access to her laboratory infrastructure. The data used are confidential. This research does not require ethical approval.

Data availability: Data are available on request due to privacy or other restrictions.

***Authorship Contributions*:** Conceptualization, Data Curation, Formal Analysis, Writing – Review & Editing- Original Draft: B.A.G., O.G.; Conceptualization, Investigation, Review & Editing- Original Draft: M.K., S.K.
All authors reviewed and approved the final version of the manuscript.

Financial Disclosure: The author declared that this study received no financial support.

References

Abd El-Wahed A A, Farag M A, Eraqi W A, Mersal G A, Zhao, C, Khalifa, S A & El-Seedi H R (2021). Unravelling the beehive air volatiles profile as analysed via solid phase microextraction (SPME) and chemometrics. *Journal of King Saud University Science* 33(5): 101449. https://doi.org/10.1016/j.jksus.2021.101449

Ağan A F & Kekeçoglu M (2020). Melittin and cancer treatment: Nanotechnological perspective. *Uludağ Arıcılık Dergisi* 20(2): 221-231. https://doi.org/10.31467/uluaricilik.784365

Al-Ghamdi A A, Al-Khaibari A M & Omar M O M (2011). Effect of honeybee race and worker age on development and histological structure of hypopharyngeal glands of honeybee. *Saudi Journal of Biological Sciences* 18(2): 113-116. https://doi.org/10.1016/j.sjbs.2011.01.001

Ali E M (2014). Contributions of some biological activities of honey bee venom. *Journal of Apicultural Research* 53(4): 441-451. https://doi.org/10.3896/IBRA.1.53.4.13

Ali H, Alqarni A S, Iqbal J, Owayss A A, Raweh H S & Smith B H (2019). Effect of season and behavioral activity on the hypopharyngeal glands of three honeybee *Apis mellifera L.* races under stressful climatic conditions of central Saudi Arabia. *Journal of Hymenoptera Research* 68: 85-101. https://doi.org/doi: 10.3897/jhr.68.29678

Ali S H & Sayed A R (2021). Review of the synthesis and biological activity of thiazoles. *Synthetic Communications* 51(5): 670-700. https://doi.org/10.1080/00397911.2020.1854787

Anandakumar P, Kamaraj S & Vanitha M K (2021). D-limonene: A multifunctional compound with potent therapeutic effects. *Journal of Food Biochemistry* 45(1), e13566. https://doi.org/10.1111/jfbc.13566

Anonymous, (2024). The goodscents company. https://www.thegoodscentscompany.com/search2.html

Ando H, Kurata A & Kishimoto N (2015). Antimicrobial properties and mechanism of volatile isoamyl acetate, a main flavor component of Japanese sake (Ginjo-shu). *Journal of Applied Microbiology* 118(4): 873-880. https://doi.org/10.1111/jam.12764

Avsar Y K, Karagul-Yuceer Y, Drake M A, Singh T K, Yoon Y & Cadwallader K R (2004). Characterization of nutty flavor in Cheddar cheese. *Journal of Dairy Science* 87(7): 1999-2010. https://doi.org/10.3168/jds.S0022-0302(04)70017-X

Bhadra S, Zhang Z, Zhou W, Ochieng F, Rockwood G A, Lippner D & Logue B A (2019). Analysis of potential cyanide antidote, dimethyl trisulfide, in whole blood by dynamic headspace gas chromatography–mass spectroscopy. *Journal of Chromatography A* 1591: 71-78. https://doi.org/10.1016/j.chroma.2019.01.058

Bogdanov S (2016). Bee Venom: Production, composition, quality. In S. Bogdanov (Ed.), *The bee venom* (pp.1-8). Bee product science. https://www.bee-hexagon.net/english/bee-products/venom/

Çaprazlı T & Kekeçoğlu M (2021). Bal arisi zehrinin kompozisyonunu ve üretim miktarini etkileyen faktörler. *Uludag Arıcılık Dergisi* (1): 132-145. https://doi.org/10.31467/uluaricilik.901279

Castro-Vázquez L, Díaz-Maroto M C, González-Viñas M A, De La Fuente E & Pérez-Coello M S (2008). Influence of storage conditions on chemical composition and sensory properties of citrus honey. *Journal of Agricultural and Food Chemistry* 56(6): 1999-2006. https://doi.org/10.1021/jf072227k

Caputi L & Aprea E (2011). Use of terpenoids as natural flavouring compounds in food industry. *Recent Patents on Food, Nutrition & Agriculture* 3(1): 9-16. https://doi.org/10.2174/2212798411103010009

Chung E S, Kim H, Lee G, Park S, Kim H & Bae H (2012). Neuro-protective effects of bee venom by suppression of neuroinflammatory responses in a mouse model of Parkinson's disease: role of regulatory T cells. *Brain, Behavior, and Immunity* 26(8): 1322-1330. https://doi.org/10.1016/j.bbi.2012.08.013

Ding Q, Sheikh A R, Gu X, Li J, Xia K, Sun N & Ma H (2021). Chinese propolis: Ultrasound-assisted enhanced ethanolic extraction, volatile compounds analysis, antioxidant and antibacterial activity comparison. *Food Science and Nutririton* 9(1): 313-330. https://doi.org/10.1002/fsn3.1997

El-Didamony S E, Amer R I & El-Osaily G H (2022). Formulation, characterization and cellular toxicity assessment of a novel bee-venom microsphere in prostate cancer treatment. *Scientific Reports* 12(1): 13213. https://doi.org/10.1038/s41598-022-17391-w

El-Saeady A A, Diab A, Shehata I A A, Nafea E A & Metwaly A A A (2016). Effect of bee venom collecting on the behavior of honeybee colonies. *Journal of Plant Protection and Pathology* 7(6): 347– 351. https://doi.org/10.21608/JPPP.2016.50576

El Sharkawi F Z, Saleh S S & El Sayed A F M (2015). Potential anti cancer activity of snake venom, bee venom and their components in liver and breast carcinoma. *International Journal of Pharmaceutical Sciences and Research* 6(8): 3224. https://doi.org/10.13040/IJPSR.0975-8232

Flanjak I, Kovačić M, Primorac L, Soldić A, Puškadija Z & Rajs B B (2021). Comparison between the quantity and quality of honey bee venom collected in the front and inside of the hive. *Journal of Apicultural Research* pp. 1-6. https://doi.org/10.1080/00218839.2021.1994262

Götz M E, Sachse B, Schäfer B & Eisenreich A (2022). Myristicin and elemicin: Potentially toxic alkenylbenzenes in food. *Foods* 11(13): 1988. https://doi.org/10.3390/foods11131988

Güneşer O & Yüceer Y K (2017). Biosynthesis of eight-carbon volatiles from tomato and pepper pomaces by fungi: *Trichoderma atroviride* and *Aspergillus sojae*. *Journal of Bioscience and Bioengineering* 123(4): 451-459. https://doi.org/10.1016/j.jbiosc.2016.11.013

Huigens III R W, Brummel B R, Tenneti S, Garrison A T & Xiao T (2022). Pyrazine and phenazine heterocycles: Platforms for total synthesis and drug discovery. *Molecules* 27(3): 1112. https://doi.org/10.3390/molecules27031112

Hussein A, El-Ansari M & Zahra A (2019). Effect of the honeybee hybrid and geographic region on the honeybee venom production. *Journal of Plant Protection and Pathology* 10(3): 171–176. https://doi.org/10.21608/jppp.2019.40922

Hwang D S, Kim S K & Bae H (2015). Therapeutic effects of bee venom on immunological and neurological diseases. *Toxins* 7(7): 2413-2421. https://doi.org/10.3390/toxins7072413

Isidorov V, Zalewski A, Zambrowski G & Swiecicka I (2023). Chemical Composition and Antimicrobial Properties of Honeybee Venom. Molecules 28(10): 4135. https://doi.org/10.3390/ molecules28104135

Jeleń H & Gracka A (2016). Characterization of aroma compounds: Structure, physico-chemical and sensory properties. In E. Guichard, C. Salles, M. Morzel, & AM. Le Bon (Eds.), *Flavour: from food to perception* (pp. 126-153). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118929384.ch6

Jung G B, Huh J E, Lee H J, Kim D, Lee G J Park H K & Lee J D (2018). Anti-cancer effect of bee venom on human MDA-MB-231 breast cancer cells using Raman spectroscopy. *Biomedical Optics Express* 9(11): 5703-5718. https://doi.org/10.1364/BOE.9.005703

Kaziur-Cegla W, Jochmann M A, Molt K, Bruchmann A & Schmidt T C (2022). In-tube dynamic extraction for analysis of volatile organic compounds in honey samples. *Food Chemistry X, 14,* Article 100337. https://doi.org/10.1016/j.fochx.2022.100337

Kekeçoğlu M, Çaprazlı T, Samancı A E T, Samancı T & Önder E Y (2022). Factors affecting quality of honeybee venom. *Journal of Apicultural Science* 66(1): 5-14. https://doi.org/10.2478/jas-2022-0001

Kekecoglu M, Sonmez E, Acar M K & Karaoglu S A (2021). Pollen analysis, chemical composition and antibacterial activity of Anatolian chestnut propolis collected from Yıgılca region. *Biology Bulletin* 48(7): 21-728. https://doi.org/10.1134/S106235902106011X

Kumar N R, Devi A, Kriti H & Kriti N (2014). Comparative biochemical studies on the poison glad and poison sac of the worker bees of three different apis species (*Apis dorsata, Apis mellifera* and *Apis florea). International Journal of Therapeutic Applications* 16: 8–16

Kumar S & Aggarwal R (2019). Thiazole: A privileged motif in marine natural products. *Mini-Reviews in Organic Chemistry* 16(1): 26-34. https://doi.org/10.2174/1570193X15666180412152743

Maga J A & Sizer C E (1973). Pyrazines in foods. Review. *Journal of Agricultural and Food Chemistry* 21(1): 22-30.https://doi.org/10.1021/jf60185a006

Melda A, Kalaycioğlu Z, Kolayli S & Berker B (2021). Comparative determination of melittin by capillary electrophoretic methods. *Journal of the Turkish Chemical Society Section A: Chemistry* 8(4): 1211-1216. https://doi.org/10.18596/jotcsa.949188

Miguel M G & Antunes M D (2011). Is propolis safe as an alternative medicine? *Journal of Pharmacy and Bioallied Sciences* 3(4): 479-495. https://doi.org/10.4103/0975-7406.90101

Mortzfeld F B, Hashem C, Vranková K, Winkler M & Rudroff F (2020). Pyrazines: Synthesis and industrial application of these valuable flavor and fragrance compounds. *Biotechnology Journal* 15(11): 2000064. https://doi.org/10.1002/biot.202000064

Ouradi H, Hanine H, Fauconnier M L, Kenne T, Rizki H, Ennahli S & Hssaini L (2021). Determination of physico-biochemical proprieties and composition in volatile constituents by solid phase micro-extraction of honey samples from different botanical and geographical origins in Morocco. *Journal of Apicultural Research* 60(1): 84-98. https://doi.org/10.1080/00218839.2020.1718339

Özkök A (2018). *Türkiye'de hızla büyüyen sektör: Arı ürünlerine genel bir bakış.* Palme Yayın Dağıtım.

Paulino B N, Silva G N, Araújo F F, Néri-Numa I A, Pastore G M, Bicas J L & Molina G (2022). Beyond natural aromas: The bioactive and technological potential of monoterpenes. *Trends in Food Science & Technology* 128: 188-201

Pawliszyn J (1999). Quantitative aspects of SPME. In: J. Pawliszyn (Ed.), *Applications of solid phase microextraction* (pp. 3-21). *Royal Society of Chemistry*. https://doi.org/10.1039/9781847550149

Ramos O Y, Salomón V, Libonatti C, Cepeda R, Maldonado L & Basualdo M (2018). Effect of botanical and physicochemical composition of Argentinean honeys on the inhibitory action against food pathogens. LWT 87: 457-463. https://doi.org/10.1016/j.lwt.2017.09.014.

Rockwood G A, Thompson D E & Petrikovics I (2016). Dimethyl trisulfide: A novel cyanide countermeasure. *Toxicology & Industrial Health,* 32(12): 2009-2016. https://doi.org/10.1177/0748233715622713

Salehi B, Upadhyay S, Erdogan Orhan I, Kumar Jugran A, Jayaweera L D S A, Dias D, Sharopov F, Taheri Y, Martins N, Baghalpour N, Cho W C & Sharifi-Rad J (2019). Therapeutic potential of α-and β-pinene: A miracle gift of nature. *Biomolecules* 9(11): 738. https://doi.org/10.3390/biom9110738

Samancı T & Kekeçoglu M (2019). Comparison of commercial and Anatolian bee venom in terms of chemical composition. *Uludag Arıcılık Dergisi 19*(1): 60-67. https://doi.org/10.31467/uluaricilik.527986

Seneme E F, Dos Santos D C, Silva E M R, Franco Y E M & Longato G B (2021). Pharmacological and therapeutic potential of myristicin: A literature review. Molecules, 26 (19), 5914. https://doi.org/10.3390/molecules26195914

Somwongin S, Chantawannakul P & Chaiyana W (2018). Antioxidant activity and irritation property of venoms from Apis species. *Toxicon,* 145: 32–39. https://doi.org/10.1016/j.toxicon.2018.02.049

Souza-Silva ÉA, Gionfriddo E & Pawliszyn J (2015). A critical review of the state of the art of solid-phase microextraction of complex matrices II. Food analysis. *TrAC Trends in Analytical Chemistry* 71: 236-248.https://doi.org/10.1016/j.trac.2015.04.018

Tanugur-Samanci A E & Kekeçoglu M (2021). An evaluation of the chemical content and microbiological contamination of Anatolian bee venom. *Plos One, 16*, 1-14, https://doi.org/10.1371/journalpone.0255161

Uzuner S Ç, Birinci E, Tetikoğlu S, Birinci C & Kolaylı S (2021). Distinct epigenetic reprogramming, mitochondrial patterns, cellular morphology, and cytotoxicity after bee venom treatment. *Recent Patents on Anti-Cancer Drug Discovery* 16(3): 377-392. https://doi.org/10.2174/1574892816666210422125058

Wager B R & Breed M D (2000). Does honey bee sting alarm pheromone give orientation information to defensive bees? *Annals of the Entomological Society of America* 93(6): 1329-1332.https://doi.org/10.1603/0013-8746(2000)093[1329:DHBSAP]2.0.CO;2

Yehye W A, Rahman N A, Ariffin A, Abd Hamid S B, Alhadi A A, Kadir F A & Yaeghoobi M (2015). Understanding the chemistry behind the antioxidant activities of butylated hydroxytoluene (BHT): A review. *European Journal of Medicinal Chemistry* 101: 295-312. https://doi.org/10.1016/j.ejmech.2015.06.026

Zidan H A E G, Mostafa Z K, Ibrahim M A, Haggag S I, Darwish D A & Elfiky A A (2018). Venom composition of Egyptian and Carniolan honeybee. *Apis mellifera* L. affected by collection methods. *Egyptian Academic Journal of Biological Sciences* 11(4): 59–71. https://doi.org/10.21608/EAJBSA.2018.17733

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

Characterization of Rosemary (*Salvia rosmarinus*) Essential Oil Obtained by Solvent-Free Microwave Extraction with Kombucha Tea *(Anthriscus sylvestris* L.) Produced by Adding Guava (*Psidium guajava* L.) Peel and Pulp

Filiz Yangılar[a*], Merve Dilara Gerek[b]

[a]*Erzincan Binali Yıldırım University Faculty of Health Sciences Department of Nutrition and Dietetics, Erzincan, TÜRKIYE*
[b]*Erzincan Binali Yıldırım University Graduate School of Health Sciences Department of Nutrition and Dietetics, Erzincan, TÜRKIYE*

ARTICLE INFO
Research Article
Corresponding Author: Filiz Yangılar, E-mail: f_yangilar@hotmail.com
Received: 28 March 2024 / Revised: 29 June 2024 / Accepted: 29 July 2024 / Online: 14 January 2025

Cite this article
Yangılar F, Gerek M D (2025). Characterization of Rosemary (*Salvia rosmarinus*) Essential Oil Obtained by Solvent-Free Microwave Extraction with Kombucha Tea (*Anthriscus sylvestris* L.) Produced by Adding Guava (*Psidium guajava* L.) Peel and Pulp. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):33-45. DOI: 10.15832/ankutbd.1460437

ABSTRACT
The study aimed to evaluate the potential of creating a functional beverage by enriching *Anthriscus sylvestris* (L.) kombucha production with guava peel and pulp, showcasing the variety and richness of the product formulations. This was done due to kombucha being a fermented beverage rich in various bioactive compounds that have significant effects on health and the possibility of enhancing product diversity through different formulations. Thus, the kombucha tea was physicochemical, phytochemical, antioxidant activity, and sensory evaluation during the 15-day fermentation period. The increased acetic acid bacteria and yeast count indicates that *Anthriscus sylvestris* is a viable substrate for the proliferation of Kombucha symbiotic microbes. Adding guava peel and rosemary oil to kombucha significantly increased its total phenolic, flavonoid, and vitamin C content ($P<0.05$). *Anthriscus sylvestris*-guava kombucha has antibacterial properties (zones of inhibition of 18.32 and 17.31 mm against *Staphylococcus aureus* ATCC 25923 and *Escherichia coli* ATCC 25922, respectively). Moreover, it also provided an inhibitory effect on α-glucosidase and α-amylase, enzymes associated with diabetes. The DPPH, ABTS^{*+}, FRAP, CUPRAC, and ORAC capacities were found to be 66.89-78.90%, 70.83-97.25%, 367-723.15 µmol Trolox/mL, 381.40-460.45 µmol Trolox/mL, and 486.50-737.50 µmol Trolox/mL, respectively. *Anthriscus sylvestris* kombucha has well-balanced and pleasant sensory properties. These findings indicate that preparing *Anthriscus sylvestris* kombucha tea with guava may have health effects and that further research is needed to determine potential health benefits.

Keywords: Kombucha, *Anthriscus sylvestris*, α-glucosidase, Guava peel and pulp, Rosemary essential oil, Functional beverage

1. Introduction

Kombucha is a beneficial beverage created by fermenting dried green tea (*Camellia sinensis*) and sugar with a symbiotic culture of yeast and bacteria called SCOBY, also referred to as "fungi tea" or "tea fungus" (Kitwetcharoen et al. 2023). Kombucha's phenolic and flavonoid content, antioxidant activity, antibacterial properties, and anti-diabetic effects have all been extensively studied (Morales 2020). These beverages include bioactive substances that can block enzymes involved in glucose digestion, potentially benefiting diabetics (Gao et al. 2013; Permatasari et al. 2021; Oliveira et al. 2023). Diversifying kombucha fermentation substrates entails using a variety of ingredients, such as fruits, vegetables, herbs, spices, tea blends, coconut water, honey, coffee, and alternative sweeteners, to create distinct and flavorful kombucha variations with potential health benefits (de Miranda et al. 2022). *Anthriscus sylvestris* (L.) Hoffm. is commonly referred to as wild chervil or cow parsley and is part of the Anthriscus genus (Umbelliferae). It can be found primarily across Asia, Europe, North America, Africa, and New Zealand (Olaru et al. 2015; Liu et al. 2023). Also recognized by the names wild chervil or cow parsley, this plant is classified under the Scandiceae tribe within the Apiaceae family (Olaru et al. 2016; Janković et al. 2024). Additionally, *Anthriscus sylvestris* is a significant source of lignan compounds, particularly highlighted in the fields of pharmacy and nutrition following the identification of the aryl tetralin lignan podophyllotoxin (Koulman et al. 2001; Berežni et al. 2024). Utilized in traditional medicine, *Anthriscus sylvestris* is recognized for its antibiotic, soothing diuretic, and cough-relieving properties (Chen et al. 2014).

Rosemary (*Rosmarinus officinalis* L.) represents a significant industrial crop within the realm of medicinal and aromatic vegetation (Sakar et al. 2023). The polyphenols derived from rosemary have garnered noteworthy attention due to their conspicuous anti-inflammatory properties, including the inhibition of inflammatory cell infiltration, suppression of inflammation-associated signalling pathways, reduction in the synthesis of inflammatory cytokines, and modulation of the

composition of intestinal microbiota (Mengoni et al. 2011; Li et al. 2016; Xia et al. 2017; He et al. 2022; Du et al. 2023; Zhang & Lu 2024). Notably, rosemary essential oil finds widespread utility across various sectors, particularly in food, phytomedicines, nutraceuticals, cosmetics, agrochemicals, and aromatherapeutic formulations (Martin-Piñero et al. 2020; Araya et al. 2022; Song et al. 2023).

The demand for kombucha as a functional beverage is rapidly growing due to its exceptional nutritional properties, which boost product diversity (Kim & Adhikari 2020). Guava (*Psidium guajava* L.) fruits are a sustainable supply of antioxidant compounds, polyphenols, and colorants, as well as raw protein, fiber, carbs, vitamin C, potassium, calcium, and phosphorus (Kamath et al. 2008; Kaur & Ghosh 2023). Guava has been proven to reduce the risk of cancer in body parts such as the abdomen (Chakrabortya & Athmaselvi 2014; Ninga et al. 2018). Many pharmacological investigations have shown that guava contains antioxidant, hepatoprotective, antibacterial, antidiabetic, anti-inflammatory, and anticancer properties, which support its traditional usage (Jaiarj et al. 1999; Oh et al. 2005; Burgin et al. 2010; Correa & Couto 2016; König et al. 2019; Zou & Liu 2023). Because of its biological qualities, guava fruit can also be used to make kombucha.

This study looked into *Anthriscus sylvestris* (L.)-based kombucha with guava pulp and peel, as well as rosemary essential oil. This study sought to identify a viable functional kombucha beverage, thereby growing the food market by introducing items with superior qualities.

2. Material and Methods

2.1. Materials

Anthriscus sylvestris (L.) plants were hand-picked in the Erzurum province of Turkey in May 2023, when the plants had grown to a height of 10 cm. Identification of the *Anthriscus sylvestris* L. was done at Department of Biology, Faculty of Science and Arts, Erzincan Binali Yıldırım University, Erzincan, Turkey. The kombucha starting cultures, SCOBY from medical companies (CENTRAGEN Laboratory Products Informatics and Consultancy Company), Guava (*Psidium guajava* L.), and Rosemary (*Salvia rosmarinus*) were ordered from an online purchasing site via the internet. Clinical isolates of pathogenic bacteria (*Enterococcus faecalis* ATCC 29212, *Staphylococcus aureus* ATCC 25923, *Escherichia coli* ATCC 25922, and *Pseudomonas aeruginosa* ATTC 27853) were obtained from the Faculty of Medicine Microbiology Laboratory at Erzincan Binali Yıldırım University.

Figure 1- The Solvent-free microwave extraction (SFME)

2.2. Methods

2.2.1. The extraction of Rosemary (Salvia rosmarinus) essential oil

The microwave extraction system (Ethos X) was used for the extraction. Rosemary (*Salvia rosmarinus*) was placed in a Pyrex disc and exposed to microwave irradiation at 622 W (50% power) for 30 minutes at 2450 MHz. No solvent or water was added. The Clevenger apparatus was positioned through an aperture on the top of the microwave device. The opening around the neck

of this apparatus is sealed with polytetrafluoroethylene to prevent microwave leakage. This procedure is carried out under atmospheric pressure. Vapour generated during the process was condensed in an external condenser. Essential oil and water were separated by decantation. Anhydrous sodium sulfate was used to dry the rosemary essential oil (REO), which was stored in amber-capped vials at 4 °C. The Solvent-free microwave extraction (SFME) is given in Figure 1.

2.2.2. Preparation of kombucha

The washed *Anthriscus sylvestris* (L.) plant samples were dried at 50 ºC and 1 m/s air flow rate. The drying process was performed until the moisture content of the plant leaves reached to approximately <7% (w/w). The dried plant leaves were then packed and stored at 4 ºC until analysis. The pulp and the seeds were separated from the peel by cutting. The guava peel or pulp was individually dried in a household microwave oven (Arçelik, ARMD580, Turkey) for 16 hours at 55 °C until it reached approximately 10±2% moisture content. A blender (Warning Commercial, USA) was used for ground and sieved (355 - µm mesh). The powders were stored at 4 ºC until analysis. After drying, the materials were crushed and mixed with hot water at a concentration of 10 g L^{-1} for 20 min. The tea broth was then sweetened with 6% (w/w) sugar. The produced tea samples were transferred to glass jars and pasteurization at 90 °C for 20 min. Based on the literature review and preliminary tests, once the beverages reached room temperature, we incorporated 10% SCOBY (w/v) and 1% REO. The codes and descriptions for various Kombucha samples are given in Table 1. Samples were analyzed on days 0, 5, 10, and 15 of fermentation.

Table 1- Coding and descriptions for various kombucha samples

Samples	***Descriptions***
ASK1	Control kombucha (no Guava peel/pulp, and no REO*)
ASK2	Kombucha with guava pulp
ASK3	Kombucha with guava peel
ASK4	Kombucha with guava pulp and 1% REO
ASK5	Kombucha with guava peel and 1% REO

*: REO: Rosemary essential oil

2.2.3. pH, titrable acidity, and color analyses

pH values using a pH meter (Eutech PH 150 Model), and the total acidity using the AOAC method (AOAC, 2000) were determined. The color was measured on a scale of *L**, *a**, and *b** using a HunterLab (Colorflex-EZ, Hunterlab, Virginia, USA) (Varela-Santos et al. 2012).

2.2.4. Acetic acid bacteria and yeast count

Microbial research included the detection of acetic acid bacteria (AAB) and yeast counts. While counting acetic acid bacteria in samples, dilutions of Glucose Yeast Extract Calcium Carbonate Agar (GYC, 10% glucose, 1% yeast extract, 2% calcium carbonate, 1.5% agar, pH 6.8±0.2) were inoculated onto the medium using the spread plate method and the petri dishes were placed at 30 °C. It was incubated at 30 °C for 5-10 days. Following the incubation period the colonies with a cream color and a clean sedimentation zone around the petri dishes were measured as typical AAB colonies (De Vero et al. 2006). To count the yeast, 10 mL of kombucha was mixed with 0.1% (w/v) peptone water and seeded on Dichloran Rose Bengal Chloramphenicol (DRBC) agar by the spreading method. The DRBC plates were then incubated at 25 °C for five days (AOAC 2000).

2.2.5. Analysis of antioxidant activity

Antioxidant capacity was found using the 2,2-diphenyl-1-picrylhydrazyl radical (DPPH), 2,20-azino-bis (3-ethylbenzothiazoline-6-sulfonic acid) diammonium salt ($ABTS^{\bullet+}$), ferric ion reducing potency method (FRAP), copper-reducing antioxidant activity (CUPRAC), and oxygen radical absorption capacity (ORAC) methods. The DPPH assay was performed according to the method used by Mazidi et al. (2012) and Asl et al. (2018), with some changes. In brief, 0.2 mL of beverage was combined with 0.8 mL of methanol and a stock solution with volumes ranging from 0.0125 to 0.2 mL/mL was prepared. Following that, 0.5 ml of the diluted sample was mixed with 1.5 mL of DPPH solution (0.1 mM). It was incubated for about an hour at room temperature in a dark atmosphere, with absorbance measurements obtained at 517 nm. The inhibitory activity was calculated using the equation:

$$Inhibition\ activity\ (\%) = [(A_{control} - A_{sample}) / A_{control}] \times 100$$

The $ABTS^{\bullet+}$ value of samples was assessed according to the method proposed by Miller et al. (1993). Peroxidase, $ABTS^{\bullet+}$ and hydrogen peroxide solutions were prepared using a buffer solution of sodium phosphate (5 mM, pH 7.6). $ABTS^{\bullet+}$ was prepared by mixing 0.3 mL peroxidase and 0.3 mL $ABTS^{\bullet+}$ with 0.4 mL H_2O_2 and incubating for approximately 30 min. in a dark environment. Then, 0.01 mL of the beverage samples was added to this solution and incubated for approximately 10 minutes $ABTS^{\bullet+}$ values were calculated by reading the absorbance at 734 nm.

The FRAP values of beverages were found using the method specified by Benzie & Strain (1999). In this method, 100 µL of sample and 900 µL of FRAP solution were added into the test tubes. The tubes were incubated for approximately 4 minutes, and their absorbance was measured at 593 nm.

The CUPRAC method involves mixing 100 µL of sample, 3 mL of CUPRAC reagent, and 1 mL of pure water. The CUPRAC solution was made from $CuCl_2$, neocuproine, and $NH_4C_2H_3O_2$. These substances are blended in equal quantities. After 30 minutes of incubation, absorbance values were measured at 450 nm using a spectrophotometer (R^2=0.9981; Apak et al. 2004).

The ORAC values of the beverages were found using the method described by Chusak et al. (2014). For analysis, the drinks were first diluted and incubated with 4.8 nM sodium fluorescein in a 75 mM PBS solution at 37 °C using a 0.1 M phosphate buffered solution for approximately 10 minutes. After incubation, 25 µL of 64 mM AAPH was added to this mixture. Spectrofluorometric readings were made using the ORAC-Florescein method at 485-520 nm.

2.2.6. Total phenolic content (TPC)

The phenolic content of the beverages was assessed using the method of Somsong et al. (2017). First, 2.4 mL of the beverage sample was diluted 10 times with pure water. Then, 150 µL of Folin Ciocalteu reagent prepared in pure water (1:10) was added and mixed for approximately 2 minutes. 7.5% (w/v) Na_2CO_3 was added and incubated at room temperature for approximately 2 hours, and the absorbance was read at 765 nm.

2.2.7. Total flavonoid content (TFC)

Flavonoids in beverages, Ahlem Rebaya et al. (2015) stated, were made according to the method. In this analysis, catechin was used at rates of 0.01-0.1 mg/mL. 25 µL of standard was added to 15 µL of 5% $NaNO_2$ solution, and after approximately 6 minutes, 30 µL of aluminum trichloride (10%) was added and incubated for 5 minutes. Then, 150 µL of NaOH was added, and after incubation for approximately 15 minutes, the color turned to a liquid pink. After this, absorbance values were read at 510 nm with a microplate reader.

2.2.8. Vitamin C extraction and determination

The ascorbic acid values of beverages were determined using the method of Zhao and Shah (2014). For this, 1 g of sample was thoroughly dissolved in 20 mL of 80% methanol, adding 0.01% trifluoroacetic acid (TFA), and agitated at 200 rpm at 50 °C for approximately 4 hours. The extract was then centrifuged at 4500 g for approximately 5 minutes. Using a rotary evaporator, the extract was evaporated at 45 °C and redissolved in 5 mL of 80% methanol (v/v) to determine vitamin C, α-glycosidase, and α-amylase inhibition activities.

The acetic acid concentration was found using the method given by Li et al. (2017). The examination of acetic acid content was performed using an Agilent 1200 HPLC system equipped with a UV detector. The separation was done using a Waters Atlantis T3 column (4.6 mm×250 mm, 5 µm) at 35 °C. The mobile phase included 0.02 mol/L NaH_2PO_4 (pH 2.7). After injecting 10 µL of the supernatant into the column. The organic acids were evacuated with a phase that moved at a flow rate of 0.9 mL/min and monitored at 210 nm. Calibration curves for the pure form containing these acids were created and assessed.

2.2.9. α-glucosidase and α-amylase inhibition activities

To test α-glucosidase inhibition, combine 200 µL of extracts or 80% methanol (control) with 100 µL of 1.0 U/mL α-glucosidase solution in 0.1 M phosphate buffer (pH of 6.9) and incubate at 37 °C for 10 min. After adding 200 µL of pNPG and incubation for 10 minutes. The action ended with 1 M Na_2CO_3, and the absorbance's were obtained at 405 nm. The initial absorbance was eliminated from both the sample and the control absorption values.

To test α-amylase inhibition in beverages, 250 µL of phenolic was dissolved by 250 µL of α-amylase in 20 mM sodium phosphate buffer (pH 6.9) and kept at 37 °C for ten minutes. Following that, 250 µL of a solution containing 1% starch was included and incubated at 37 °C for roughly ten minutes. 500 µL of 3, 5-dinitrosalicylic acid was supplied. The blend was left in boiling water for 5 minutes. After being cooled to ambient temperatures, absorbance was detected at 540 nm.

2.2.10. Antimicrobial activity

The antibacterial effect of beverages was determined using the agar well diffusion method against *Enterococcus faecalis* ATCC 29212, *Staphylococcus aureus* ATCC 25923, *Escherichia coli* ATCC 25922, and *Pseudomonas aeruginosa* ATTC 27853. A sterile copper tube was used to create 6 mm-diameter wells on Mueller-Hinton agar plates. A sterile cotton swab was used to evenly distribute bacterial solutions containing 10^6 cfu mL^{-1}.

To produce a sterile residual of fermented kombucha, the resulting infusions were centrifuged and passed through a 0.22 µm sterilized microfilter. After adding 100 µL of each sample to the wells, the plates were kept at 4 °C for 2 hours to allow pre-diffusion into the agar. The plates were then left to incubate at 37 °C for a total of 18 hours. Kanamycin (50 µg/well) was used as positive guidance, while distilled water acted as the negative standard (Khazi et al. 2024).

2.2.11. Sensory analysis

Participants used a hedonic scale to rate the color, appearance, aroma, flavor, sourness, and overall acceptability of kombucha drinks. The scale ranged from one (severe dislike) to five (great liking). The sensory acceptability tests were done by panellists of 15 untrained judges of both sexes (9 women and 6 men).

2.2.12. Statistical analysis

A factorial experimental design was used for the research trial, which included five applications and four fermentation days (0^{th}, 5^{th}, 10^{th}, and 15^{th} days). The acquired data were subjected to variance analyses, and Duncan's multiple comparison tests were used to determine the significant veracity with SPSS software package version 25.0.

3. Results and Discussion

3.1. pH and acidity

pH of kombucha is critical for determining its microbiological safety because of the antibacterial effects caused by the buildup of organic acids (Villarreal-Soto et al. 2019). In the current study, pH values taken during fermentation days showed a progressive drop along with the formation of acetic acid (Table 2). The pH of kombucha samples fluctuated significantly ($P<0.05$). The highest pH value of the samples (3.47±0.04) was observed in the ASK1 sample on the 0th day of fermentation, while the lowest value (2.37±0.04) was found in the ASK4 sample on the 15^{th} day of fermentation. Furthermore, in samples containing especially plain guava pulp and peel with added rosemary essential oil, a decrease in pH values has been observed. The mean pH values for ASK1, ASK2, ASK3, ASK4, and ASK5 were 3.07±0.36, 2.82±0.29, 2.86±0.27, 2.74±0.26, and 2.85±0.33, respectively. Yavari et al. (2011) and Ayed & Hamdi (2015) both reported similar results.

According to the findings in Table 2, the initial total acidity values of samples increased during fermentation, particularly acutely until the 15^{th} day. The concentration of acetic acid in *Anthriscus sylvestris* (L.) kombucha samples differed by 1.27% and 10.65%, respectively. The acetic acid values of the samples were found to be statistically significant both in terms of the fermentation duration and the treatments applied ($P<0.05$). The acidity value has increased in all samples depending on the fermentation time. Particularly, the highest values were observed on the 15^{th} day. Additionally, among the samples, the highest acidity value on the 15^{th} day (10.65±0.91%) was found in the ASK4 sample. The results may have been influenced by the presence of pulp and rosemary essential oil. Zubaidah et al. (2018) fermented the snake fruit juices they produced using kombucha consortium for 14 days. The acidity of the beverages ranged from 0.92 to 1.65%. These study results outperformed those of the aforementioned researchers.

3.2. Acetic acid bacteria and yeast counts

Table 2 shows the numbers of acetic acid bacteria and yeast. Following inoculation, there were substantial differences in the numbers of yeast and acetic acid bacteria between the samples. This variation could be related to the use of three different symbiotic bacterium and yeast cultures (SCOBYs), each with a different number of viable cells. The acetic acid bacteria count of the samples was found to be statistically significant both in terms of the fermentation duration and the treatments applied ($P<0.05$). The sample with the highest acetic acid bacteria count was ASK1 on the 0^{th} day of fermentation (7.47±0.03 log cfu^{-1}), while the sample with the lowest count was ASK2 on the 15^{th} day of fermentation (2.55±0.11 log cfu^{-1}). Additionally, samples with rosemary essential oil contained fewer acetic acid bacteria compared to the other samples. This result could be attributed to the antimicrobial effect of the rosemary essential oil. These findings confirm that the unusual portions of *Anthriscus sylvestris* (L.) are suitable substrates for the production of these compounds. The mean yeast count was 4.56±0.27, 4.98±0.14, 5.10±0.40, 5.22±0.14, and 4.70±0.26 log cfu/mL for samples ASK1, ASK2, ASK3, ASK4, and ASK5. These findings confirm that the leaves and stems of *Anthriscus sylvestris* (L.) are acceptable substrates, as they possess values comparable to those of typical kombuchas with a shorter fermentation time. The highest yeast count on the 5^{th} day of fermentation was observed in the ASK4 sample (5.18±0.16 log cfu^{-1}), while the lowest was in the ASK1 sample (4.46±0.42 log cfu^{-1}). Similarly, Sreeramulu et al. (2000) discovered 4.48 log cfu mL^{-1} cells of yeast and 5.3 log cfu mL^{-1} cells of bacteria in fermented liquids following six days of fermentation. Teoh et al. (2004) discovered that the counts of yeast on the sixth day of the process were between 5 and 7 log cfu/mL. Because kombucha has no established microbiological or chemical makeup, some variations in chemical parameters, process length, and cell counts are to be expected (Teoh et al. 2004). The yeast results obtained by the researchers on the 6^{th} day of fermentation align with the findings of this study. According to De Filippis et al. (2018), the most common microorganisms in kombucha culture are acetic acid bacteria and yeasts. These microorganisms' activity is complementary since yeasts hydrolyse

sucrose to produce fructose and glucose for ethanol synthesis. Acetic bacteria finally convert ethanol to acetic acid and other organic acids (Zubaidah et al. 2019).

3.3. Total phenolic content (TPC)

Secondary metabolites in plant-based products include phenolic chemicals, which have significant antioxidant properties and a positive impact on human health. These molecules affect the products' aroma, taste, and color features (Chakravorty et al. 2016; Basli et al. 2017; Delgado et al. 2019; Alara et al. 2021). As a result, alternative substrates derived from other plant species may contain phenolic chemicals (Leonarski et al. 2022). During the fermentation period, the overall phenolic content within ASK1 consistently exhibited lower levels than the other samples ($P<0.05$). Towards the conclusion of the fermentation duration, significant increments in total phenolic content were observed in ASK2, ASK3, ASK4, and ASK5 ($P<0.05$) (Figure 2A). The TPC of guava pulp and peel kombucha samples differed significantly from the control sample ($P<0.05$). The investigation found that the total phenolic content of the samples varied from 29.15±0.13 µmol GAE/100ml to 48.60±0.70 µmol GAE/100mL depending on the raw material utilized in manufacturing ($P<0.05$). ASK5 samples had a higher total phenolic content, while control samples had lower values ($P<0.05$). Furthermore, the total phenolic values of the samples, especially in the sample with rosemary essential oil and guava peel (ASK5), were found to be at the highest level. This result is likely influenced by the higher presence of phenolic compounds in the fruit peel.

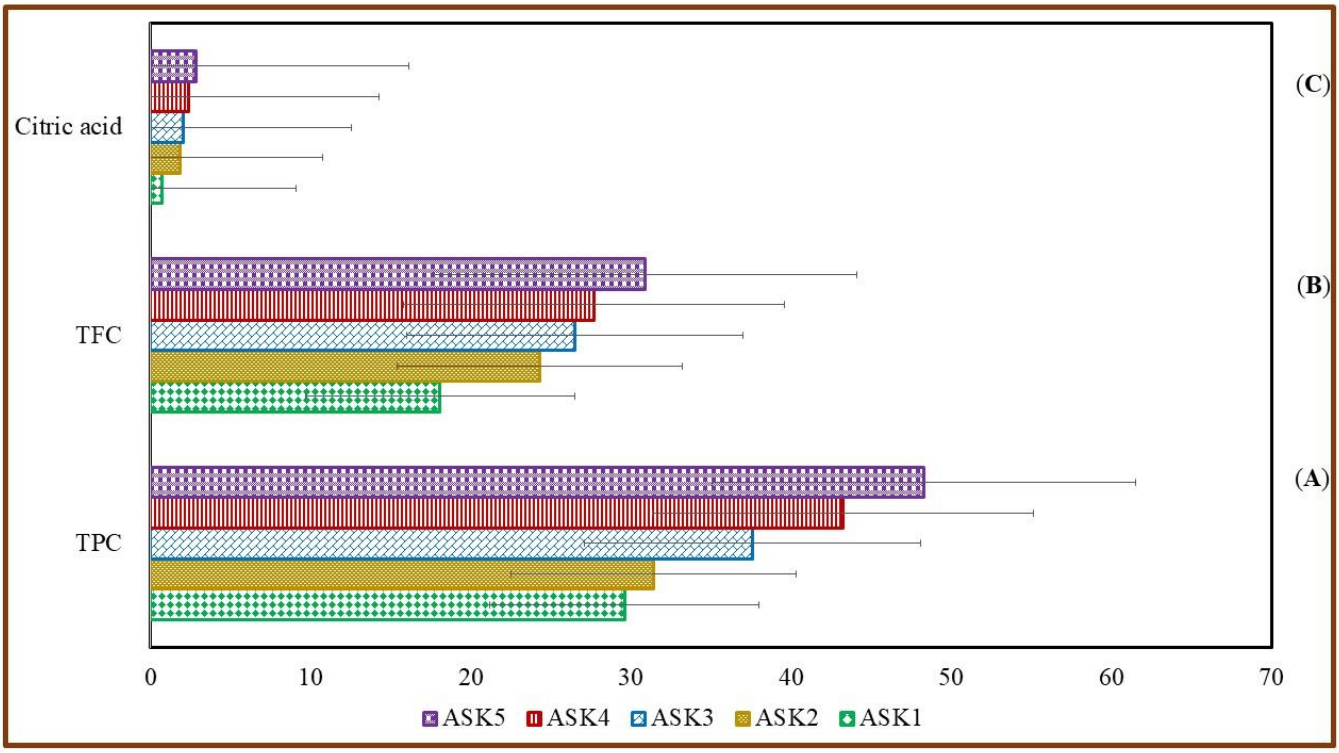


Figure 2- (A) Total phenolic content (TPC; µmol GAE/100mL); (B) Total flavonoid content (TFC; mg CE/100mL), and (C) citric acid cg/mL) values of the control kombucha (ASK1), kombucha infused with guava pulp (ASK2), kombucha infused with guava peel (ASK3), kombucha infused with guava pulp and 1% REO (ASK4), and kombucha infused with guava peel and 1% REO (ASK5)

3.4. Total flavonoid content (TFC)

The TFC ranged from 16.87±1.73 to 30.12±0.04 mg CE/100 mL (Figure 2B). The Kombucha samples containing guava peel and rosemary essential oil had the highest total flavonoid concentration, while the control samples had the lowest ($P<0.05$). The flavonoid content of ASK1, ASK2, ASK3, ASK4, and ASK5 was determined to be 16.87±1.73, 23.88±0.59, 26.06±0.60, 27.13±0.92, and 30.12±0.04 mg CE/100 mL, respectively. Several researchers have evaluated the total flavonoid content of kombucha made from diverse raw materials in the literature. These investigations found that flavonoid content can vary based on the raw material used in manufacture. Gamboa-Gomez et al. (2016) found that kombucha made with *Litsea glaucescens* had a lower total flavonoid content of 110.6 mg CE/L to 66.2 µg/mL, while kombucha made with *Eucalyptus camaldulensis* had a higher total flavonoid content of 29.3 µg/mL to 53.4 µg/mL after 7 days of fermentation at 25 °C. The total flavonoid contents of the samples ASK1, ASK2, ASK3, ASK4, and ASK5 on the 10th day of fermentation were determined to be 9.10±0.14, 13.54±1.53, 15.92±0.78, 20.38±1.29, and 21.86±1.31 (mg CE/100 mL) respectively. The inclusion of *Anthriscus sylvestris* (L.) plant, guava peel, and rosemary essential oil components in the formulation of sample ASK5 may have contributed to the high flavonoid content of this sample. The flavonoid content found in the kombucha sample with *Eucalyptus camaldulensis*, as

reported by Gamboa-Gomez et al. (2016) after 7 days of fermentation, was found to be slightly higher than the values obtained in this study on the 10th day of fermentation. This disparity in values is likely attributed to the differences in the materials used in the production of kombucha.

3.5. Citric organic acid

Figure 2C presents the fluctuations in citric acid concentrations in kombucha over time. Fermentation led to a considerable rise in citric acid concentration ($P<0.05$). After 15 days of storage, the acetic acid values of ASK1, ASK2, ASK3, ASK4, and ASK5 kombucha samples were 0.73±0.02, 1.82±0.04, 2.09±0.04, 2.39±0.02, and 2.83±0.09 g/mL, respectively. Fermentation resulted in a considerable rise in vitamin C levels ($P<0.05$). Generally, kombucha drinks contain the most vitamin C (15.19 mg/L for black tea and 7.88 mg/L for green tea kombucha) (Malbaša et al. 2011; Jayabalan et al. 2014; Geraris Kartelias et al. 2023). According to Hraš et al. (2000), citric acid may enhance the antioxidant properties of rosemary extract. Guava fruit also has 68 calories per 100 grams, is high in fiber, and has four times more vitamin C than oranges. The majority of the vitamin C is located in the fruits' skin, which is typically not peeled before ingestion (Mercadante et al. 1999). Citric acid can be produced by yeast and degraded by cells (Ye et al. 2014). The ASK5 sample made with guava peel, rosemary essential oil, and *Anthriscus sylvestris* (L.) had a greater citric acid value than the control sample. These findings suggest that the impact of fermentation on citric acid levels may vary depending on the specific plant species.

3.6. Antioxidant characteristics

The antioxidant capacity of various kombucha samples is shown in Figure 3A-E. Antioxidant activity increased significantly ($P<0.05$) with the addition of guava fruit, rosemary essential oil, and *Anthriscus sylvestris* (L.) compared to the control sample. The highest DPPH, ABTS^{*+}, FRAP, CUPRAC, and ORAC values were 78.90±0.08%, 97.25±1.57%, 723.15±2.61 µmol Trolox/mL, 460.45±0.58 µmol Trolox/mL, and 737.50±3.60 µmol Trolox/mL, respectively. The highest DPPH was determined in ASK5 (78.90±0.08%) and in ASK4 (77.15±0.24%). Using guava fruit and rosemary essential oil along with *Anthriscus sylvestris* (L.) further increased the DPPH value of ASK5 and ASK4 kombucha. The ASK1 sample (66.89±1.40%) had the lowest DPPH value (Figure 3A). The DPPH radical scavenging capacity of kombucha varies based on processing methods and polyphenol content, with yellow and green tea containing more physiologically active components (Jayabalan et al. 2008; Malbaša et al. 2011). Guava and rosemary essential oils have been shown to boost the bioactive potential of kombucha drinks. According to the antioxidant capacity results determined by ABTS$^{\bullet+}$, the highest values were detected in the ASK5, ASK4, ASK3, ASK2, and ASK1 samples as 97.25±1.57, 95.96±0.70, 93.05±0.72, 92.80±0.91, and 70.83±0.26%, respectively. All demonstrated a favorable link between higher guava peel and REO. Furthermore, as previously demonstrated, the observed increase in antioxidant capacity was favourably related to increased amounts of polyphenols and flavonoids (Abduh et al. 2023). Each kombucha formulation has significantly different FRAP values compared to the control ($P<0.05$). The ASK5 sample had the greatest FRAP value (723.15±2.61 µmol Trolox/mL). The ASK1 kombucha exhibited the lowest FRAP value (367±3.81 µmol Trolox/mL) (Figure 3C). The ASK1 sample was fermented with just sugar, and SCOBY was found to have the lowest bioactive potential ($P<0.05$). Significant variations were observed across kombucha tea samples in terms of ASK1 and ASK5, as well as the bioaccessibility (%) of CUPRAC values ($P<0.05$). CUPRAC values for ASK1, ASK2, ASK3, ASK4, and ASK5 samples were 381.40±1.33, 394.90±0.09, 414.80±2.43, 430.2430.08±1.02, and 460.45±0.58 µmol Trolox/mL, respectively. It was thought that the criteria impacting these changes in the composition of kombucha tea types (growing conditions, maturity duration, agricultural procedure, and climatic impact), guava fruit, rosemary essential oil, and tea leaf processing technique. Figure 3E shows the ORAC antioxidant activity. ORAC capacity was determined to be between 486.50±1.69 µmol Trolox/mL and 737.50±3.60 µmol Trolox/mL. The ORAC values of the kombucha beverages were statistically significant ($P<0.05$). The highest ORAC values were found in the ASK5 sample. Abuduaibifu & Tamer (2019) reported that DPPH, CUPRAC, and FRAP activities of black tea kombucha samples were 93.99±0.24, 56.08±0.71, and 81.77±1.48 µmol Trolox/g w.s.d.m, respectively. Previous research has demonstrated that increased quantities of polyphenols and flavonoids are connected with an increase in antioxidant capacity (Abduh et al. 2023). Furthermore, the presence of lactic acid bacteria in SCOBY may release many bioactive substances, such as exopolysaccharides, which contribute to the free radical scavenging ability of kombucha (Amjadi et al. 2023). Kombucha tea's health advantages are also due to its antioxidant content. Thus, kombucha samples, which are *Anthriscus sylvestris* (L.), guava, and rosemary oil, have the potential to be antioxidant-rich products, meaning that they may give health advantages by neutralizing free radicals.

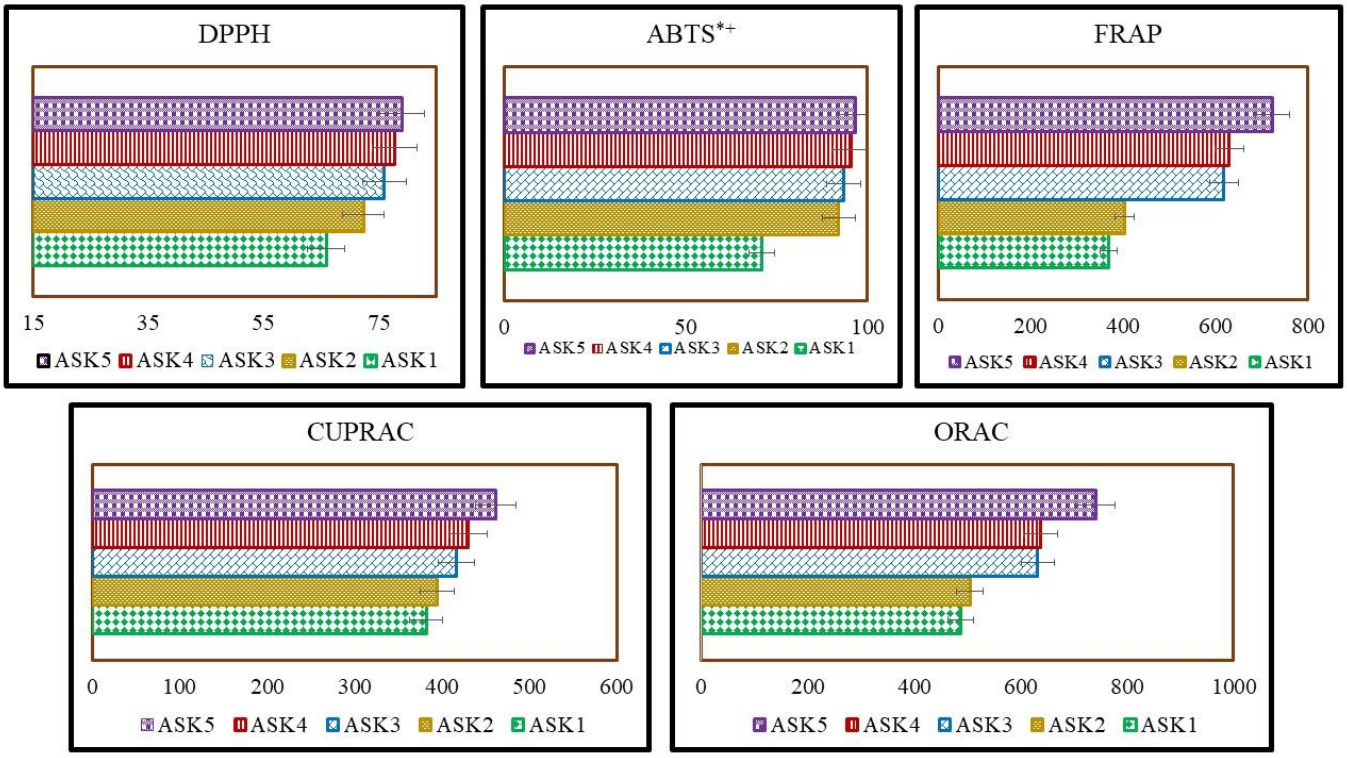


Figure 3- (A) DPPH (%); (B) ABTS*$^{+}$ (%); (C) FRAP (µmol Trolox/mL); (D) CUPRAC (µmol Trolox/mL); (E) ORAC (µmol Trolox/mL) values of the control kombucha (ASK1), kombucha infused with guava pulp (ASK2), kombucha infused with guava peel (ASK3), kombucha infused with guava pulp and % REO (ASK4), and kombucha infused with guava peel and 1% REO (ASK)

3.7. Color parameters

Table 3 displays statistically significant variations (P<0.05) in the color changes of all samples. All kombucha beverages showed statistically significant *L**, *a**, and *b** values during the fermentation process (P<0.05). The control kombucha with only *Anthriscus sylvestris* (L.) and it had the greatest *L** value after storage according to the raw material used in the production. ASK5 kombucha had the greatest *a** value after 10 days of storage. Adding guava peel and REO to kombucha teas resulted in significant differences across groups (P<0.05). On the 15th day of fermentation, the ASK2 group had the lowest *a** value (-0.53), while the ASK5 group had the highest (-4.07). Chemical alterations to phenolic molecules, particularly anthocyanins and carotenoids, affect their color diversity values. However, the inclusion of plant infusions reduced these values. According to the research of Chakravorty et al. (2016), during fermentation, thearubigin fractions might be completely or partially transformed into theaflavin, which may have been effective in converting the kombucha beverage's color to a reddish-light brown. Abuduaibifu & Tamer (2019) analyzed the *L**, *a**, and *b** values of black tea with kombucha culture after 48 h of fermentation at 28±2 °C. The results were 4.72±0.23, 5.91±0.19, and 4.52±0.27, respectively.

3.8. Antimicrobial activities

Kombucha's antibacterial activities have been extensively explored and recorded, both in control kombucha and in those containing *Anthriscus sylvestris* (L.) extracts and guava fruit (Table 4). Kombucha's antibacterial characteristics come from the presence of organic acids, plant-derived phenolic chemicals, enzymes, proteins, and bacteriocins (Battikh et al. 2013). This study observed antibacterial effects in kombucha against gram-negative bacteria (*Escherichia coli* ATCC 25922 and *Pseudomonas aeruginosa* ATTC 27853) and gram-positive bacteria (*Enterococcus faecalis* ATCC 29212 and *Staphylococcus aureus* ATCC 25923). Furthermore, it was shown that samples containing rosemary essential oil (ASK4 and ASK5) exhibited higher antibacterial activity. However, kombucha made with guava peel had a stronger antibacterial impact than kombucha made with guava pulp. The activity was increased with the guava peel against both *Enterococcus faecalis* ATCC 29212 (zone of inhibition 13.05 mm) and *Pseudomonas aeruginosa* ATTC 27853 (zone of inhibition 13.67 mm) strains. This is probably due to the existence of acetic acid, which has been recognized as the primary antibacterial substance in kombucha. Numerous bioactive compounds or metabolites created through the fermentation process, such as flavonoids, bacteriocins, polyphenols, and enzymes, could potentially be the cause of kombucha's antibacterial qualities. Understanding the specific mechanisms responsible for kombucha's antibacterial properties enables the strategic use of this natural antimicrobial ability in the creation of functional foods, allowing the development of products that provide consumers with both culinary and health benefits. Furthermore, kanamycin demonstrated effectiveness against all tested bacteria. In light of these findings, it is recommended that kanamycin be considered as the antibiotic of choice for selection. In their 21-day fermentation research, Pure & Pure (2016) utilized

Escherichia coli and *Staphylococcus aureus* but did not detect antibacterial properties in the traditional kombucha and other fermented samples they examined. These results differ from the current study findings, indicating that the choice of raw materials may have influenced this outcome.

3.9. Inhibition activities of α-glucosidase and α-amylase

In dietary starch, the hydrolase inhibitors block α-amylase and α-glucosidase, decreasing glucose assimilation. These two enzymes on the small intestine's boundary are vital for carbohydrate hydrolysis. Preventing the action of these digestive enzymes decreases glucose absorption into the circulatory system by dropping the level of glucose in the body (Geraris Kartelias et al. 2023). Figure 4A-B shows the inhibitory qualities of α-glucosidase and α-amylase. The ASK5 sample inhibited α-glucosidase and α-amylase by 115.08% and 86.7%, respectively, compared to 75.3% and 67.5% in the control group. These results collectively affirm that the total acidity and pH levels of kombucha beverages are subject to variation based on the specific raw materials, SCOBY utilized in production, and the fermentation conditions employed.

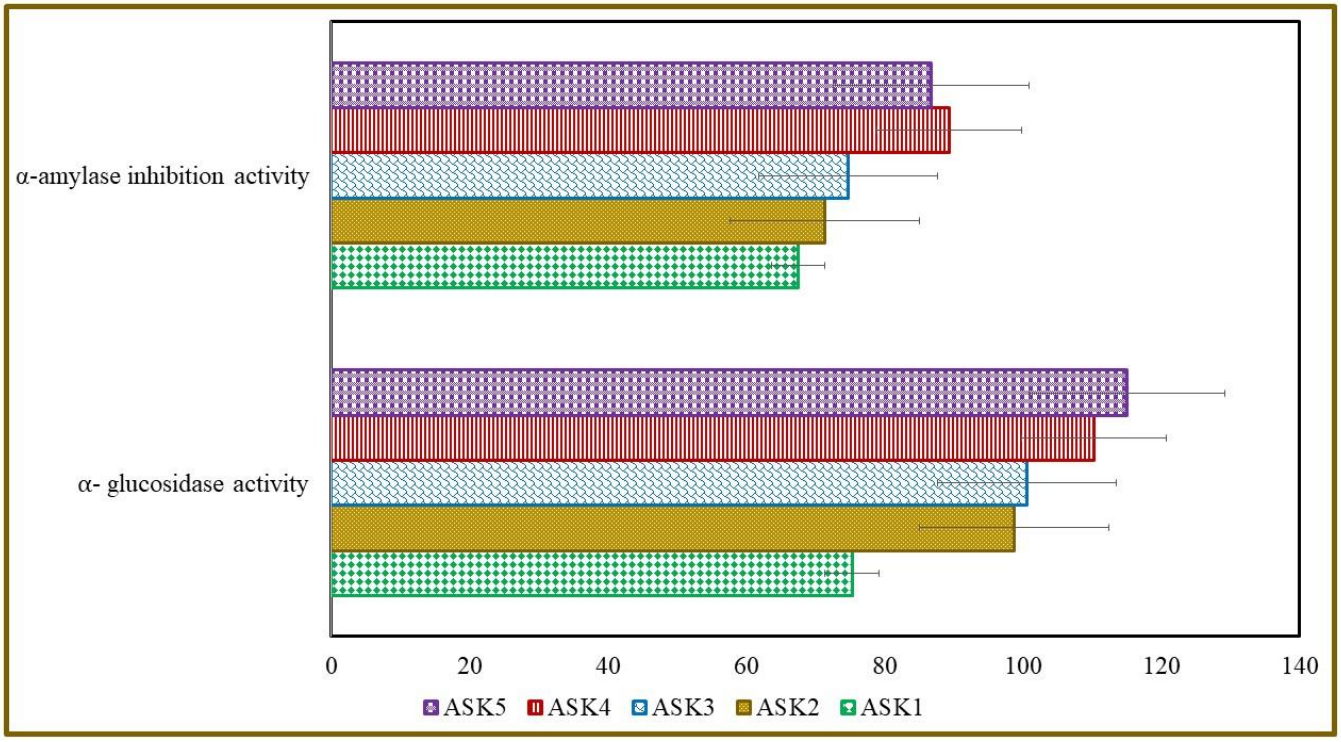


Figure 4- (A) α-glucosidase activity (%); (B) α–amylase inhibition activity (%) of the control kombucha (ASK1), kombucha inused with guava pulp (ASK2), kombucha infused with guava peel (ASK3), kombucha infused with guava pulp and 1% REO (ASK4), and kombucha infused with guava peel and 1% REO (ASK5)

The α-glucosidase enzyme inhibition assay measures the sample's inhibitory capacity against the enzyme and is expressed as a percentage of inhibition. The α-glucosidase is the most important enzyme in the absorption of carbohydrates, accelerating the decomposition of disaccharides and oligosaccharides involving maltose, maltotriose, and α-dextrins in the small intestinal tract, causing glucose (Proença et al. 2017). Both *Anthriscus sylvestris* (L.) and guava have inhibitory capabilities against the α-glucosidase enzyme. However, guava peel and rosemary essential oil outperformed kombucha in this regard. Thus, it improves glucose metabolism, glycaemic management, and insulin sensitivity (Koh et al. 2010). Wang et al. (2018) observed that fermentation can considerably boost the inhibitory efficiency of guava leaf tea against α-glucosidase. Inhibiting the carbohydrate metabolism enzymes α-glucosidase and α-amylase has been shown to effectively treat diabetes and hypertension (Ortiz-Andrade et al. 2007; Ademiluyi & Oboh 2013). Phenolic chemicals can regulate carbohydrate and lipid metabolism by inhibiting α-glucosidase and α-amylase due to their ability to chelate, modify structure, and limit enzyme performance (Lin et al. 2016; Nagappan et al. 2017).

3.10. Sensory properties

ASK1, ASK2, ASK3, ASK4, and ASK5 were sensory evaluated during 15 days of fermentation. Kombucha samples were evaluated in terms of appearance, color, aroma, flavor, sourness and general acceptability. There was no significant difference in color, appearance, aroma, sourness, or overall acceptance values among the kombucha samples. Guava pulp-fermented kombucha scored considerably higher on all sensory measures compared to guava peel kombucha ($P<0.05$). ASK2 and ASK3 had better overall sensory evaluation scores than ASK4 and ASK5, as seen in Figure 5. The findings suggest that guava pulp and

peel-derived kombucha, notably ASK2 and ASK3, may appeal to consumers because of their superior sensory properties. The kombucha drink made with guava pulp and rosemary (ASK5) did not receive favourable feedback after the 10-day storage period, and the sample prepared with guava pulp and rosemary (ASK4) was also not well-received after the 15th day of fermentation. Consequently, sensory analysis was terminated at the end of the 15-day fermentation period.

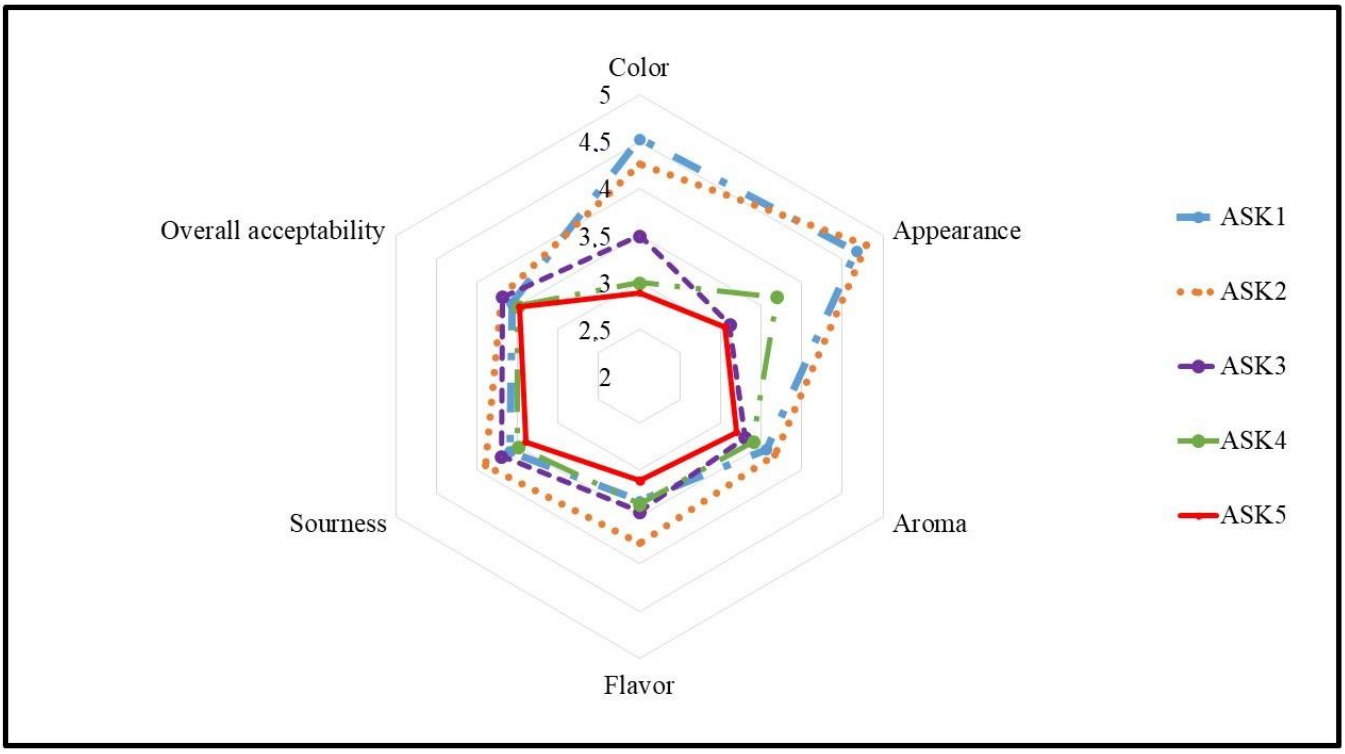


Figure 5- The sensory property assessment of the control kombucha (ASK1), kombucha infused with guava pulp (ASK2), kombucha infused with guava peel (ASK3), kombucha infused with guava pulp and 1% REO (ASK4), and kombucha infused with guava peel and 1%REO (ASK5)

4. Conclusions

Anthriscus sylvestris (L.)-based kombucha has more bioactive components and functional characteristics than ordinary kombucha. The inclusion of antioxidant and antimicrobial components in *Anthriscus sylvestris* (L.) kombucha resulted in enhanced total phenolic content and better antioxidant and antibacterial activity. Kombucha tea outperformed other samples in inhibiting the α-glucosidase enzyme, indicating medicinal promise in diabetes. Based on the higher presence of phenolic compounds in the sample (ASK5) produced with peel, it has been determined that they exhibit stronger antioxidant activity, as expected. The phenolic and flavonoid levels increased significantly, as did antioxidant capacity and a-amylase inhibitory action. Furthermore, the sensory evaluation revealed that guava pulp kombucha has a higher overall acceptance than *Anthriscus sylvestris* (L.) kombucha, emphasizing its palatability. Taken together, these findings highlight the potential of *Anthriscus sylvestris* (L.) leaves combined with guava fruit and REO as a unique and healthy ingredient for making kombucha tea with a variety of health benefits.

Author contributions Conceptualization: F.Y. Investigation: F.Y., M.D.G. Data curation: F.Y., M.D.G. Formal analysis: F.Y., M.D.G. Methodology: F.Y. Validation: F.Y. Writing - original draft: F.Y. Writing - review & editing: F.Y.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Ethical approval Ethics approval was not required for this research.

References

Abduh M Y, Nofitasari D, Rahmawati A, Eryanti A Y & Rosmiati M (2023). Effects of brewing conditions on total phenolic content, antioxidant activity and sensory properties of cascara. Food Chemistry Advances 2: 100183 https://doi.org/10.1016/j.focha.2023.100183

Abuduaibifu A & Tamer C E (2019). Evaluation of physicochemical and bioaccessibility properties of goji berry kombucha. *Journal of Food Processing and Preservation* 43(9): e14077. https://doi.org/10.1111/jfpp.14077

Ademiluyi A O & Oboh G (2013). Aqueous extracts of Roselle (Hibiscus sabdariffa Linn.) varieties inhibit α-amylase and α-glucosidase activities in vitro. *Journal of Medicinal Food* 16(1): 88-93. https://doi.org/10.1089/jmf.2012.0004

Alara O R, Abdurahman N H & Ukaegbu C I (2021). Extraction of phenolic compounds: A review. Current Research in Food Science 4: 200-214. https://doi.org/10.1016/j.crfs.2021.03.011

Amjadi S, Armanpour V, Ghorbani M, Tabibiazar M, Soofi M & Roufegarinejad L (2023). Determination of phenolic composition, antioxidant activity, and cytotoxicity characteristics of kombucha beverage containing Echium amoenum. *Journal of Food Measurement and Characterization* pp. 1-11. https://doi.org/10.1007/s11694-023-01856-1

AOAC (2000). (17th ed.), Microbiological methods, Vol. I, Official Methods of Analysis of AOAC International, USA (2000).

Apak R, Güçlü K, Özyürek M & Karademir S E (2004). Novel total antioxidant capacity index for dietary polyphenols and vitamins C and E, using their cupric ion reducing capability in the presence of neocuproine: CUPRAC method. *Journal of Agricultural and Food Chemistry* 52(26): 7970-7981. https://doi.org/10.1021/jf048741x

Araya J, Esquivel M, Jimenez G, Navia D & Poveda L (2022). Antimicrobial activity and physicochemical characterization of thermoplastic films based on bitter cassava starch, nanocellulose and rosemary essential oil. *Journal of Plastic Film & Sheeting* 38(1): 46-71. https://doi.org/10.1177/87560879211023882

Asl R M Z, Niakousari M, Gahruie H H, Saharkhiz M J, Khaneghah & A M (2018). Study of two-stage ohmic hydro-extraction of essential oil from Artemisia aucheri Boiss: Antioxidant and antimicrobial characteristics. Food Research International 107: 462-469. https://doi.org/10.1016/j.foodres.2018.02.059

Ayed L& Hamdi M (2015). Manufacture of a beverage from cactus pear juice using "tea fungus" fermentation. Annals of Microbiology, 65(4): 2293-2299. https://doi.org/10.1007/s13213-015-1071-8

Basli A, Belkacem N & Amrani I (2017) Health benefits of phenolic compounds against cancers. Phenolic Compounds-Biological Activity, 10: 5772-67232. https://dx.doi.org/10.5772/67232

Battikh H, Chaieb K, Bakhrouf A & Ammar E (2013). Antibacterial and antifungal activities of black and green kombucha teas. *Journal of Food Biochemistry* 37(2): 231-236. https://doi.org/10.1111/j.1745-4514.2011.00629.x

Benzie I F& Strain J J (1999). Ferric reducing/antioxidant power assay: direct measure of total antioxidant activity of biological fluids and modified version for simultaneous measurement of total antioxidant power and ascorbic acid concentration. In Methods in Enzymology, 299: 15-27. https://doi.org/10.1016/S0076-6879(99)99005-5

Berežni S, Mimica-Dukić N, Domina G, Raimondo F M, Orčić D (2024). Anthriscus sylvestris—Noxious Weed or Sustainable Source of Bioactive Lignans?. Plants 13(8): 1087. https://doi.org/10.3390/plants13081087

Burgin A B, Magnusson O T, Singh J, Witte P, Staker B L, Bjornsson J M & Gurney M E (2010). Design of phosphodiesterase 4D (PDE4D) allosteric modulators for enhancing cognition with improved safety. Nature Biotechnology 28(1): 63-70. https://doi:10.1038/nbt.1598

Chakrabortya I & Athmaselvi K A (2014). Changes in physicochemical properties of guava juice during ohmic heating. *Journal of Ready to Eat Food* 1(4): 152-157

Chakravorty S, Bhattacharya S, Chatzinotas A, Chakraborty W, Bhattacharya D & Gachhui R (2016). Kombucha tea fermentation: Microbial and biochemical dynamics. *International Journal of Food Microbiology* 220: 63-72. https://doi.org/10.1016/j.ijfoodmicro.2015.12.015

Chen H, Jiang H Z, Li Y C, Wei G Q, Geng Y & Ma C Y (2014). Antitumor constituents from Anthriscus sylvestris (L.) Hoffm. *Asian Pacific Journal of Cancer Prevention* 15(6): 2803-2807. https://doi.org/10.7314/APJCP.2014.15.6.2803

Chusak C, Thilavech T & Adisakwattana S (2014). Consumption of Mesona chinensis attenuates postprandial glucose and improves antioxidant status induced by a high carbohydrate meal in overweight subjects. *The American Journal of Chinese Medicine* 42(02): 315-336. https://doi.org/10.1142/S0192415X14500219

Correa M G & Couto J S (2016) Anticancer properties of Psidium guajava-a mini-review. *Asian Pacific Journal of Cancer Prevention* 17(9): 4199-4204

De Filippis F, Troise A D, Vitaglione P& Ercolini D (2018). Different temperatures select distinctive acetic acid bacteria species and promotes organic acids production during Kombucha tea fermentation. Food Microbiology 73: 11-16. https://doi.org/10.1016/j.fm.2018.01.008

de Miranda J F, Ruiz L F, Silva C B, Uekane T M, Silva K A, Gonzalez A G M & Lima A R (2022). Kombucha: A review of substrates, regulations, composition, and biological properties. *Journal of Food Science* 87(2): 503-527. https://doi.org/10.1111/1750-3841.16029

De Vero L, Gala E, Gullo M, Solieri L, Landi S & Giudici P (2006). Application of denaturing gradient gel electrophoresis (DGGE) analysis to evaluate acetic acid bacteria in traditional balsamic vinegar. Food Microbiology 23(8): 809-813. https://doi.org/10.1016/j.fm.2006.01.006

Delgado A M, Issaoui M & Chammem N (2019). Analysis of main and healthy phenolic compounds in foods. *Journal of AOAC International* 102(5): 1356-1364. https://doi.org/10.1093/jaoac/102.5.1356

Du C, Li Z, Zhang J, Yin N, Tang L, Li J & Chen X (2023). The protective effect of carnosic acid on dextran sulfate sodium-induced colitis based on metabolomics and gut microbiota analysis. Food Science and Human Wellness 12(4): 1212-1223. https://doi.org/10.1016/j.fshw.2022.10.003

Gamboa-Gómez C I, González-Laredo R F, Gallegos-Infante J A, Pérez M D M L, Moreno-Jiménez M R, Flores-Rueda A G & Rocha-Guzmán N E (2016). Antioxidant and angiotensin-converting enzyme inhibitory activity of Eucalyptus camaldulensis and Litsea glaucescens infusions fermented with kombucha consortium. Food Technology and Biotechnology 54(3): 367. https://doi.org/10.17113/ftb.54.03.16.4622

Gao J, Xu P, Wang Y, Wang Y & Hochstetter D (2013). Combined effects of green tea extracts, green tea polyphenols or epigallocatechin gallate with acarbose on inhibition against α-amylase and α-glucosidase in vitro. Molecules 18(9): 11614-11623. https://doi.org/10.3390/molecules180911614

Geraris Kartelias I, Karantonis H C, Giaouris E, Panagiotakopoulos I & Nasopoulou C (2023). Kombucha Fermentation of Olympus Mountain Tea (Sideritis scardica) Sweetened with Thyme Honey: Physicochemical Analysis and Evaluation of Functional Properties. Foods, 12(18): 3496. https://doi.org/10.3390/foods12183496

He X, Zhang M, Li S T, Li X, Huang Q, Zhang K & Ma Y Y (2022). Alteration of gut microbiota in high-fat diet-induced obese mice using carnosic acid from rosemary. Food Science & Nutrition 10(7): 2325-2332. https://doi.org/10.1002/fsn3.2841

Hraš A R, Hadolin M, Knez Ž & Bauman D (2000). Comparison of antioxidative and synergistic effects of rosemary extract with α-tocopherol, ascorbyl palmitate and citric acid in sunflower oil. Food Chemistry, 71(2): 229-233. https://doi.org/10.1016/S0308-8146(00)00161-8

Jaiarj P, Khoohaswan P, Wongkrajang Y, Peungvicha P, Suriyawong P, Saraya M S& Ruangsomboon O (1999). Anticough and antimicrobial activities of Psidium guajava Linn. leaf extract. *Journal of Ethnopharmacology* 67(2): 203-212. https://doi.org/10.1016/S0378-8741(99)00022-7

Janković M, Berežni S& Orčić D (2024). Quantitative Analysis of Lignans from the Fruits of Wild Chervil (Anthriscus sylvestris (L.) Hoffm.). Facta Universitatis, Series: Physics, Chemistry and Technology, 039-046. https://doi.org/10.2298/FUPCT2301039J

Jayabalan R, Malbaša, R V, Lončar E S, Vitas J S & Sathishkumar M (2014). A review on kombucha tea-microbiology, composition, fermentation, beneficial effects, toxicity, and tea fungus. Comprehensive Reviews in Food Science and Food Safety 13(4): 538–550. https://doi.org/10.1111/1541-4337.12073

Jayabalan R, Subathradevi P, Marimuthu S, Sathishkumar M& Swaminathan K (2008). Changes in free-radical scavenging ability of kombucha tea during fermentation. Food Chemistry 109(1): 227-234. https://doi.org/10.1016/j.foodchem.2007.12.037

Jiang B & Zhang Z W (2012). Comparison on phenolic compounds and antioxidant properties of Cabernet Sauvignon and Merlot wines from four wine grape-growing regions in China. Molecules 17(8): 8804-8821. https://doi.org/10.3390/molecules17088804

Kamath J V, Rahul N, Kumar C A & Lakshmi S M (2008). Psidium guajava L: A review. *International Journal of Green Pharmacy* (IJGP) 2(1)

Kaur H & Ghosh M (2023). Probiotic fermentation enhances bioaccessibility of lycopene, polyphenols and antioxidant capacity of guava fruit (Psidium guajava L). *Journal of Agriculture and Food Research* 14: 100704. https://doi.org/10.1016/j.jafr.2023.100704

Khazi M I, Liaqat F, Liu X, Yan Y & Zhu D (2024). Fermentation, functional analysis, and biological activities of turmeric kombucha. *Journal of the Science of Food and Agriculture* 104(2): 759-768. 10.1002/jsfa.12962

Kim J & Adhikari K (2020). Current trends in kombucha: Marketing perspectives and the need for improved sensory research. Beverages 6(1): 15. https://doi.org/10.3390/beverages6010015

Kitwetcharoen H, Phung L T, Klanrit P, Thanonkeo S, Tippayawat P, Yamada M & Thanonkeo P (2023). Kombucha healthy drink-recent advances in production, chemical composition and health benefits. Fermentation 9(1): 48. https://doi.org/10.3390/fermentation9010048

Koh K K, Quon M J, Han S H, Lee Y, Kim S J & Shin E K (2010). Atorvastatin causes insulin resistance and increases ambient glycemia in hypercholesterolemic patients. *Journal of the American College of Cardiology* 55(12): 1209-1216. https://doi:10.1016/j.jacc.2009.10.053

Koulman A, Bos R, Medarde M, Pras N & Quax W J (2001). A fast and simple GC MS method for lignan profiling in Anthriscus sylvestris and biosynthetically related plant species. Planta Medica 67(09): 858-862. https://doi.org/10.1055/s-2001-18849

König A, Schwarzinger B, Stadlbauer V, Lanzerstorfer P, Iken M, Schwarzinger C& Weghuber J (2019). Guava (Psidium guajava) fruit extract prepared by supercritical CO2 extraction inhibits intestinal glucose resorption in a double-blind, randomized clinical study. Nutrients 11(7): 1512. https://doi.org/10.3390/nu11071512

Leonarski E, Guimarães A C, Cesca K & Poletto P (2022). Production process and characteristics of kombucha fermented from alternative raw materials. Food Bioscience 49: 101841. https://doi.org/10.1016/j.fbio.2022.101841

Li Y, Wang S, Lei D, He Y B, Li B & Kang F (2017). Acetic acid-induced preparation of anatase TiO2 mesocrystals at low temperature for enhanced Li-ion storage. *Journal of Materials Chemistry A* 5(24): 12236-12242. https://doi.org/10.1039/C7TA02361H

Li H, Sun J J, Chen G Y, Wang W W, Xie Z T, Tang G F & Wei S D (2016). Carnosic acid nanoparticles suppress liver ischemia/reperfusion injury by inhibition of ROS, Caspases and NF-κB signaling pathway in mice. Biomedicine & Pharmacotherapy 82: 237-246. https://doi.org/10.1016/j.biopha.2016.04.064

Lin D, Xiao M, Zhao J, Li Z, Xing B, Li X & Chen S (2016). An overview of plant phenolic compounds and their importance in human nutrition and management of type 2 diabetes. Molecules 21(10): 1374. https://doi.org/10.3390/molecules21101374

Liu Y L, Cao Y G, Niu Y, Zheng Y J, Chen X, Ren Y J & Feng W S (2023). Diarylpentanoids and phenylpropanoids from the roots of Anthriscus sylvestris (L.) Hoffm. Phytochemistry 216: 113865. https://doi.org/10.1016/j.phytochem.2023.113865

Malbaša R V, Lončar E S, Vitas J S & Čanadanović-Brunet J M (2011). Influence of starter cultures on the antioxidant activity of kombucha beverage. Food Chemistry 127(4): 1727-1731. https://doi.org/10.1016/j.foodchem.2011.02.048

Martin-Piñero M J, García M C, Santos J, Alfaro-Rodriguez M C & Muñoz J (2020). Characterization of novel nanoemulsions, with improved properties, based on rosemary essential oil and biopolymers. *Journal of the Science of Food and Agriculture* 100(10): 3886-3894. https://doi.org/10.1002/jsfa.10430

Mazidi S, Rezaei K, Golmakani M, Sharifan A & Rezazadeh S (2012). Antioxidant activity of essential oil from Black zira (Bunium persicum Boiss.) obtained by microwave-assisted hydro-distillation. *Journal of Agricultural Science and Technology* 14(5): 1013-1022

Mengoni E S, Vichera G, Rigano L A, Rodriguez-Puebla M L, Galliano S R, Cafferata E E & Vojnov A A (2011). Suppression of COX-2, IL-1β and TNF-α expression and leukocyte infiltration in inflamed skin by bioactive compounds from Rosmarinus officinalis L. Fitoterapia 82(3): 414-421. https://doi.org/10.1016/j.fitote.2010.11.023

Mercadante A Z, Steck A & Pfander H (1999). Carotenoids from guava (Psidium guajava L.): isolation and structure elucidation. *Journal of Agriculture and Food Chemistry* 47(1): 145-51. https://doi.org/10.1021/jf980405r

Miller N J, Rice-Evans C, Davies M J, Gopinathan V & Milner A (1993). A novel method for measuring antioxidant capacity and its application to monitoring the antioxidant status in premature neonates. Clinical Science (London, England: 1979), 84(4): 407-412. https://doi.org/10.1042/cs0840407

Morales D (2020). Biological activities of kombucha beverages: The need of clinical evidence. Trends in Food Science & Technology 105: 323-333. https://doi.org/10.1016/j.tifs.2020.09.025

Nagappan H, Pee P P, Kee S H Y, Ow J T, Yan S W, Chew L Y & Kong K W (2017). Malaysian brown seaweeds Sargassum siliquosum and Sargassum polycystum: Low density lipoprotein (LDL) oxidation, angiotensin converting enzyme (ACE), α-amylase, and α-glucosidase inhibition activities. Food Research International 99: 950-958. https://doi.org/10.1016/j.foodres.2017.01.023

Ninga K A, Sengupta S, Jain A, Desobgo Z S C, Nso E J & De S (2018). Kinetics of enzymatic hydrolysis of pectinaceous matter inguava juice. *Journal of Food Engineering* 221: 158-166. https://doi.org/10.1016/j.jfoodeng.2017.10.022

Oh W K, Lee C H, Lee M S, Bae E Y, Sohn C B, Oh H & Ahn J S (2005). Antidiabetic effects of extracts from Psidium guajava. *Journal of Ethnopharmacology* 96(3): 411-415. https://doi.org/10.1016/j.jep.2004.09.041

Olaru O T, Niţulescu G M, Orţan A & Dinu-Pîrvu C E (2015). Ethnomedicinal, phytochemical and pharmacological profile of Anthriscus sylvestris as an alternative source for anticancer lignans. Molecules 20(8): 15003-15022. https://doi.org/10.3390/molecules200815003

Olaru O T, Nițulescu G M, Orţan A, Băbeanu N, Popa O, Ionescu D & Dinupîrvu C E (2016). Polyphenolic content and toxicity assessment of Anthriscus sylyestris Hoffm. Romanian Biotechnological Letters 22(6): 12054.

Oliveira J T, da Costa F M, da Silva T G, Simões G D, dos Santos Pereira E, da Costa P Q & Pieniz S (2023). Green tea and kombucha characterization: Phenolic composition, antioxidant capacity and enzymatic inhibition potential. Food Chemistry 408: 135206. https://doi.org/10.1016/j.foodchem.2022.135206

Orčić D, Berežni S, Škorić D, Mimica-Dukić N (2021). Comprehensive study of Anthriscus sylvestris lignans. Phytochemistry 192: 112958. https://doi.org/10.1016/j.phytochem.2021.112958

Orčić D, Berežni S, Mimica-Dukić N (2022). Phytochemical and biochemical studies of wild chervil (Anthriscus sylvestris). Biologia Serbica 44(1). https://doi.org/10.5281/zenodo.7074977

Ortiz-Andrade R R, Garcia-Jimenez S, Castillo-Espana P, Ramirez-Avila G, Villalobos-Molina R, Estrada-Soto S (2007). α-Glucosidase inhibitory activity of the methanolic extract from Tournefortia hartwegiana: an anti-hyperglycemic agent. *Journal of Ethnopharmacology* 109(1): 48-53. https://doi.org/10.1016/j.jep.2006.07.002

Permatasari H K, Nurkolis F, Augusta P S, Mayulu N, Kuswari M, Taslim N A & Gunawan W B (2021). Kombucha tea from seagrapes (Caulerpa racemosa) potential as a functional anti-ageing food: in vitro and in vivo study. Heliyon 7(9). https://doi.org/10.1016/j.heliyon.2021.e07944

Proença C, Freitas M, Ribeiro D, Oliveira E F, Sousa J L, Tome S M & Fernandes E (2017). α-Glucosidase inhibition by flavonoids: an in vitro and in silico structure–activity relationship study. *Journal of Enzyme Inhibition and Medicinal Chemistry* 32(1): 1216-1228. https://doi.org/10.1080/14756366.2017.1368503

Pure A E & Pure M E (2016). Antioxidant and antibacterial activity of kombucha beverages prepared using banana peel, common nettles and black tea infusions. Applied Food Biotechnology 3(2): 125-130. https://doi.org/10.22037/afb.v3i2.11138

Rebaya A, Belghith S I, Baghdikian B, Leddet V M, Mabrouki F, Olivier E & Ayadi M T (2015). Total phenolic, total flavonoid, tannin content, and antioxidant capacity of Halimium halimifolium (Cistaceae). *Journal of Applied Pharmaceutical Science* 1(01): 052-057. https://doi.org/10.7324/JAPS.2015.50110

Sakar E H, Zeroual A, Kasrati A, Gharby S (2023). Combined effects of domestication and extraction technique on essential oil yield, chemical profiling, and antioxidant and antimicrobial activities of rosemary (Rosmarinus officinalis L.). *Journal of Food Biochemistry* 2023(1): 6308773. https://doi.org/10.1155/2023/6308773

Song X, Sui X & Jiang L (2023). Protection Function and Mechanism of Rosemary (Rosmarinus officinalis L.) Extract on the Thermal Oxidative Stability of Vegetable Oils. Foods 12(11): 2177. https://doi.org/10.3390/foods12112177

Somsong P, Tiyayon P & Srichamnong W (2017). Antioxidant of green tea and pickle tea product, miang, from northern Thailand. Paper presented at the IV Asia Symposium on Quality Management in Postharvest Systems 1210.

Sreeramulu G, Zhu Y & Knol W (2000). Kombucha fermentation and its antimicrobial activity. Journal of Agricultural and Food Chemistry 48(6): 2589-2594. https://doi.org/10.1021/jf991333m

Teoh A L, Heard G & Cox J (2004). Yeast ecology of Kombucha fermentation. *International Journal of Food Microbiology* 95(2): 119-126. https://doi.org/10.1016/j.ijfoodmicro.2003.12.020

Varela-Santos E, Ochoa-Martinez A, Tabilo-Munizaga G, Reyes J E, Pérez-Won M, Briones-Labarca V & Morales-Castro J (2012). Effect of high hydrostatic pressure (HHP) processing on physicochemical properties, bioactive compounds and shelf-life of pomegranate juice. Innovative Food Science & Emerging Technologies 13: 13-22. https://doi.org/10.1016/j.ifset.2011.10.009

Villarreal-Soto S A, Beaufort S, Bouajila J, Souchard J P, Renard T, Rollan S & Taillandier P (2019). Impact of fermentation conditions on the production of bioactive compounds with anticancer, anti-inflammatory and antioxidant properties in kombucha tea extracts. Process Biochemistry 83: 44-54. https://doi.org/10.1016/j.procbio.2019.05.004

Wang L, Luo Y, Wu Y, Liu Y & Wu Z (2018). Fermentation and complex enzyme hydrolysis for improving the total soluble phenolic contents, flavonoid aglycones contents and bio-activities of guava leaves tea. Food Chemistry 264: 189-198. https://doi.org/10.1016/j.foodchem.2018.05.035

Xia G, Wang X, Sun H, Qin Y& Fu M (2017). Carnosic acid (CA) attenuates collagen-induced arthritis in db/db mice via inflammation suppression by regulating ROS-dependent p38 pathway. Free Radical Biology and Medicine 108: 418-432. https://doi.org/10.1016/j.freeradbiomed.2017.03.023

Yavari N, Assadi M M, Moghadam M B & Larijani K (2011). Optimizing glucuronic acid production using tea fungus on grape juice by response surface methodology. *Australian Journal of Basic and Applied Sciences* 5(11): 1788-1794. http://www.insipub.com/ajbas/2011/November-2011/1788-1794.pdf

Ye M, Yue T & Yuan Y (2014). Evolution of polyphenols and organic acids during the fermentation of apple cider. *Journal of the Science of Food and Agriculture* 94(14): 2951-2957. https://doi.org/10.1002/jsfa.6639

Zhang L & Lu J (2024). Rosemary (Rosmarinus officinalis L.) polyphenols and inflammatory bowel diseases: Major phytochemicals, functional properties, and health effects. Fitoterapia, 106074. https://doi.org/10.1016/j.fitote.2024.106074

Zhao D & Shah N P (2014). Changes in antioxidant capacity, isoflavone profile, phenolic and vitamin contents in soymilk during extended fermentation. LWT – Food Science and Technology 58(2): 454–462. https://doi.org/10.1016/j.lwt.2014.03.029

Zou X & Liu H (2023). A review of meroterpenoids and of their bioactivity from guava (Psidium guajava L.). *Journal of Future Foods* 3(2): 142-154. https://doi.org/10.1016/j.jfutfo.2022.12.005

Zubaidah E, Afgani C A, Kalsum U, Srianta I & Blanc P J (2019). Comparison of in vivo antidiabetes activity of snake fruit Kombucha, black tea Kombucha and metformin. Biocatalysis and Agricultural Biotechnology 17: 465-469. https://doi.org/10.1016/j.bcab.2018.12.026

Zubaidah E, Apriyadi T E, Kalsum U, Widyastuti E, Estiasih T, Srianta I & Blanc P J (2018). In vivo evaluation of snake fruit Kombucha as hyperglycemia therapeutic agent. *International Food Research Journal* 25(1): 453-457. http://www.ifrj.upm.edu.my

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Evaluation of Pesticide Use in Greenhouse Tomato Production in the Context of Sustainability, Food Safety and Export Residue Notifications

Merve Mürüvvet Dag[a*], Hasan Yilmaz[a]

[a]*Department of Agricultural Economics, Faculty of Agriculture, Isparta Applied Sciences University, Isparta, TÜRKİYE*

ARTICLE INFO
Research Article
Corresponding Author: Merve Mürüvvet Dag, E-mail: mervedag@isparta.edu.tr
Received: 16 April 2024 / Revised: 16 July 2024 / Accepted: 29 July 2024 / Online: 14 January 2025

Cite this article
Dag M M, Yilmaz H (2025). Evaluation of Pesticide Use in Greenhouse Tomato Production in the Context of Sustainability, Food Safety and Export Residue Notifications. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):46-58. DOI: 10.15832/ankutbd.1469332

ABSTRACT
The sustainability of agricultural production and access to healthy food are important issues in the agricultural sector. From a food safety perspective, pesticide residue problems are becoming an increasing concern for consumers worldwide. This study aims to evaluate pesticide use in greenhouse tomato production in terms of sustainability, food safety and export residue notifications. In this research, original data obtained from a face-to-face survey conducted with 115 greenhouse tomato farmers in Kumluca district of Antalya province was used. It was found that farmers on the surveyed farms used pesticides containing 57 different active ingredients. Various types of pesticides were used, including 12.9 kg/ha of insecticides, 10.9 kg/ha of nematicides, 8.8 kg/ha of fungicides, and 3.8 kg/ha of acaricides, resulting in a total of 36.5 kg/ha of pesticide use. The study revealed that 55.88% of the pesticides used by farmers were classified as hazardous according to the World Health Organization's (WHO) classification. The increasing number of pesticide residue notifications from European Union countries in recent years serves as a warning for Turkish farmers. Therefore, Türkiye needs to make more efforts to comply with international standards for the sustainability of tomato production and exports. As a result, in the research area, more efforts should be made to increase farmer training activities, tighten controls in greenhouse production, and strengthen safe food control mechanisms in order to reduce the use of pesticides for sustainable agriculture and safe food. In addition, greenhouse production support policies should be redesigned to encourage farmers to use biological control methods.

Keywords: Tomato, Pesticide, Türkiye, Residue, Sustainable agriculture

1. Introduction

The agricultural sector is an important sector that contributes to the country's economy and plays a role in people's nutrition. Therefore, ensuring both sustainability in agricultural products and access to healthy food appear to be very important issues. In recent years, there has been an increasing demand not only for producing larger quantities but also for the development of sustainable agriculture, where production is environmentally friendly, socially fair, and economically beneficial. It has been emphasized that agricultural food production practices should be improved for all types (conventional, integrated, organic agriculture, etc.) of agriculture (Wezel et al. 2014).

After potato, tomato is the second most important vegetable produced worldwide (Pishgar-Komleh et al. 2017; FAO 2023). According to Maham et al. (2020), tomato is among the most widely produced vegetables globally. Tomato production not only provides income for many rural families but also serves as a significant source of employment (Çetin & Vardar 2008). In 2021, the total tomato production area worldwide was determined to be 5 167 388 hectares (ha). The countries with the highest share of global tomato production are China (22.15%), India (16.35%), Nigeria (16.35%), Türkiye (3.20%), and Egypt (2.90%). These countries account for 60.95% of the world's tomato production area (FAO 2023). In Türkiye, tomato is primarily produced in the Antalya province, covering 62.10% of the total production (TÜİK 2023).

Türkiye's top tomato export destinations are determined as follows: Syria (18.58%), Ukraine (13.23%), Romania (10.78%), Bulgaria (9.15%), and Israel (8.90%) (TRADE MAP 2022). Russia was the largest importer of tomato in 2019; however, tomato exports to Russia have gradually decreased since then. Sezgin (2022) expresses the expectation that Türkiye's tomato exports to these countries will diminish due to the conflict between Russia and Ukraine.

In a report published by the Ministry of Trade in 2021 within the framework of European Union harmonization laws in Türkiye, it was emphasized that due to Türkiye being among the countries most affected by climate change, ensuring sustainability in agriculture and the importance of the Green Deal Action Plan for uninterrupted agricultural trade with the EU

have been highlighted (Republic of Türkiye Ministry of Trade 2021). The Green Deal is a comprehensive policy framework by the European Union with the goal of becoming climate neutral by 2050. This action plan includes the target of climate neutrality by 2050, the widespread use of renewable energy sources, making industrial processes sustainable, promoting electric vehicles and clean fuels, reducing the use of pesticides and fertilizers, adopting sustainable farming practices, protecting ecosystems and biodiversity, and improving waste management (Szpilko & Ejdys 2022). In accordance with this target, efforts are being made to reduce pesticide and chemical fertilizer usage and to promote the widespread adoption of biological and biotechnical control methods. Regulations supporting certain sustainable agricultural practices indirectly encourage the reduction or proper and effective usage of pesticides.

Support for greenhouse production is promoted worldwide due to reasons such as countries' desire to ensure food security and food safety. However, the increasing use of intensive chemical inputs in greenhouse agriculture, reaching threatening levels for the environment and human health, and the growing concerns regarding the environment, food security, and human health have prompted governments to take certain measures (Yılmaz & Tanç 2019). Technologies used in greenhouse production and environmentally friendly agricultural practices are particularly supported by governments in developed countries (European Commission 2024). In developed countries, food security is considered part of preventive health policy and is given special importance (Wilson 1989). In Türkiye, the area and quantity of greenhouse production are constantly increasing. Supported by reasons such as providing significant agricultural export income for Türkiye, contributing to food security, and numerous farmers engaging in greenhouse production, greenhouse agriculture is encouraged to increase technology usage levels and implement environmentally friendly practices. Greenhouse agriculture in Türkiye is given special importance, especially due to greenhouse products being significant export commodities and food security issues, and various support policies are implemented by the government for different purposes using different tools (Anonymous 2024).

One of the significant food safety issues in agricultural production is pesticide residue problems in products. The purpose of pesticide use is to control and prevent plant pests and diseases. Improper use of pesticides in agricultural production, not adhering to the recommended dosage and techniques, leads to various adverse effects. Among these adverse effects are environmental pollution, health problems, and a decrease in soil and water quality. Furthermore, during pesticide application, certain pesticides may drift beyond the intended area of use, potentially causing harm to soil, water, and wildlife. There are many studies showing a positive relationship between pesticides and various diseases, such as leukemia, pancreatic cancer, breast cancer, etc., and their negative effects on the environment (Infante-Rivard et al. 1999; Alguacil et al. 2000; Dolapsakis et al. 2001; Bassil et al. 2007; Geiger et al. 2010; Kumar et al. 2012; Yılmaz, 2015a; Yılmaz 2015b; Mahmood et al. 2016; Sharma et al. 2019; Wang et al. 2024). These negative effects have prompted some countries to reduce pesticide use (Van Driesche & Bellows 2012).

It has been observed that there has been a certain increase in pesticide usage worldwide from the past to the present. Since 1990, there has been a 50.83% increase in pesticide usage globally. In Türkiye, this increase rate is 114.81%. Despite the relatively high increase in pesticide usage in Türkiye compared to 1990, it is observed that pesticide usage is lower than in developed countries. Pesticide usage per unit area was reported as 1.81 kg/ha globally and 3.18 kg/ha in the European Union in 2020. In Türkiye, the amount of pesticide usage per unit area is 2.32 kg/ha (FAOSTAT 2020).

Eyhorn et al. (2015) stated that with current knowledge, technology, and alternative production systems, pesticide usage can be reduced without compromising yield or increasing production costs. Additionally, they noted that various alternative methods for reducing pesticide usage are being developed worldwide. Among these methods are biological and biotechnical control, integrated pest management, the use of natural enemies, and sustainable farming techniques. In this context, it is important to first identify some information regarding farmers' pesticide usage. The study aims to determine the diseases and pests encountered by greenhouse tomato farmers, identify the types and amounts of pesticides used in tomato production, evaluate them in terms of sustainability, food safety and export residue notifications. Additionally, the hazard classes of pesticide types used in tomato production will be determined according to the classification of the World Health Organization (WHO), and the pesticide residue notifications made by the countries to which Türkiye exports the most tomatoes were also determined.

2. Material and Methods

The primary data for the study were obtained through face-to-face surveys conducted with greenhouse tomato farmers in the Kumluca district of Antalya province, Türkiye. Antalya province, which holds a significant position in tomato production in Türkiye (62.10%), was selected, and within it, Kumluca district, known for its intensive tomato production (16.30%), was chosen to represent the province. According to the data obtained from the Agriculture and Forestry District Directorate, there are 6 213 farmers engaged in greenhouse tomato cultivation in Kumluca district. A proportional approach was used to determine the sample size to best represent the target population. The number of farmers to be surveyed was calculated using the "proportional sampling" method (Miran 2010). The following formula was used to determine the sample size (Equation 1). As a result of the calculations, a total of 115 surveys were conducted in twenty-one villages, with a 95% confidence interval and an 8% rate of error. Consultations were held with personnel from the District Agriculture and Forestry Directorate to conduct the survey in settlements that would represent the greenhouse production activities and socio-economic aspects of Kumluca district. The greenhouse farms surveyed were selected randomly.

$$\mathbf{n} = \frac{\mathbf{Np(1-p)}}{\mathbf{(N-1)\sigma^2+p(1-p)}} \quad (1)$$

In the formula:
n = Sample Size,
N = Population Size (6213),
p= Estimation Rate (0.05 for maximum sample size),
$\boldsymbol{\sigma^2}$ = Population Variance.

Descriptive statistics (N, %, x̄, σ) were used in the analysis of the data. Additionally, one of the significant aims of the study was to classify the types of pesticides used in tomato production according to the toxicity classification of the WHO. The active ingredients of these pesticides and the licensing status of the drugs used in tomato production by the Republic of Türkiye Ministry of Agriculture and Forestry were also determined. Subsequently, an independent two-sample t-test was conducted to determine if there was a significant difference between the quantity of product loss in tomato production and the specific diseases and pests encountered by farmers. Before conducting the t-test, it was checked whether the data were normally distributed and whether the variances of the groups were homogeneous; after performing the necessary checks, the t-test was applied.

3. Results and Discussions

The socio-demographic characteristics measured in research are important parameters in interpreting producers' decisions. Table 1 provides the socio-demographic characteristics of the farmers. Accordingly, the average age of the farmers was found to be 49.6 years. It was determined that 87.83% of the farmers have social security. The average level of education for the farmers was determined to be 8 years. Additionally, it was found that 54.78% of the farmers have non-agricultural income (Table 1). According to an OECD report, in Türkiye, 58% of individuals aged 25–64 have completed primary education, 20% have completed secondary education, and 22% have completed higher education (OECD 2022). According to the United Nations Development Program's Human Development Report, the average level of education in the world and Türkiye in 2021 was reported as 8.6 years (UNDP 2022). The results of the study indicate that the level of education is consistent with both the world average and the average in Türkiye. Yılmaz (2014) determined the average age of farmers in greenhouse tomato production using bumblebees as 44.02 years, while it was 48.32 years in farms not using bumblebees. Additionally, the education level of farmers in farms using bumblebees was found to be 7.59 years, compared to 6.86 years in farms not using bumblebees. In another study, the average age of farmers practicing in greenhouse vegetable production was found to be 40.45 years, with an average education level of 6.09 years (Engindeniz et al. 2010). Özkan et al. (2011) found that the average age of greenhouse tomato farmers was 42 years, the household size was 3.9 people, and 66.80% of the farmers were primary school graduates. In a study by Metin (2020), the average age of tomato farmers was found to be 42.27 years, with 38.13% having completed primary education and 20% having non-agricultural income. Another study conducted in the Kumluca district of Antalya, Türkiye, found that the average age of tomato farmers was 47.55 years, with an average education level of 6.44 years (Demircan et al. 2019). Sanga & Elia (2020) found that 96% of tomato farmers in Tanzania were aged 45 or younger, and 45% had completed secondary education. In another study, the average age of farmers was found to be 45 years, with 28% having completed secondary school or higher education (Ceylan et al. 2020). In another study, it was determined that 56.1% of farmers were aged between 41 and 60 years, 86.4% had a secondary education level or below, and 59.6% had non-agricultural income (Varoğlu 2022).

The agricultural production experience of the farmers was found to be 27.01 years, while their greenhouse production experience was found to be 25.1 years (Table 1). In another study, the average agricultural production experience of greenhouse tomato farmers was found to be 15.7 years (Özkan et al. 2011). In a study by Metin (2020), the average agricultural experience of tomato farmers was determined to be 22.49 years, while their greenhouse production experience was 19.11 years. Differences among farmers implementing biological and biotechnical control methods were identified in the study by Sayın et al. (2020). Accordingly, the average age of farmers implementing these control methods was 52.3 years, with an education level of 8.1 years and an agricultural experience of 28.3 years, whereas those not implementing these methods had an average age of 51.3 years, an education level of 8.4 years, and an agricultural experience of 25.5 years. It was found that 77.39% of the farmers were not members of any farmer organization (Table 1). In another study, it was determined that 93.77% of the farmers were registered with any farmer organization (Varoğlu 2022).

The average farm size was determined to be 1.01 ha, with an average greenhouse area of 0.71 ha (Table 1). Karaman & Yılmaz (2007) found that tomato farmers had an average farm width of 2.81 ha in their study. Çanakcı & Akıncı (2009) determined that the average greenhouse size of farms in another study involving greenhouse farmers was 0.14 ha. In another study focusing on tomato greenhouses, the average farm size was determined to be 0.46 ha (Özkan et al. 2011). In a study conducted in the Kumluca district of Antalya province, Demircan et al. (2019) found that the greenhouse width of farmers averaged 0.33 ha. In another study, the average farm width of tomato farmers was found to be 1.01 ha, with an average greenhouse width of 0.83 ha (Metin 2020). Yılmaz (2014) showed that the greenhouse size was 0.57 ha in farms where bumblebees were used, while it was determined to be 0.55 ha in farms where bumblebees were not used. In the study by Sayın et al. (2020), they revealed that farms using biological and biotechnical control methods (5.14 ha) had a higher average cultivated land area compared to farms not using them (4.19 ha).

Table 1 also includes some general information about tomato production by farmers in the Kumluca district of Antalya province, Türkiye. The average tomato production area on the surveyed farms was determined to be 0.50 ha, with an average greenhouse age of 17.07 years. The average yield in tomato production was found to be 152.5 tons per hectare. These yield values vary depending on the agricultural techniques applied by tomato farmers, the quality of seedlings used, growing conditions, and greenhouse conditions. The average price of tomato sold by farmers in Türkiye was determined to be $0.41/kg. The research recorded the lowest price at $0.21/kg and the highest price at $0.91/kg. Tomato prices fluctuate depending on seasonal factors, supply and demand balance, export pathways, production quantity, and quality. The product loss in tomato production is 9,175.9 kg/hectare. It was determined that product losses occurred in all of the examined farms due to the effects of diseases and pests, as well as the harvesting, storage and marketing processes. Özkan et al. (2011) specified the average usage period of plastic greenhouses used by tomato farmers in their studies. Accordingly, it was found that 49.13% had a usage period of 1–10 years, 45.95% had a usage period of 10–20 years, and 4.91% had a usage period of 20 years or more. In another study conducted in the Kumluca district of Antalya province, the average tomato yield was determined to be 158 tons/hectare (Demircan et al. 2019). In another study, the average tomato yield was found to be 108.8 tons/hectare on farms using bumblebees and 94.30 tons/hectare on farms not using bumblebees (Yılmaz 2014).

In the study, it was determined that 49.57% of the farmers implement biotechnical control. However, it was found that not all farmers adopt biological control methods (Table 1). These significant findings indicate that many farmers in the greenhouse tomato production prefer biotechnical control, yet not all farmers have adopted biological control methods. Studies conducted by Mason and Huber (2002), Topuz (2005), and Pérez-García et al. (2011) emphasize the importance of environmentally friendly biopesticides and biotechnical/biological control methods, highlighting the necessity for further research and implementation in this regard.

Table 1- Socio-demographic and technical characteristics of farmers and farms

Socio-Demographic Characteristics	$\bar{x}$ **/ %**
Age (years) (std. deviation: 10.969)	49.6
Education level (years) (std. deviation: 8.0)	8.0
Percentage of those with social security (%)	87.83
Percentage of those with non-agricultural income (%)	54.78
Years of agricultural production experience (years) (std. deviation: 13.243)	27.01
Years of greenhouse production experience (years) (std. deviation: 12.061)	25.10
Membership in farmer organizations (%)	22.61
Technical Characteristics	$\bar{x}$ **/ %**
Total farm area (ha) (std. deviation: 0.831)	1.01
Total greenhouse area (ha) (std. deviation: 0.669)	0.71
Age of greenhouse (years) (std. deviation: 9.531)	17.07
Tomato production area (ha) (std. deviation: 0.447)	0.50
Tomato yield (kg/ha) (std. deviation: 3.290)	152 456
Tomato selling price ($/kg) (std. deviation: 0.228)	0.41
Product loss (kg/ha) (std. deviation: 492.180)	9 175.90
The rate of farms experiencing product losses (%)	100.00
Percentage of those implementing biotechnical control (%)	49.57

Table 2 presents the purpose of tomato production and the countries to which tomato are exported from the Antalya province. According to the table, it was determined that 14.78% of the farmers produce solely for domestic consumption, 33.91% produce solely for export, and 51.31% produce for both domestic consumption and export purposes. When examining the countries to which tomato are exported from the Antalya province, it was found that the highest exports are to Russia (73.04%), Ukraine (39.13%), Romania (25.22%), and Bulgaria (16.52%). In 2022, among the top 10 countries to which Türkiye exported tomato, Syria (18.58%), Ukraine (13.23%), Romania (10.78%), Bulgaria (9.15%), Israel (8.90%), Russia (6.97%), Germany (4.52%), Poland (4.26%), Georgia (3.76%), and the Netherlands (2.95%) were identified (TRADE MAP 2022).

Table 2- Purpose of tomato production and exported countries

	N	***%***
Purpose of Tomato Production		
Domestic consumption	17	14.78
Export	39	33.91
Both	59	51.31
Exported Countries[1]		
Russia	84	73.04
Ukraine	45	39.13
Romania	29	25.22
Bulgaria	19	16.52
Czech Republic	3	2.61
Moldova	3	2.61
Germany	1	0.87
Georgia	1	0.87
Poland	1	0.87
Serbia	1	0.87

[1]: Multiple options have been selected.

Table 3 presents the most common diseases and pests encountered by farmers in greenhouse tomato production. It was determined that the most encountered pest in greenhouse tomato production, referred to as *Tuta absoluta*, accounted for 99.13%, while the most prevalent disease was *Leveillula tauric* (Powdery mildew in Solanaceae), accounting for 79.13%. Following *Tuta absoluta*, the next most encountered pests were *Tetranychus urticae* (Red spider mite) at 94.78%, *Bemisia tabaci* (Whitefly) at 89.57%, and *Helicoverpa armigera* (African bollworm) at 86.96%. After *Leveillula tauric*, the most common diseases encountered were *Phytophthora infestans* (Late blight) at 71.30% and *Botrytis cinerea* (Gray mold) at 59.13%.

Table 3- Diseases and pests encountered by farmers in greenhouse tomato production[1]

Pests	***Number of Farmers (N)***	***%***
Tuta absoluta (Tomato leafminer)	114	99.13
Tetranychus urticae (Red spider mite)	109	94.78
Bemisia tabaci (Whitefly)	103	89.57
Helicoverpa armigera (Corn earworm)	100	86.96
Thysanoptera (Flower thrips)	12	10.43
Myzus persicae (Green peach aphid)	6	5.22
Diseases	***Number of Farmers (N)***	***%***
Leveillula tauric (Powdery mildew in Solanaceae)	91	79.13
Phytophthora infestans (Late blight)	82	71.30
Botrytis cinerea (Gray mold)	68	59.13
Fusarium oxysporum (Fusarium with in tomato)	23	20.00
Meloidogyne spp. (Root-knot nematodes)	18	15.65
Aculops lycopersici (Tomato russet mite)	7	6.09
Rhizoctonia solani (Root rot)	3	2.61
Alternaria solani (Early blight)	2	1.74

[1]: Multiple options were selected

Figure 1- *Phytophthora infestans* (Mildew)

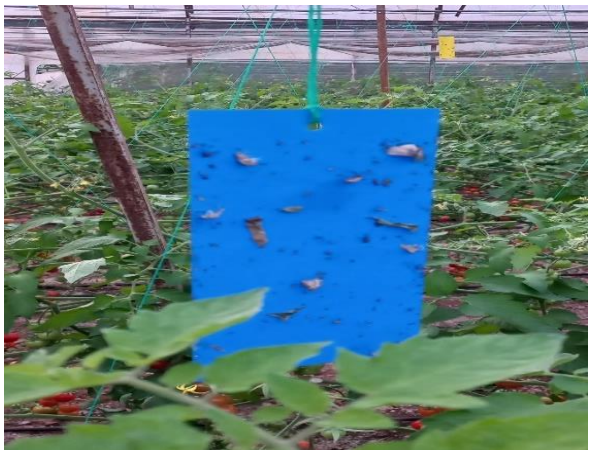


Figure 2- Biological control (Blue sticky trap)

Table 4 presents the relationship between specific diseases and pests encountered by tomato farmers and the amount of product loss in tomato production. The differences in the means of product loss due to the occurrence of diseases or pests in tomato were determined using an independent t-test. Accordingly, the average product loss was found to be 9,256 kg/ha in tomato affected by *Tuta absoluta* (Tomato leafminer), 9,345 kg/ha in tomato affected by *Tetranychus urticae* (Red spider mite), and 9,850 kg/ha in tomato affected by *Phytophthora infestans* (Late blight). The independent t-test results indicated that these differences were statistically significant. These findings demonstrate that the production quantity of tomato is affected by specific diseases or pests, leading to losses in tomato production.

Table 4- Product loss of tomato farmers according to diseases and pests encountered in the greenhouse

The amount of tomato product loss (kg/ha)		***N***	***Mean ±Std. Deviation***	***t-test***
Tuta absoluta (Tomato leafminer)	Present	114	9 256±486.692	-2.125*
	Absent	1	0.00±0	
Tetranychus urticae (Red spider mite)	Present	109	9 345±494.193	-2.125*
	Absent	6	6 091±356.858	
Phytophthora infestans (Late blight)	Present	82	9 850±488.454	-2.361**
	Absent	33	7 500±467.413	

*: 0.10 and **: 0.05 are significant at the level.

In greenhouse tomato production, the use of pesticides and their quantities have a significant impact on yield. Table 5 shows the amounts of insecticides, fungicides, nematicides, and acaricides used by farmers in tomato production.

In the examined farms, it was found that an average of 12.9 kg/ha of insecticide (35.5%), 10.9 kg/ha of nematicide (29.9%), 8.8 kg/ha of fungicide (24.1%), and 3.8 kg/ha of acaricide (10.4%) were used, total 36.5 kg/ha of pesticide. In a study conducted on tomato production in greenhouses in the Serik district of Antalya province, Türkiye, it was reported that farms certified with EurepGAP used a total of 12.2 kg/ha of pesticides, including 9.7 kg/ha of fungicide, 0.5 kg/ha of insecticide, and 2 kg/ha of nematicide. Conversely, farms without EurepGAP certification were found to use a total of 43.7 kg/ha of pesticides, including 34.1 kg/ha of fungicide, 2.3 kg/ha of insecticide, and 9.3 kg/ha of nematicide (Bayramoğlu et al. 2010). In another study conducted in Antalya province, it was reported that for winter tomato production in greenhouses, 14.8 kg/ha of pesticides were used, including 1.5 kg/ha of insecticide, 11.8 kg/ha of fungicide, and 1.3 kg/ha of nematicide. In the same study, for summer tomato production, 8.2 kg/ha of pesticides were used, including 1.6 kg/ha of insecticide, 5.9 kg/ha of fungicide, and 0.7 kg/ha of nematicide. Additionally, for spring tomato production, 4.3 kg/ha of pesticides were used, including 0.6 kg/ha of insecticide and 3.7 kg/ha of fungicide (Özkan et al. 2008). In a study conducted with tomato producers in Kenya, it was found that producers used insecticides (97.66%), fungicides (91.93%), and herbicides (16.67%) in excessive doses (Kinuthia, 2019). Comparing with other studies, it can be said that pesticide use in greenhouse tomato production varies according to cultivation techniques, cultivation periods, density of pests and diseases, whether farms have EurepGAP certification, and whether use biotechnical and biological control techniques. EurepGAP, represents an international standard used to certify that agricultural products are produced in a safe and sustainable manner (Kurek 2007). European Union countries and other advanced nations may require compliance with this standard for agricultural products. Having this certification can enhance the acceptability and reliability of products in international markets. However, this requirement may vary depending on the type of product, target markets, and buyers.

It has been revealed that farmers do not comply with the recommended pesticide doses from agricultural engineers and agricultural extension agents (Table 5). The reason for farmers not adhering to the recommended pesticide doses may be attributed to their lack of sufficient knowledge or awareness about proper usage practices. Increasing agricultural extension

activities is considered a potential solution to prevent unconscious pesticide use, as it would provide guidance to farmers on proper usage. In their study, Ombeni et al. (2021) indicated that farmers' knowledge about pesticide use in tomato, their age, the number of their children, and the nutritional values of agricultural products produced with and without pesticides are influenced by product prices. Another study highlighted that intensive pesticide use prevails in greenhouse vegetable farming in Antalya province, and most farmers tend to excessively use pesticides without proper consideration (Özkan et al. 2022).

It was determined that 59.13% of the surveyed farms have regular communication with an agricultural consultant, while 40.87% do not have an agricultural consultant. Additionally, the frequency of visits by agricultural consultants was determined in the study, with 41.18% of farmers stating that consultants visit their greenhouse once a week, 32.35% every two weeks, 19.12% once a month, and 7.35% every ten days (Table 5).

In the study area, farmers mostly make decisions about pesticide use themselves (91.30%), but they also receive support from agricultural engineers, agricultural consultants, and technical staff (74.78%) regarding its use (Table 5). In a study conducted in the Kumluca district of Antalya province, Türkiye, it was stated that farmers mainly base their decisions on agricultural pesticide use on recommendations from pesticide dealers (61.7%) (Kan, 2002). In the research by Nguetti et al. (2018), 63.5% of farmers expressed that their own experience influences their decisions on pesticide use. Another study indicated that farmers determine the pesticides they use in vegetable production based on information obtained from agricultural pesticide dealers, elders in the family, and other farmers (Malgie et al. 2015).

Table 5- Pesticide usage and pesticide usage behaviour in surveyed farms

Type of agricultural chemicals	***Used amount***	
	(gr-ml-cc)/ha	***%***
Insecticides	12 948.07	35.46
Nematicides	10 944.65	29.98
Fungicides	8 807.95	24.12
Acaricides	3 810.80	10.44
Total pesticides	36 511.47	100.00
	N	**%**
Do you have a regular agricultural consultant		
Yes	68	59.13
No	47	40.87
How often do agricultural consultant visit?		
Weekly	28	41.18
Every ten days	5	7.35
Biweekly	22	32.35
Monthly	13	19.12
Decision making on pesticide usage[1]		
Personal choice	105	91.30
Agricultural Engineer/Agricultural Consultant/Technical Staff	86	74.78
Family members	7	6.09
Other farmers	7	6.09

[1]: Multiple options were selected.

Due to the significant role of tomato exports, the agricultural chemicals used in the production process are crucial. Furthermore, examining this issue is of great importance in the context of sustainable agricultural production. In Table 6, the pesticides used by farmers are grouped according to the toxicity classification of the World Health Organization (WHO) for tomato production. It was determined that farmers on the surveyed farms used pesticides containing 57 different active ingredients.

Farmers commonly use insecticides, including Laser (60.00%), Radiant 120 SC (56.52%), Movento SC 100 (37.39%), Mospilan 20 SP (32.17%), Uphold 360 SC (20.87%), and Voliam Targo (19.13%). It has been determined that 55.88% of the pesticides used by farmers are classified as hazardous according to the World Health Organization's (WHO) classification. The most commonly used acaricides by farmers are Agrimec EC (33.91%), Actinmor (23.48%), Capella (19.13%), Bamilda (10.43%), and Ariphut (10.43%). Most of the acaricides used are classified as highly hazardous or fall into the low risk acute hazard category under normal use. Commonly used fungicides by farmers include Signum WG (34.78%), Dikotan M-45 (24.35%), Trimaton (20.00%), Nirmal (17.39%), and Collis SC (16.52%). The fungicides used generally fall into the WHO hazard classification categories of slightly hazardous, moderately hazardous, or low risk of acute hazard under normal use. It has been found that nematicides are used in low quantities by farmers, with Pagos 10 G (6.09%) and Tervigo 20 SC (4.35%) being the most commonly used. The nematicides used are generally classified as slightly hazardous or highly hazardous in toxicity.

It has been revealed that a significant amount of pesticides are used in greenhouse tomato production in the research area. It was determined that farmers not only use agricultural chemicals when diseases/pests occur but also for preventive purposes

before diseases/pests appear. In a study conducted in Egypt, the average contents of HCB, lindane, dieldrin, heptachlor epoxide, and DDT derivatives in tomatoes were found to be 0.009 mg/kg, 0.003 mg/kg, 0.006 mg/kg, 0.008 mg/kg, and 0.083 mg/kg, respectively. Additionally, the levels of dimethoate, profenofos, and pirimiphos-methyl were determined to be 0.461 mg/kg, 0.206 mg/kg, and 0.114 mg/kg, respectively (Abou-Arab 1999). In the study by Nguetti et al. (2018), it was noted that 6% of farmers conducted post-harvest spraying to protect their crops. Kan's (2002) study revealed that 49.5% of farmers considered pre-emptive spraying necessary before the appearance of pests. Additionally, it was indicated that 63% of farmers who sprayed more than 22 times during a production season operated in plastic greenhouses. Engindeniz et al. (2010) found that farmers used 36% more pesticides for tomato grown in glass greenhouses. The reason for this was attributed to 43.3% of farmers believing that the recommended dosage was not effective enough and 36.7% stating that diseases and pests had developed resistance. In the study conducted by Hepsağ & Kızıldeniz (2021), it was found that in tomatoes grown in the Mediterranean region of Türkiye, 61.5% of the samples had detectable residues of one or more pesticides. The most commonly found residues included methyl chlorpyrifos, cyfluthrin, deltamethrin, and acetamiprid.

Table 6- The types of pesticides used in tomato production according to the WHO Hazard Classification

	Trade name	*Chemical family*	*Toxicity class[1]*	*Status[2]*	*Number of farmers (N)****	*%*
Insecticide	Laser	480 g/l Spinosad	III	Registered	69	60.00
	Radiant 120 SC	120 g/l Spinetoram	U	Registered	65	56.52
	Movento SC 100	100 g/l Spirotetramat	III	Registered	43	37.39
	Mospilan 20 SP	%20 Acetamiprid	II	Registered	37	32.17
	Uphold 360 SC	300 g/l Methozyfenozide + 60 g/l Spinetoram	U	Registered	24	20.87
	Voliam Targo	45 g/l Chlorantraniliprole + 18 g/l Abamectin	U + Ib	Registered	22	19.13
	Hektaş Zodiac	100 g/l Spirotetramat	III	Registered	5	4.35
	Imperator 25 EC	250 g/l Cypermethrin	II	Registered	4	3.48
	Breaker 240 SC	240 g/l Sulfoxaflor	II	Registered	9	7.83
	No-fly	%20 Acetamiprid	II	Registered	10	8.70
	Liberator 10 EC	100 g/l Pyriproxyfen	U	Registered	11	9.57
	Atercle	240 g/l Mataflumizone	U	Unregistered	13	11.30
	Dacron WP	32000 IU/mg Bacillus thuringiensis var kurstaki strain SA-11	III	Registered	9	7.83
	Zoomilda	Malathion	III	Unregistered	4	3.48
	Neemarin	%5 Emamectin benzoate	II	Registered	8	6.96
	Bethrin 2.5 EC	25 g/l Deltamethrin	II	Registered	12	10.43
	Circaden 200 SC	200 g/l Cyantraniliprole	U	Registered	15	13.04
	Evander	%5 Emamectin benzoate	II	Registered	6	5.22
	Modesta	350 g/l Imidacloprid	II	Unregistered	9	7.83
	Tunga	222 g/l Flubendiamide	III	Registered	17	14.78
	Spintor 240 SC	240 g/l Spinosad	III	Registered	13	11.30
	Benevia 100 OD	100 g/l Cyantraniliprole	U	Registered	6	5.22
	Plemax	240 g/l Indoxacarb + 80 g/l Novaluron	II + U	Registered	7	6.09
	Altacor 35 WG	%35 Chlorantraniliprole	U	Registered	11	9.57
	Klarbon M22	227 g/l Chlorpyrifos methyl	III	Registered	6	5.22
	Dicarzol 50 SP	%50 Formetanate hydrochloride	Ib	Registered	8	6.96
Acaricide	Agrimec EC	18 g/l Abamectin	Ib	Registered	39	33.91
	Actinmor	18 g/l Abamectin	Ib	Registered	27	23.48
	Capella	240 g/l Bifenazate	U	Registered	22	19.13
	Bamilda	Bifenazate	U	Unregistered	12	10.43
	Ariphut	Paraffin oil	Not listed	Unregistered	12	10.43
	Hebanon	Hesoythiatepo	Not listed	Unregistered	7	6.09
	Puzzle 20 WP	%20 Pyridaben	II	Registered	7	6.09
	Efdal Hekzarun 50 EC	50 g/l Hexythiazox	U	Registered	6	5.22
	Zrmilda	Abamectin-propois	Not listed	Unregistered	6	5.22
	Tripsol	22.5 g/l Acrinathrin + 12.6 g/l Abamectin	U + Ib	Registered	4	3.48

Table 6 (Continue)- The types of pesticides used in tomato production according to the WHO Hazard Classification

	Trade name	*Chemical family*	*Toxicity class*[1]	*Status*[2]	*Number of farmers (N)****	*%*
	Signum WG	%26.7 Boscalid + %6.7 Pyraclostrobin	U	Registered	40	34.78
	Dikotan M-45	%80 Mancozeb[4]	U	Unregistered	28	24.35
	Trimaton	500 g/l Metam-sodium	II	Registered	23	20.00
	Nirmal	100 g/l Penconazole	III	Registered	20	17.39
	Collis SC	200 g/l Boscalid + 100 g/l Kresoxim-methyl	U + III	Registered	19	16.52
	Lusen SC 500	250 g/l Fluopyram + 250 g/l Trifloxystrobin	III + U	Registered	14	12.17
	Hektaş Aktor	500 g/l Imazalil	II	Registered	9	7.83
	Qualy 300 EC	300 g/l Cyprodinil	Not listed	Registered	10	8.70
	Ippon 500 SC	Iprodione	III	Registered	6	5.22
	Yoma	375 g/l Fluazinam + 150 g/l Azoxystrobin	U	Registered	8	6.96
	Switch 62.5 WG	%37.5 Cyprodinil + %25 Fludioxonil	U	Registered	4	3.48
	Orvego	300 g/l Ametoctradin + 225 g/l Dimethomorph	III	Registered	6	5.22
	Ridomil Gold MZ 68 WG	500 g/l Chloratholonil + 37.5 g Metalaxyl-M	II	Registered	8	6.96
	Mastercop	65.82 g/l Copper oxychloride Copper sulfate	II	Registered	3	2.61
	Adolax	%80 Mancozeb[4]	U	Unregistered	5	4.35
Fungicide	Amaline	266.6 g/l Metallic copper equivalent basic copper sulfate + 40 g/l Zoxamide	II + U	Registered	5	4.35
	Bonko	500 g/l Chlorothalonil	U	Unregistered	9	7.83
	Korconil 500 SC	500 g/l Chlorothalonil	U	Unregistered	10	8.70
	Luna Tranquility SC 500	375 g/l Pyrimethanil + 152 g/l Fluopyram	III	Registered	10	8.70
	Equation Pro	%30 Cymoxanil + 22.5 Famoxadone	II + U	Registered	1	0.87
	Albacore	65.82 g/l Copper oxychloride Copper sulfate	II	Registered	1	0.87
	Quadris Maxx	200 g/l Azoxystrobin + 125 g/l Difenoconazole	U + II	Registered	4	3.48
	Topguard EQ	296 g/l Azoxystrobin + 218 g/l Flutriafol	U + II	Registered	6	5.22
	Embrelia 140 SC	100 g/l Isopyrazam + 40 g/l Difenoconazole	II	Registered	11	9.57
	Origam FS	75 g/l Azoxystrobin + 37.5 Metalaxyl-m + 12.5 g/l Fludioxonil	U + II + U	Registered	1	0.87
	Tennis 360 SL	360 g/l Hymexazol	III	Registered	2	1.74
	Top-Copp 5 E	51.4 g/l Copper salts of fatty and rosin acids equivalent to metallic copper	II	Registered	2	1.74
	Pagos 10 G	%10 Fosthiazate	Not listed	Registered	7	6.09
	Tervigo 20 SC	20 g/l Abamectin	Ib	Registered	5	4.35
	Velum Prime SC 400	400 g/l Fluopyram	III	Registered	3	2.61
Nematicide	Bioact DC 216	$4.7*10^{10}$ * 216 adet spor/l Purpureocillium lilacinus 251 (Paecilomyces lilacinus strain 251)	Not listed	Registered	2	1.74
	Nemacur 400 EC	400 g/l Fenamiphos	Ib	Registered	2	1.74

[1]Toxicity class (Ia = Extremely hazardous; Ib = Highly hazardous; II = Moderately hazardous; III = Slightly hazardous; U = Unlikely to cause acute hazard under normal use; FM = Fumigant, unclassified; O = Former pesticide, unclassified. WHO (World Health Organization), 2019. (Accessed: 09.06.2023); [2]Anonymous, 2023. (Erişim Tarihi: 09.06.2023); [3]Farmers have declared the use of multiple plant protection products for diseases/pests; [4]The use and sale of fungicides that active ingredient is Mancozeb is prohibited by the Ministry of Agriculture and Forestry. However, farmers who have fungicides containing the active ingredient Mancozeb are allowed to use them until 31.12.2023. During the survey process of this research, farmers were encountered who used fungicides with the active ingredient Mancozeb within the allowed period. As a matter of fact, it was determined that 24.35% of the producers used Dikotan M-45 with the active ingredient mancozeb and 4.35% used Adolax; [5]The registration status of pesticides is based on the information provided by the Ministry of Agriculture and Forestry at https://bku.tarimorman.gov.tr/.

The RASFF (Rapid Alert System for Food and Feed) reports on non-compliance issues detected in exports from RASSF Window countries to European Union countries. These RASFF notifications, classified as information exchange or border rejection, form an online database. To maintain a balance between transparency and the protection of commercial information, individual company names and identities are not disclosed. In exceptional cases where the protection of human health demands greater transparency, the Commission takes appropriate measures through regular communication channels. The Commission also informs third-country authorities about notifications concerning products produced, distributed, or sent from these countries. Table 7 provides some information regarding pesticide residue notifications related to tomato exports from European Union (EU) countries to Türkiye in the year 2023. These notifications reflect the seriousness of various countries regarding chemical substances in agricultural products and the measures taken. In 2023, a total of seven notifications were made to Turkey from EU countries concerning tomato products. All of these notifications are related to pesticide residues. Three of these notifications are from Romania, while the others are from Germany, Croatia, Italy, and Slovenia. Out of these residue notifications, 5 are considered to be notifications requiring attention, while 2 are in the form of border rejection notifications. Border rejection is a measure taken against food and feed shipments tested at the external borders of the European Union and found to pose health risks. During this process, notifications are sent to all EU border points to increase the effectiveness of controls and prevent rejected products from re-entering the EU through another border entry point. This helps to ensure that food safety and health standards are effectively maintained at the EU's external borders.

As a result of pesticide residue notifications to Türkiye, education should be provided to farmers and other stakeholders in the agricultural sector regarding the use of pesticides, and transition to biological/biotechnical control methods should be encouraged. Additionally, export controls to the EU from Türkiye should be strengthened and rigorously monitored. Türkiye should prioritize improving its customs and food inspection processes and ensuring compliance with EU standards.

Table 7- Pesticide residue notifications from the EU to Türkiye in tomato exports in 2023

Subject	*Date*	*Classification*	*Risk Decision*	*Notifying Country*	*Hazards*
Unauthorized substance chlorpyrifos-methyl and exceeding pesticide pirimiphos methyl	23.06.2023	Information notification for attention	Potentially serious	Romania	Chlorpyrifos-methyl unauthorised substance,pirimiphos-methyl
Unauthorized substance (chlorpyrifos-methyl, 0,048 mg/kg - ppm)	08.05.2023	Information notification for attention	Potentially serious	Romania	Chlorpyrifos-methyl
Exceeding LMA pesticid chlorpirifos	11.04.2023	Information notification for attention	Serious	Romania	Chlorpyrifos-methyl unauthorised substance
Chlorpirifos-methyl in frozen oven semi-dried tomato	13.02.2023	Border rejection notification	Potentially serious	Italy	Chlorpyrifos-methyl unauthorised substance
Chlorpyrifos	26.05.2023	Information notification for attention	Potentially serious	Slovenia	Chlorpyriphos-ethyl
Unauthorized substance (Chlorothalonil)	29.03.2023	Information notification for attention	Serious	Germany	Chlorothalonil
Chlorothalonil	23.03.2023	Border rejection notification	No risk	Croatia	Chlorothalonil

Source: RASFF Window 2023

4. Conclusions

Adopting sustainable methods for combating plant diseases and pests, aiming to preserve the sustainability of agriculture and ecosystem health, offers a more environmentally friendly approach compared to traditional agricultural practices. Within this context, environmentally friendly methods such as promoting natural enemies, preserving biological diversity, and prioritizing soil health are crucial for controlling plant diseases and pests.

While pesticides commonly used in traditional farming practices may be effective in controlling plant diseases and pests, their overuse can lead to negative effects such as the development of resistance in harmful organisms, soil, water, and air pollution, and adverse impacts on human health. Therefore, farmer awareness campaigns should be conducted regarding the use of sustainable approaches, and the adoption of these practices should be encouraged.

The socio-demographic characteristics of greenhouse tomato farmers in the Kumluca district of Antalya province, Türkiye, as well as the technical features of their farms, were examined in the study. It was found that the education levels of the farmers were relatively low, and the size of agricultural farms consisted mostly of small-scale farms. Nearly half of the farmers were found to lack regular agricultural extension activities. Additionally, it was determined that a high proportion of farmers (91.30%) made pesticide usage decisions based on their own knowledge and experience.

It has been determined that the most common pests encountered by tomato farmers in greenhouse tomato production are *Tuta absoluta*, *Tetranychus urticae*, *Bemisia tabaci*, and *Helicoverpa armigera*, while the most common diseases encountered are *Leveillula tauric*, *Phytophthora infestans*, and *Botrytis cinerea*. Accordingly, it has been observed that the presence of *Tuta absoluta* and *Tetranychus urticae* pests, along with *Phytophthora infestans* disease, contributes to increased product losses in tomato production.

It has been determined that the average pesticide usage in the examined farms is 36.5 g/ha. Within the relevant production period, it was found that pesticides in the insecticide (35.46%), fungicide (24.12%), and acaricide (10.44%) groups were heavily utilized. It was also observed that farmers generally do not adhere to the recommended pesticide usage doses provided by agricultural engineers and extension agents. Increasing agricultural extension activities and providing guidance to farmers on the proper use of pesticides would be a crucial step in reducing indiscriminate pesticide usage.

Tomato is among the most significant export products in Türkiye. The study results indicate that nearly all farmers focus on tomato production for export purposes. The countries most commonly exported to are typically Russia, Ukraine, Romania, and Bulgaria. The use of pesticides in exported products is of significant importance. It is expected that these products have no residues and comply with the standards of the importing country. In this context, the diseases and pests encountered by farmers have been identified, and the pesticides used in their control have been determined. It was found that farmers on the examined farms used 57 different active substances in pesticides. Most of the insecticides used by farmers fall into the category of slightly hazardous, moderately hazardous, or posing a low acute risk under WHO hazard classification; acaricides mostly fall into the category of highly hazardous or posing a low acute risk under normal use; fungicides generally fall into the category of slightly hazardous, moderately hazardous, or posing a low acute risk under WHO hazard classification; and nematicides generally fall into the category of slightly hazardous or highly hazardous toxicity. Excessive pesticide usage has been detected in tomato grown in greenhouses in the research area. Farmers not only use pesticides when diseases or pests emerge but also as a preventive measure before the onset of diseases or pests. This practice not only has negative environmental impacts but also leads to immunity and resistance against pesticides. To promote the adoption of environmentally friendly practices, it would be beneficial to organize regular training programs through agricultural engineers and experts to raise awareness among farmers about the correct use of pesticides and the identification of diseases and pests specific to tomato cultivation. To minimize these adverse effects and preserve environmental health, farmers should be incentivized to adopt these methods through agricultural support mechanisms. Steps taken in this direction could help farmers adopt a more sustainable and environmentally friendly approach to agricultural practices.

As a result, the risks posed by pesticide usage to the sustainability of Türkiye's tomato exports and production should not be overlooked. It was determined in the study that 55.88% of the pesticides used by farmers are classified as hazardous, according to the World Health Organization (WHO). The increasing number of pesticide residue notifications from European Union countries in 2023 serves as a warning for Turkish farmers. Therefore, more efforts need to be made towards compliance with international standards to ensure the continuity of Türkiye's tomato exports. In this context, inspections need to be tightened, more selective pesticide use should be encouraged, quality control mechanisms should be strengthened, and supports provided to farmers should be redesigned to encourage sustainable agriculture. It is important to increase inspections in greenhouses where tomato production is carried out and to develop training programs for farmers to ensure the production of products that are safe for the environment and human health. Taking these measures will result in Turkey's entry into export markets, strengthening Turkey's image in export markets and increasing export revenues.

In order for the recommendations to be realized and the goals of safe food in tomato production to be achieved, it is critical to ensure the continuity of the inspection, monitoring and support mechanism and to implement it decisively and to clearly demonstrate of political and administrative willpower.

Acknowledgements

This study is derived from the Master's Thesis named "Investigation of the Effects Agricultural Support Policies on Agroecological Applications in Tomatoes Production in Greenhouse: A Case Study from Kumluca District of Antalya Province (In Turkish)" in the department of Agricultural Economics, Institute of Graduate Education, Isparta University of Applied

Sciences.The thesis work was supported by the PRIMA (Partnership for Research and Innovation in the Mediterranean Area) program under TÜBİTAK project number 122N053. I am very grateful to the Scientific and Technological Research Council of Turkey (TÜBİTAK) for their support.

References

Abou-Arab A A K (1999). Behavior of pesticides in tomatoes during commercial and home preparation. *Food chemistry* 65(4):509-514. https://doi.org/10.1016/S0308-8146(98)00231-3

Alguacil J, Kauppinen T, Porta M, Partanen T, Malats N, Kogevinas M, Benavides F G, Obiols J, Bernal F, Rifa J & Carrato A (2000). Risk of pancreatic cancer and occupational exposures in Spain. *Annals of Occupational Hygiene* 44(5): 391-403. https://doi.org/10.1093/annhyg/44.5.391

Anonymous (2023). *Plant Protection Products Department of the Ministry of Agriculture and Forestry.* Retrieved June 9, 2023, https://bku.tarimorman.gov.tr/

Anonymous (2024). *Supports Provided for Greenhouse Production*. Retrieved June 18, 2024, https://www.tarimorman.gov.tr/Konular/Bitkisel-Uretim/Tarla-Ve-Bahce-Bitkileri/Ortu-Alti-Yetistiricilik

Bassil K L, Vakil C, Sanborn M, Cole D C, Kaur J S, & Kerr K J (2007). Cancer health effects of pesticides: Systematic review. *Canadian Family Physician* 53(10): 1704-1711

Bayramoğlu Z, Gündoğmuş E, & Tatlıdil F F (2010). The impact of EurepGAP requirements on farm income from greenhouse tomatoes in Turkey. *African Journal of Agricultural Research* 5(5): 348-355

Çanakcı M & Akıncı İ (2009). Agricultural practices and human labor utilization in greenhouse vegetable cultivation in Antalya province. *Journal of Agricultural Machinery Science* 5(2): 193-202 (In Turkish)

Ceylan R F, Özkan B, Sayın C & Akpınar M G (2020). Factors effecting profit efficiency of greenhouse producers in Antalya. *International Conference Emerging New Word (ICENWE).* July 26-28, India, 117-128

Demircan V, Turan A & Dalgıç A (2019). Labor use in greenhouse tomato production: a case study from Kumluca district of Antalya province, Turkey. *Lucrări Ştiinţifice* 62(2): 9-12

Dolapsakis G, Vlachonikolis I G, Varveris C & Tsatsakis A M (2001). Mammographic findings and occupational exposure to pesticides currently in use on crete. *European Journal of Cancer*, 37(12): 1531-1536. https://doi.org/10.1016/S0959-8049(01)00159-9

Engindeniz S, Yılmaz İ, Durmuşoğlu E, Yağmur B, Eltez R Z, Demirtaş B, Engindeniz D & Tatarhan A H (2010). Comparative input analysis of greenhouse vegetables. *Ecology*, 19(74): 122-130. https://doi.org/10.5053/ekoloji.2010.7415

European Commission (2024). *The Common Agricultural Policy at A Glance*. Retrieved June 18, 2024, https://agriculture.ec.europa.eu/common-agricultural-policy/cap-overview/cap-glance_en

Eyhorn F, Roner T & Specking H (2015). Reducing Pesticide Use and Risks-What Action is Needed. *Swiss Intercooperation*. September 1-30, Helvetas 1-32

FAO (2023). *Tomatoes.* Retrieved June 20, 2023, https://www.fao.org/land-water/databases-and-software/crop-information/tomato/en/

FAOSTAT (2020). *Pesticides Indicators.* Retrieved May 14, 2023, https://www.fao.org/faostat/en/#data/EP

Geiger F, Bengtsson J, Berendse F, Weisser W W, Emmerson M, Morales M B, Ceryngier P, Liira J, Tscharntke T, Winqvist C, Eggers S, Bommarco R, Part T, Retagnolle V, Plantegenest M, Clement L W, Dennis C, Palmer C, Onate J J, Guerrero I, Hawro V, Aavik T, Thies C, Flohre A, Hanke S, Fischer C, Goedhart P W, & Inchausti P. (2010). Persistent negative effects of pesticides on biodiversity and biological control potential on European farmland. *Basic and Applied Ecology*, 11(2): 97-105. https://doi.org/10.1016/j.baae.2009.12.001

Hepsağ F & Kizildeniz T (2021). Pesticide residues and health risk appraisal of tomato cultivated in greenhouse from the Mediterranean region of Turkey. *Environmental Science and Pollution Research* 28(18): 22551-22562

Infante-Rivard C, Labuda D, Krajinovic M & Sinnett D (1999). Risk of childhood leukemia associated with exposure to pesticides and with gene polymorphisms. *Epidemiology* 10(5): 481-487. https://www.jstor.org/stable/3703333

Kan M (2002). *Issues and Solution Suggestions in the Use of Agricultural Pesticides in Greenhouse Vegetable Cultivation in Kumluca District of Antalya Province.* (Published Master's Thesis, Ankara University, Institute of Natural Sciences) (In Turkish)

Karaman S & Yılmaz İ (2007). Analysis of factors determining the use of Bombus bees for pollination in greenhouse tomato production. *Tekirdağ Journal of Agricultural Faculty* 4(1): 99-107 (In Turkish)

Kinuthia C W (2019). *Determinants of pesticide use and uptake of alternative pest control methods among small scale tomato farmers in Nakuru County, Kenya* (Doctoral dissertation, Egerton University).

Kumar N, Pathera A K, Saini P & Kumar M (2012). Harmful effects of pesticides on human health. *Annals of Agri-Bio Research* 17(2): 125-127

Kurek A (2007). Eurepgap–The principles of certification and their implementation in horticultural holdings. *Acta Scientiarum Polonorum. Oeconomia*, 6(3):85-94.

Maham S G, Rahimi A, Subramanian S & Smith D L (2020). The environmental impacts of organic greenhouse tomato production based on the nitrogen-fixing plant (Azolla). *Journal of Cleaner Production*, 245: 118679. https://doi.org/10.1016/j.jclepro.2019.118679

Mahmood I, Imadi S R, Shazadi K, Gul A & Hakeem K R (2016). Effects of pesticides on environment. *Plant, soil and microbes*, 1: 253-269. https://doi.org/10.1007/978-3-319-27455-3_13

Malgie W, Ori L & Ori H (2015). A study of pesticide usage and pesticide safety awareness among farmers in Commewijne in Suriname. *Journal of Agricultural Technology*, 11(3): 621-636.

Mason P G & Huber J T (2002). *Biological Control Programmes in Canada, 1981–2000*. CABI Publishing.

Metin F D (2020). *Determination of Purchasing Power of Producers Engaged in Greenhouse Tomato Production in Antalya Province.* (Published Master's Thesis, Akdeniz University, Institute of Natural Sciences) (In Turkish)

Nguetti J H, Imungi J K, Okoth M W, Wang'ombe J, Mbacham W F & Mitema S E (2018). Assessment of the knowledge and use of pesticides by the tomato farmers in Mwea Region, Kenya. *African Journal of Agricultural Research*, 13(8): 379-388.

OECD (2022). *Education at a Glance.* Retrieved May 22, 2023, https://doi.org/10.1787/3197152b-en

Ombeni J, Munyuli T & Makelele J (2021). Use of pesticides in tomato crop and its impact on the nutritional quality of vegetables and the health status of market gardeners in rural health zones of South Kivu Province, Eastern of the Democratic Republic of Congo. *Academia Letters*, 2.

Özkan B, Akçaöz H V, Karaman S & Taşcıoglu Y (2002). An economic assessment of pesticide use in greenhouse vegetable production in Antalya Province.

Özkan B, Hatırlı S A, Öztürk E & Aktaş A R (2008). Profitability Analysis of Greenhouse Tomato Production in Antalya Province. *The Scientific And Technical Research Council Of Turkey.* PROJECT NO: TOVAG-106 O 026. (In Turkish)

Özkan B, Hatırlı S, Öztürk E & Aktaş A (2011). Profitability analysis of greenhouse tomato production in Antalya province. *Journal of Agricultural Sciences,* 17(1): 34-42. https://doi.org/10.1501/Tarimbil_0000001154

Pérez-García A, Romero D & De Vicente A (2011). Plant protection and growth stimulation by microorganisms: biotechnological applications of bacilli in agriculture. *Current opinion in biotechnology,* 22(2): 187-193. https://doi.org/10.1016/j.copbio.2010.12.003

Pishgar-Komleh S H, Akram A, Keyhani A, Raei M, Elshout P M F, Huijbregts M A J & Van Zelm R (2017). Variability in the carbon footprint of open-field tomato production in Iran-A case study of Alborz and East-Azerbaijan provinces. *Journal of cleaner production*, 142: 1510-1517. https://doi.org/10.1016/j.jclepro.2016.11.154

RASFF Window (2023). *European Commisssion.* Retrieved January 2, 2024, https://webgate.ec.europa.eu/rasff-window/screen/list

Republic of Türkiye Ministry of Trade (2021). *The Green Deal Action Plan.* Retrieved June 5, 2023, https://ticaret.gov.tr/data/60f1200013b876eb28421b23/MUTABAKAT%20YE%C5%9E%C4%B0L.pdf

Sanga E E & Elia E F (2020). Socio-demographic determinants of access to climate change information among tomato growing farmers in Mvomero District, Tanzania. *University of Dar es Salaam Library Journal* 15(2): 121-136

Sayın B, Bayav A, Beşen T, Karamürsel D, Çelikyurt M A, Emre M, Kuzgun M, Yılmaz Ş G & Arslan S (2020). Determination of producers' perspectives on biological and biotechnical pest control practices and their environmental awareness. *Kahramanmaraş Sütçü İmam University Journal of Agriculture and Nature* 23(2): 453-466. https://doi.org.tr/10.18016/ksutarimdoga.vi.599085 (In Turkish)

Sezgin V (2022). How will Ukraine-Russia war affect Turkish trade? *Journal of Abant Social Sciences* 22(2): 546-557

Sharma A, Kumar V, Shahzad B, Tanveer M, Sidhu G P S, Handa N, Kohli S K, Yadav P, Bali A S, Parihar R D, Dar O I, Singh K, Jasrotia S, Bakshi P, Ramakrishnan M, Kumar S, Bhardwaj R & Thukral A K (2019). Worldwide pesticide usage and its impacts on ecosystem. *SN Applied Sciences* 1: 1-16

Szpilko D & Ejdys J (2022). European Green Deal–research directions. a systematic literature review. *Ekonomia i Środowisko*. 2.

Topuz E (2005). Some alternative methods to chemical pesticides in agricultural pest control. *Derim* 22(2): 53-59 (In Turkish)

TRADE MAP (2022). *Foreign Trade.* Retrieved April 28, 2023, https://www.trademap.org/

TÜİK (2021). *Tomato Production Quantity.* Retrieved March 24, 2022, https://biruni.tuik.gov.tr/medas/?locale=tr

UNDP (2022). *Human Development Report.* Retrieved May 22, 2023, https://hdr.undp.org/system/files/documents/global-report-document/hdr2021-22pdf_1.pdf

Van Driesche R & Bellows Jr T S (2012). *Biological Control*. Springer Science & Business Media. https://doi.org/10.1007/978-1-4613-1157-7

Varoğlu S T (2022). *Agricultural Producers' Approaches to Sustainable Agriculture in Iğdır Province*. (Published Doctoral Dissertation, Uludağ University, Institute of Natural Sciences) (In Turkish)

Wang A, Hussain S & Yan J (2024). Evaluating the interlinkage between pesticide residue regulation and agricultural green total factor productivity: empirical insights derived from the threshold effect model. *Environment Development and Sustain*ability. https://doi.org/10.1007/s10668-024-04744-w

Wezel A, Casagrande M, Celette F, Vian J F, Ferrer A & Peigné J (2014). Agroecological practices for sustainable agriculture: A review. *Agronomy for Sustainable Development* 34: 1-20. https://doi.org/10.1007/s13593-013-0180-7

Wilson G (1989). Family food systems, preventive health and dietary change: a policy to increase the health divide. *Journal of Social Policy* 18(2): 167-185

Yılmaz H & Tanç Z A (2019). Biological control in pest management in Turkey: Comparison of the attributes of participant and non-participant greenhouse farmers in government-subsidized biological control practices. *Future of Food: Journal on Food, Agriculture and Society* 7(2): 1-8. https://doi.org/10.17170/kobra-20190709594

Yılmaz H (2014). An Analysis on Factors Influencing government supported Bumble bees use as pollinators by greenhouse producers' in The Mediterranean Coastal Region Of Turkey. *Acta Scientiarum Polonorum Hortorum Cultus* 13(6): 59-70

Yilmaz H (2015a) Farm level analysis of pesticide use in sweet cherry (Prunus avium L.) growing in west Mediterranean Region of Turkey. *Acta Scientiarum Polonorum Hortorum Cultus* 14(3): 115–129

Yilmaz H (2015b). Analysis in terms of environmental awareness of farmers' decisions and attitudes in pesticide use: The case of Turkey. *Bulg. Chem. Commun* 47: 771–775

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Environmental and Ecological Risks Posed by Sediment Heavy Metals in Reservoirs: A Preliminary Study from Northwest Türkiye

Murat Tekiner[a*], Tülay Tunçay[b], Mehmet Parlak[c]

[a]*Department of Agricultural Structures and Irrigation, Faculty of Agriculture, Çanakkale Onsekiz Mart University, 17800 Çanakkale, TURKIYE*

[b]*Soil Fertilizer and Water Resources Central Research Institute, 06172 Ankara, TURKIYE*

[c]*Department of Chemistry and Chemical Processing Technologies, Vocational School of Lapseki, Çanakkale Onsekiz Mart University, 17800 Çanakkale, TURKIYE*

ARTICLE INFO
Research Article
Corresponding Author: Murat Tekiner, E-mail: mtekiner@comu.edu.tr
Received: 19 May 2024 / Revised: 16 July 2024 / Accepted: 29 July 2024 / Online: 14 January 2025

Cite this article
Tekiner M, Tunçay T, Parlak M (2025). Environmental and Ecological Risks Posed by Sediment Heavy Metals in Reservoirs: A Preliminary Study from Northwest Türkiye. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):59-70. DOI: 10.15832/ankutbd.1486524

ABSTRACT
Reservoir sediments are an important component of aquatic ecosystems. Concentrations, sources, pollution and ecological risks of heavy metals pose serious risks on sustainable management of these ecosystems. This research focuses on heavy metal contents, physicochemical properties, environmental and ecological risks of sediments in four reservoirs (Ayvacık, Bayramdere, Bayramiç, and Umurbey) in Northwest Türkiye. Bayramiç reservoir had greater sediment Al, Cu, Fe, Mn concentrations, clay and silt contents than the other reservoirs (Ayvacık, Bayramdere, and Umurbey). In all four reservoirs, sediment heavy metals were generally of natural origin. Although sediment pollution index was identified as "considerable contamination" for Mn, such a case was not detected for the other heavy metals (Cd, Co, Cr, Cu, Fe, Ni, Pb, and Zn). An ecological risk assessment was made for reservoir sediments and a "low contamination" was detected.

Keywords: Pollution indices, Reservoirs, Sediment pollution, Ecological risk assessment

1. Introduction

Aquatic ecosystems such as oceans, seas, rivers and dams are under pressure of various factors including urbanization, industrialization, domestic and agricultural activities and mining activities (Fang et al. 2019; Githaiga et al. 2021; Muhammad 2023). Direct release of pollutants into the aquatic ecosystems results in serious destruction of existing ecosystems. Heavy metals can intrude into aquatic environments as pollutants and cause toxic effects on living ecosystem through the food chain (Xu et al. 2017; Rezapour et al. 2022; Şavran & Küçük 2022).

Dams and reservoirs have significant contributions to development of humanity. They are mostly constructed on rivers and streams for drinking water supply, irrigation of agricultural fields, flood control, electricity generation, industrial water supply, water quality improvement, recreational activities, river and inland waterway transportation, development and protection of fisheries, sediment retention and control (Alla & Liu 2021). Sediments accumulate in dam reservoirs in time and these dams complete their lifespan.

In aquatic environments, heavy metals enter into various physicochemical reactions. Sediment heavy metal concentrations may vary with the structure of sediment, particle size, specific surface area, and organic matter (Ma et al. 2023; Toller et al. 2022). With their distribution, persistence, non-degradability and toxicity, heavy metals pose a potential threat to aquatic ecosystems (Dong & Li 2023; Kang et al. 2024). Sediments constitute an accumulation site for heavy metals. Pollution loads of heavy metals threaten not only the habitats of aquatic ecosystems but also the entire ecosystem through accumulation and proliferation in the food chain (Li et al. 2024; Wang et al. 2022). Therefore, heavy metal concentrations of aquatic ecosystems should regularly be monitored.

A detailed research has not been conducted on sediment heavy metals of four reservoirs (Ayvacık, Bayramdere, Bayramiç, and Umurbey) in Northwest Türkiye. This study was conducted 1) to determine sediment heavy metal concentrations of four reservoirs used for different purposes (usually drinking and/or irrigation water supply), 2) to determine pollution status and ecological risks of each reservoir with the use of pollution and ecological risk indices, 3) to identify sources of sediment heavy

metals. Present hypothesis was set as “There is a low heavy metal pollution and ecological risk since there are no industrial facilities around the reservoirs; pollution sources are not coming from anthropogenic effects”.

2. Materials and Method

2.1. Study area

Ayvacık, Bayramdere, Bayramic and Umurbey reservoirs are located within the boundaries of Çanakkale Province in Northwest Türkiye (Figure 1).

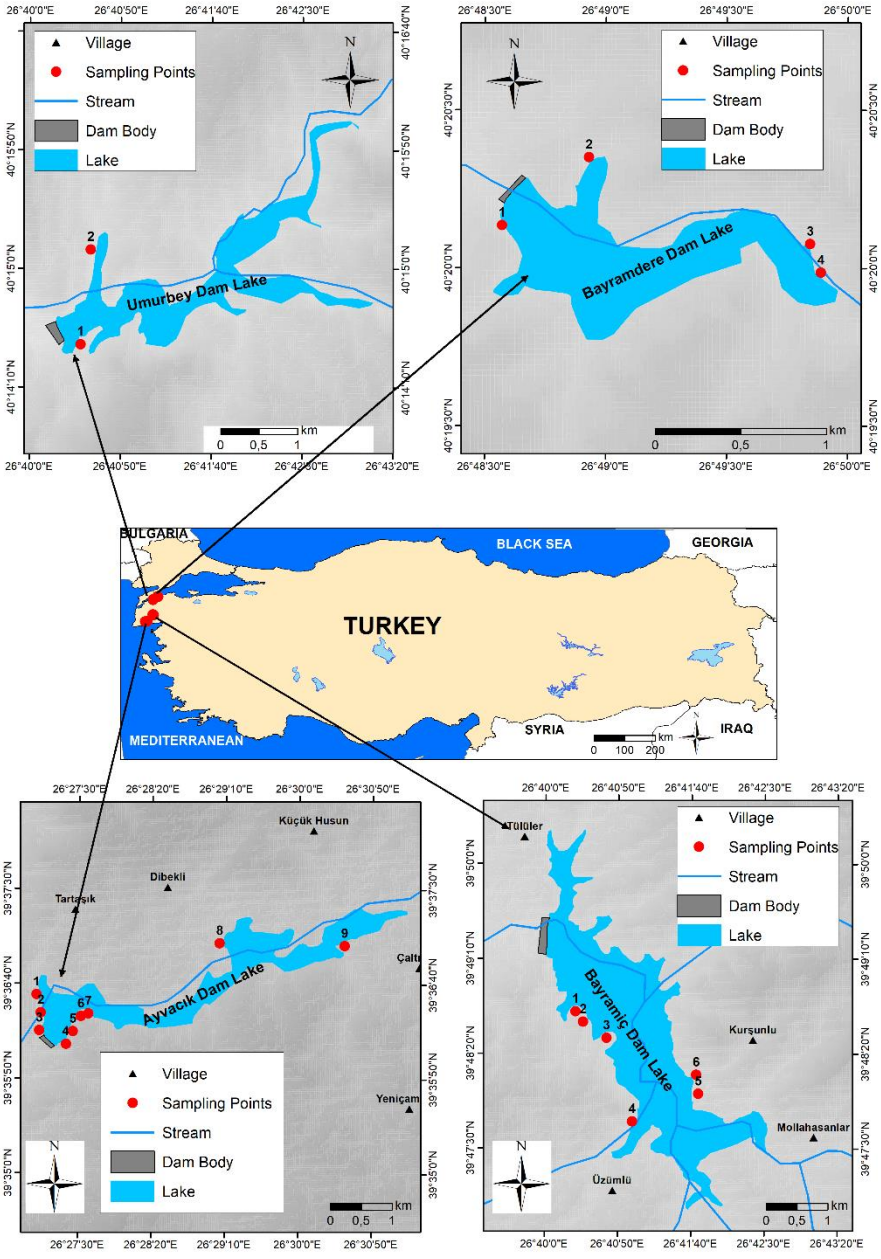


Figure 1- Location of four reservoirs and sediment sampling stations

There are no industrial facilities, but rural settlements and agricultural lands around the reservoirs. Ayvacık Reservoir (AYR) is located 8 km east of Ayvacık district center and was completed in 2008. It has a clay-core inside and sand-gravel embankment, a water storage capacity of 39 hm^3 and used for drinking and irrigation water supply (Taş et al. 2023) (Figure 2a). Geology of Ayvacık reservoir consists of melange, volcanic rocks and sedimentary rocks. Melange (complex series) was formed by sedimentary, volcanic and magmatic rocks being moved, torn off, dragged and stored in a certain place as a result of a certain effect. It is composed of radiolarite, mudstone, tuff, serpentinite, diabase, gabbro, marble, meta sand stone and limestone blocks. Volcanic rocks generally consist of altered andesite, agglomerate formed by the loose bonding of andesite blocks and pebbles,

mostly with tuff cement and basalt, a product of Plioquaternary basaltic magmatism. Sedimentary rocks include intercalated layers containing clayey limestones, claystone, sandstone, conglomerate and mudstones, old alluviums formed by the cementation of materials such as sand, gravel, silt and clay, slope debris formed by the deposition of andesite blocks and pebbles rolled down from the slopes at the foot of the slopes as a result of the erosion of agglomerates, alluviums formed through erosion and deposition of vegetative soil and rocks along the stream bed at lower elevations (Yağcı 1995).

Figure 2- Images of reservoirs from which sediment samples were taken (a: Ayvacık Reservoir, b: Bayramdere Reservoir, c: Bayramic Reservoir, d: Umurbey Reservoir) (DSİ 2023)

Bayramdere Reservoir (BDR) is located 19 km east of Lapseki District center and 7 km downstream of Hacıomerler village. It was completed in 2010 for irrigation and drinking water supply on Karanlık Stream. It has an irrigation area of 1050 ha. It has a clay core inside and rock-fill type embankment. It is 56 m high from the thalweg and 60 m high from the foundation (Taş et al. 2023) (Figure 2b). Geologically, Bayramdere Plain consists of Eocene flysch at the bottom, Neogene-aged sediments and vulcanite and alluviums above them. Bayramdere reservoir contains sedimentary rocks, volcanic rocks and metamorphic rocks. Çanakkale Formation (Cenozoic-Tertiary-Upper Miocene), which is observed in the form of conglomerate-sandstone in sedimentary rocks is composed of alluviums formed by the accumulation of the material carried by Bayramdere, old alluviums (stream terraces, Cenozoic-Quaternary) suspended as the stream bed deepens and slope debris (Cenozoic-Quaternary) formed by the accumulation of material from the stream on the plains and alluviums (Quaternary) formed through the accumulation of material rolling down steep slopes. Although volcanic rocks are found in most of the reservoir area, they are divided into two sections: andesite (plagioclase, hornblende, quartz and biotite) and andesitic tuff (feldspar, mica and glassy paste). Metamorphic rocks include schistosity developed as amphibolite, chlorite, serice and graphite schist (Kırmızı Erdal 2019).

Bayramic Reservoir (BCR), completed in 1975, is located 4 km northeast of Bayramic district center. It is an earth-fill dam with a surface area of 585 ha, an average depth of 8-10 meters and is fed by Küçük Menderes River, Çavuşlu and Ayazma streams. BCR was built for 92% irrigation, 4% energy and 4% drinking water purposes (Taş et al. 2023) (Figure 2c). BCR is located in Evciler Basin in the north of Kaz Mountains region and consists of Paleozoic metamorphic lands. Presence of Paleozoic metamorphics around the Evciler, Çırpılar, Yeşilköy villages and Sakardağı indicates that Evciler basin developed on metamorphics also called Kaz Mountain group. It is also covered by Oligocene volcanics (andesite, dacite, rhyolite-rhyodazite lavas and proclastics). Oligo-Miocene granite-granodiorite terrains are widely observed. Pliocene lands in Evciler basin consist of both volcanic and sedimentary rocks (Koç 2007).

Umurbey Reservoir (UBR) is located 17 km south of Lapseki district center, 6 km downstream from Umurbey Town. UBR, which was built in 2008 as a clay-core sand-gravel, rock fill type on Umurbey stream, irrigates 3661 ha agricultural area (Taş et al. 2023) (Figure 2d). The low areas located in the northern and southern parts of the research area consist of limestones formed in the Miocene and Pliocene and marine sediments deposited in the Tertiary period. While the coastal part of the research area

consists of Paleozoic granitoids and metamorphic rocks, upper parts are composed of Quaternary alluvial deposits. There are units consisting of marl, claystone and sandstone located between Umurbey district and southern part of Lapseki district. While the upper part of Umurbey Stream basin consists of Paleocene-Eocene aged volcanics, the lower parts consist of Pliocene sediments (Kırmızı Erdal 2019).

2.2. Sediment sampling and analytical procedures

Reservoir sediment samplings were carried out in August 2023. Since physical facilities were not available in present reservoirs, a small number of sediment samples could be taken. Therefore, every accessible sampling location was chosen a station. For sampling, 9 stations were selected in AYR, 4 stations in BDR, 6 stations in BCR and 2 stations in UBR (Figure 1). Station coordinates were taken with a GPS device (±5 m). A single sediment sampling was done at each station. Surface sediment samples (0-10 cm depth) were taken with a plastic shovel. Three sediment samples taken randomly at each station were mixed and turned into a single sample. Then, sediment samples weighing about 1-1.5 kg were placed into nylon bags and brought to the laboratory. Each sample was air-dried, passed through a 63 µm sieve for heavy metal analysis and stored at +4 ºC until analysis. Sediment samples were sieved through a 63 µm sieve since the particle size at which metal adsorption is most effective is <63 µm (Cüce et al. 2022). For other analyses such as particle size distribution, pH, EC, and organic matter, sediment samples were passed through a 2 mm sieve (El-Radaideh et al. 2017; Parlak et al. 2021).

Sediment samples were passed through acid-digestion process in 1/3 perchloric acid/nitric acid mixture and total heavy metals (Al, Cd, Co, Cu, Cr, Fe, Mn, Ni, Pb, and Zn) were determined in an ICP-OES device (Varian 710-ES model) (USEPA1996). Acid mixture-supplemented sediment samples were kept in a microwave oven (Mars 6 one touch technology model) at 180 ºC for 1 hour. Particle size distribution (clay, silt, fine sand, coarse sand percent) was determined with the use of Bouyoucous hydrometer and sieve (Gee & Or 2002). Sediment pH values were determined in saturation pastes in accordance with (Thomas 1996). Sediment EC values were also determined in saturation pastes with an EC-meter (Rhoades 1996). Organic matter content was determined with the modified Walkley-Black method (Nelson & Sommers 1996). Certified reference material (CRM) (NCS DC73371, sediment) was used to determine analytical accuracy and precision. Recovery percentages of the reference material varied between 93.98 - 100.55% (Table 1).

Table 1- The certificate values of the certified reference material (NCS DC73371, sediment), and the values (mean ± standard deviation) found in this study and the recovery rates, n = 3

Metals	*Certified (µg/g)*	*Determined (µg/g)*	*Recovery (%)*
Al	153600	155733±5021	98.69
Cd	0.11	0.13±0.01	100.55
Co	20	20.66±1.53	97.14
Cu	28	29.33±2.08	95.78
Cr	128	136.33 ± 5.50	93.98
Fe	65000	65833±1412	98.76
Mn	910	921±23.3	98.84
Ni	56	57±3.60	98.51
Pb	31	31.33±2.08	99.23
Zn	90	92.33±4.04	97.59

2.3. Sediment quality guidelines (SQG)

Sediment heavy metals may pose an ecological risk for aquatic organisms (Proshad et al. 2022). Therefore, a comparison was made with Threshold Effect Level (TEL) and Probable Effect Level (PEL) values to assess the harmful effects of heavy metals on benthic organisms. If the heavy metal concentrations are lower than TEL, they don't cause harmful effects; if the concentration is greater than PEL, harmful effects may occur. TEL and PEL values of heavy metals are provided in Table 4.

2.4. Evaluation of environmental and ecological risks

In this study, enrichment factor (EF), geoaccumulation index (I_{geo}), contamination factor (CF), and pollution load index (PLI) were used to determine the pollution level of sediments. Ecological risks of sediment heavy metals were calculated with potential ecological risk factor (Er^1), potential ecological risk index (RI), and toxic risk index (TRI). The equations, explanations and classifications used in calculation of these indices are given in Table 2.

2.5. Statistical analysis

Data normality was checked with Shapiro-Wilk test. While Cd, Fe, Mn, Pb, fine sand, coarse sand, and pH data showed normal distribution, other data (Co, Cr, Cu, Ni, Zn, clay, silt, and organic matter) did not show normal distribution. Some transformations (logarithmic transformation for Co, Cu, and organic matter; square root transformation for Cr, Ni, clay, and silt; sinus transformation for Zn) were made to make the data that did not show a normal distribution show a normal distribution. One-way

ANOVA test was performed to compare the differences between sediment heavy metal concentrations and physicochemical properties such as particle size distribution, pH, EC**,** and organic matter content. Significant means were compared with the use of Tukey's test ($P<0.05$). Multivariate statistical analyses such as principal component analysis (PCA) and correlation analysis were also performed. Kaiser-Meyer-Olkin (KMO) and Bartlett's tests were applied to determine the suitability of the data for PCA. Those with KMO values greater than 0.6 were used in PCA. Significance levels of 1% and 5% were taken into consideration in correlation tests. Statistical analyses were performed with use of SPSS 22 statistical software.

Table 2- Environmental and ecological indices used in this study*

Index	*Equation*	*Explanations*	*Classification*	*Contamination degree*
Enrichment factor, EF	$EF=\frac{(Cn/Al)Sample}{(Bn/Al)Background}$	where, C_n is the measured elemental concentration in sediment and B_n is the background heavy metal concentration (Jia et al. 2018)	EF<2 2≤EF<5 5≤EF<20 20≤EF<40	Minimal enrichment Moderate enrichment Significant enrichment Very high enrichment
Geoaccumulation index, I_{geo}	$Igeo = log2(\frac{Cn}{1.5xBn})$	where, C_n is the measured elemental concentration in sediment and B_n is the background heavy metal concentrations (Jia et al. 2018)	$I_{geo}\leq 0$ $0 < I_{geo} \leq 1$ $1 < I_{geo} \leq 2$ $2 < I_{geo} \leq 3)$ $3 < I_{geo} \leq 4$ $4 < I_{geo} \leq 5$ $I_{geo}>5$	Uncontaminated Uncontaminated to moderate contaminated Moderate contaminated Moderately to heavily contaminated Heavily contaminated Heavily to extremely contaminated Extremely contaminated
Contamination factor, CF	CF=Cn / Bn	where, C_n and B_n are the measured and background elemental concentrations in sediment (Hakanson 1980)	CF < 1 1 ≤ CF < 3 3 ≤ CF < 6 CF ≥ 6	Low contamination Moderate contamination Considerable contamination Very high contamination
Pollution load index, PLI	PLI = $(CF_1 \times CF_2 \times CF_n)^{1/n}$	where, CF is the contamination factor (Pobi et al. 2019) (n=8 in this study).	PLI ≤ 0 PLI > 0–1 PLI > 1	Perfection Baseline level Contaminated
Potential ecological risk factor, Er^i	$Er^i = Tr^i \times CF^i$	where, Tr^i is the biological toxic metal response factor (Cd = 30, Cu = 5, Cr = 2, Ni = 5, Pb = 2 and Zn = 1; Hakanson 1980) and CF is the contamination factor	$Er^i < 40$ $40 \leq Er^i < 80$ $80 \leq Er^i < 160$ $160 \leq Er^i <320$ $Er^i \geq 320$	Low risk Moderate risk Considerable risk High risk Very high risk
Potential ecological risk index, RI	$RI = \sum_{i=1}^{n} Er\imath$	where, Er^i is the potential ecological risk factor	RI < 150 150 ≤ RI <300 300 ≤ RI <600 RI ≥ 600	Low risk Moderate risk Considerable risk High risk
Toxic risk index, TRI	$TRIi = \sqrt{\frac{(C_i/TEL)^2 + (C_i/PEL)^2}{2}}$ $TRI = \sum_{i=1}^{n} TRI$	where, C_i is the measured content of heavy metal; TEL and PEL are threshold effect level and probable effect level of heavy metals, respectively; n is the number of heavy metals (Zhang et al. 2016)	TRI ≤ 5 5 < TRI ≤ 10 10 < TRI ≤ 15 15 < TRI ≤ 20 TRI > 20	No toxic risk Low toxic risk Moderate toxic risk Considerable toxic risk Very high toxic risk

***** In our study, Al was chosen as the heavy metal used for background due to its abundant content and stability in sediment (Canpolat et al. 2022).

3. Results and Discussion

3.1. Sediment heavy metal concentrations and physicochemical characteristics

Sediment heavy metal concentrations for studied 10 heavy metals and some physicochemical properties of 4 reservoirs are given in Table 3.

Table 3- Reservoir sediment heavy metal concentrations (in mg kg^{-1}) and some physicochemical characteristics*

	Al	*Cd*	*Co*	*Cr*	*Cu*	*Fe*	*Mn*	*Ni*	*Pb*	*Zn*
AYR										
Mean±std.dev.	6264±1843^{ab}	0.005±0.002	0.39±0.94	11.99±8.31^{a}	7.82±2.18^{b}	8788±1726^{ab}	1275±550^{b}	13.63±14.45	5.25±1.81	15.48±2.09^{b}
Min.-Max.	3665-8915	0.003-0.008	0.01-2.89	5.53-31.30	4.97±11.95	6375-11233	77-2080	1.91-47.81	3.48-12.27	12.27-19.28
BDR										
Mean±std.dev.	2888±1136^{b}	0.008±0.003	0.08±0.02	0.93±0.92^{b}	5.65±0.74^{b}	6363±1759^{b}	1865±611^{ab}	0.002±0.001	2.07±2.01	13.74±5.93^{b}
Min.-Max.	2015-4434	0.005-0.011	0.06-0.11	0.07-2.18	4.88-6.41	5300-8980	1241-2662	0.001-0.003	0.01-4.77	7.14-21.01
BCR										
Mean±std.dev.	10158±5220^{a}	0.008±0.005	0.10±0.07	9.10±4.72^{ab}	19.32±9.04^{a}	11522±3056^{a}	3047±1034^{a}	5.09±6.71	5.01±3.33	23.56±9.63^{ab}
Min.-Max.	4979-19505	0.003-0.014	0.03-0.22	5.17-18.15	13.02-35.22	7696-15549	1546-4167	0.98-18.47	0.01-9.46	15.10-38.83
UBR										
Mean±std.dev.	4763±317^{ab}	0.060±0.003	0.09±0.03	3.75±0.34^{ab}	15.42±7.90^{ab}	7124±214^{ab}	2380±892^{ab}	1.76±2.48	6.76±4.85	47.78±45.63^{a}
Min.-Max.	4538-4987	0.040-0.080	0.07-0.11	3.51-3.99	9.83-21.01	6973-7275	1749-3010	0.002-3.51	3.32-10.19	15.51-80.05

Table 3 (Continue)- Reservoir sediment heavy metal concentrations (in mg kg^{-1}) and some physicochemical characteristics*

	C	*Si*	*F.S*	*C.S*	*pH*	*EC*	*O.M*
AYR							
Mean±std.dev.	5.69±3.34^{b}	12.28±9.44^{b}	74.47±14.67^{a}	7.29±2.89	7.75±0.16	0.40±015^{b}	0.72±0.53
Min.-Max.	2.04-12.25	2.04-28.57	46.96-86.77	3.72-12.25	7.46-7.92	0.22-0.68	0.15-1.78
BDR							
Mean±std.dev.	12.22±6.19^{ab}	18.15±6.67^{ab}	64.42±10.70^{ab}	5.21±1.85	7.49±0.42	1.07±0.68^{a}	1.22±0.98
Min.-Max.	8.16-21.28	12.25-24.44	52.74-73.94	2.58-6.53	7.02-7.87	0.27-1.92	0.36-2.49
BCR							
Mean±std.dev.	23.92±11.38^{a}	33.64±20.87^{a}	34.69±27.66^{b}	7.73±2.83	7.34±0.32	1.02±0.33^{a}	1.76±1.75
Min.-Max.	6.12-34.69	8.16-64.58	8.96-78.75	4.37-12.31	6.87-7.70	0.42-1.39	0.25-4.40
UBR							
Mean±std.dev.	15.62±7.36^{ab}	12.50±2.95^{ab}	61.69±12.35^{ab}	10.18±2.04	7.57±0.08	0.73±0.07^{ab}	0.65±0.52
Min.-Max.	10.42-20.83	10.41-14.59	52.96-70.43	8.74-11.63	7.52-7.63	0.68-0.78	0.28-1.02

* C (Clay) (%), Si (Silt) (%), F.S (Fine sand) (%), C.S (Coarse sand) (%), EC (Electrical conductivity) (dS m^{-1}), O.M (Organic matter) (%): Means in the same column followed by the different letter for each criterion are significantly different at the 0.05 level (Tukey's test).

There were significant differences in average Al, Cr, Cu, Fe, Mn, Zn, clay, silt, fine sand, and EC of reservoir sediments ($P<0.05$). BCR had a higher Al concentration than BDR; AYR had a higher Cr concentration than BDR; BCR had a higher Cu concentration than AYR and BDR. BCR had a greater Fe concentration than BDR. BCR had a higher Mn content than AYR, UBR had a higher Zn content than AYR and BDR. Such differences in some heavy metal concentrations (Al, Cr, Cu, Fe, Mn, and Zn) were attributed to lithogenic properties. Buccione et al. (2021) indicated the reasons of Cr, Cu, and Zn in the Pietro del Pertusillo (Italy) reservoir sediments as geogenic/lithogenic processes. BCR had greater clay and silt contents than AYR. AYR had higher fine sand content than BCR. BDR and BCR had greater EC values than AYR. AYR sediments were classified as 55.56% sandy, 22.22% loamy-sand, 22.22% sandy-loam; BDR sediments were classified as 50% loamy sand, 25% sandy clay loam**;** BCR sediments were classified as 16.67% sandy loam, %16.67 loamy sand, 33.32% silt loam, 16.67% clay loam, 16.67% silt clay loam; UBR sediments were classified as 50% sandy-loam and 50% sandy-clay-loam. There were no significant differences in Cd, Co, Ni, Pb, coarse sand, pH, and organic matter contents of reservoir sediments. Sediment pH value was measured as 7.75 for AYR, 7.49 for BDR, 7.34 for BCR and 7.57 for UBR and they were all slightly alkaline.

3.2. Comparison of reservoir sediment heavy metal concentrations with the other studies

In this study, heavy metal concentrations of sediments taken from four reservoirs were compared with the other reservoirs of Türkiye and the reservoirs of the other countries (Table 4). While Al concentration of AYR, BDR and UBR was lower than in the other reservoirs of Türkiye, Al concentration of BDR was found to be higher than Süreyyabey reservoir and lower than the others (Çubuk II, Atıkhisar, and Değirmendere) and the average shale value. Present Cd concentrations were found to be lower than the average shale value of reservoirs and water bodies of Türkiye and other countries. Co concentrations were lower than Süreyya Bey, Çubuk II, Atıkhisar, Değirmendere, Wadi Al-Aqiq Reservoirs and the average shale value. While Cr concentrations of AYR and BCR were higher than reservoirs of Algeria and Poland and lower than the others (Çubuk II, Süreyyabey, Atikhisar, Alemşah, Değirmendere, Hammaz Grouz Dam, Wadi Al-Aqiq Water Reservoir Dam, Lake Nasser), Cr concentrations of BDR and UBR were lower than the reservoirs of other countries and higher than the average shale value of Türkiye. Fe concentrations of 4 reservoirs were lower than the other reservoirs (Süreyyabey, Çubuk II, Alemşah, Değirmendere, Wadi al-Aqiq, Konsin River, and Lake Nasser) and average shale value. Mn concentration of AYR was higher than the reservoirs of others countries,

except for Nigeria. Mn concentrations of BDR, BCR and UBR were higher than the other studies. While Ni concentration of AYR was higher than Alemşah Earth Fill Dam and Jezewoir Reservoirs and lower than the other water bodies, Ni concentrations of three reservoirs (BDR, BCR, and UBR) were lower than the mean values of the other reservoirs. Present Pb concentrations were lower than the values of the other reservoirs of Türkiye and the other countries, except for Algeria, and the average shale value. Zn concentrations of present reservoirs were lower than the other studies and the average shale value. Present findings revealed that there was no significant heavy metal contamination, except Mn, in the sediments of AYR, BDR, BCR and UBR. Mn, which makes up approximately 0.1% of the Earth's crust, is usually found in olivine, clay minerals, feldspar, apatite, anorthite and biotite minerals (Atabey 2015; Post 1999). Parlak et al. (2023) stated that Mn was formed by geochemical weathering of rocks (pedogenic processes). It is also estimated that sediment heavy metal concentrations of different reservoirs varied greatly with natural factors such as anthropogenic sources and rock weathering (Table 4). Heavy metal concentrations (Cd, Cu, Cr, Ni, Pb, and Zn) of the sediments sampled from reservoirs were determined to be lower than TEL and PEL. This result shows that heavy metals in sediments do not have a harmful effect on benthic organisms. Ustaoğlu et al. (2022) reported that TEL and PEL values for Pb, Cd, Zn, Cu, Ni, and Cr in Sera Lake Nature Park sediments did not pose a risk to sediment biota. Kankılıç et al. (2013) found that Zn, Cu, Pb, and Cd in Kapulukaya Dam Lake sediments did not have a negative effect on aquatic organisms.

Table 4- Comparison of present heavy metal concentrations with the heavy metal concentrations of the sediments in various parts of the world and sediment quality

Study location	*Al*	*Cd*	*Co*	*Cu*	*Cr*	*Fe*	*Mn*	*Ni*	*Pb*	*Zn*	*References*
ADL,Türkiye	6264	0.005	0.39	7.82	11.99	8788	1275	13.63	5.25	15.48	This study
BDL,Türkiye	2888	0.008	0.08	5.65	0.93	6363	1865	0.02	2.07	13.74	This study
BcDL,Türkiye	10158	0.08	0.10	19.31	9.10	11522	3048	5.09	5.01	23.56	This study
UDL,Türkiye	4763	0.06	0.09	15.42	3.75	7124	2380	1.76	6.76	47.78	This study
Süreyyabey Dam Lake, Türkiye	8455	0.10	105.5	61.70	201.5	52651	721	628.60	12.40	182.70	Erdoğan et al. 2023
Çubuk II Dam Lake, Türkiye	20725	-	21.14	31.23	48.81	37510	696	61.21	10.94	66.45	Fikirdeşici Ergen et al. 2021
Atıkhisar Dam Lake, Türkiye	22742	0.37	12.80	40.26	24.50	28021	732	29.10	35.51	74.40	Fural et al. 2021
Alemşah Earth Fill Dam, Türkiye	-	-	-	23.03	21.52	18505	296	9.44	8.39	30.99	Parlak et al. 2021
Degirmendere Dam Lake, Türkiye	20397	0.11	7.10	16.00	41.00	16161	387	22.00	9.00	26.00	Varol et al. 2022
Hammam Grouz Dam, Algeria	-	1.59	-	2.85	5.60	-	-	-	1.86	68.00	Aissaoui et al. 2017
Wadi Al-Aqiq Water Reservoir Dam, Saudi Arabia	-	-	14.20	50.29	50.58	22670	482	30.04	6.81	37.07	Alghamdi et al. 2019
Konsin River and Igboho Dam Reservoir, Nigeria	-	0.86	-	32.17	33.43	33850	1536	-	44.20	119.24	Asomba et al. 2023
Hussain Sagar Lake, India	-	19.89	-	90.11	90.00	-	-	47.04	79.89	273.14	Ayyanar & Thatikonda 2020
Lake Nasser, Egypt	-	0.18	-	21.78	30.79	12418	280	27.56	10.91	35.38	Goher et al. 2014
Jezewo Reservoir, Poland	-	0.40	-	10.10	6.50	-	-	5.90	17.6	903.70	Sojka et al. 2018
Mean Sediment Values	-	0.17	-	33.00	72.00	-	-	52.00	19.00	95.00	Salomons & Förtsner 1984
Average Shale Values	80000	0.30	19.00	45.00	90.00	47200	850	68.00	20.00	95.00	Turekian & Wedepohl1961
Threshold Effect Level (TEL)	-	0.59	-	35.70	37.30	-	-	18.00	35.00	123.00	Mac Donald et al. 2020
Probable Effect Level (PEL)	-	3.53	-	197.00	90.00	-	-	36.00	91.30	315.00	Mac Donald et al. 2020

3.3. Determination of sediment heavy metal sources

Correlation analysis and factor analysis were performed to determine the sources of heavy metals in reservoir sediments. Table 5 shows the correlation coefficients between heavy metals and some physicochemical properties (particle size distribution, pH, and organic matter) of reservoir sediments. Significant positive correlations was seen between Cd and silt (r= 0.44) ($P<0.05$); highly significant positive correlations between Co and Cr (r= 0.74), between Co and Ni (r= 0.82) ($P<0.01$); highly significant positive correlations between Cu and Fe (r= 0.71), between Cu and Mn (r= 0.68), between Cu and Zn (r= 0.64), between Cu and clay (r= 0.57), between Cu and silt (r= 0.68), between Cu and organic matter (r=0.67) ($P<0.01$); significant positive correlations between Cu and Pb (r= 0.43) ($P<0.05$); significant negative correlations between Cu and fine sand (r= 0.69) ($P<0.01$); significant negative correlations between Cu and pH (r= 0.53) ($P<0.05$); significant positive correlations between Cr and Ni (r= 0.96) ($P<0.01$); significant positive correlations between Fe and Mn (r= 0.64) ($P<0.01$); significant positive correlations between Fe and silt (r=0.53), between Fe and organic matter (r=0.60) ($P<0.05$); significant negative correlations between Fe and fine sand (r= 0.47) ($P<0.05$); significant positive correlations between Mn and clay (r= 0.79), between Mn and silt (r= 0.78), between Mn and organic matter (r=0.65) ($P<0.01$); significant negative correlations between Mn and fine sand (r= 0.82) ($P<0.01$); significant negative correlations between Mn and pH (r=0.49) ($P<0.05$); significant positive correlations between Pb and Zn (r= 0.56)

(P<0.01); highly significant negative correlations between clay and silt (r= 0.73) (P<0.01); significant positive correlations between clay and pH (r= 0.50), between clay and organic matter (r= 0.51), (P<0.05); significant negative correlations between silt and fine sand (r= 0.95), between silt and organic matter (r= 0.69), and between pH and organic matter (r= 0.69) (P<0.01). Fikirdeşici Ergen et al. (2021) worked on sediments of reservoirs in Ankara and reported highly significant positive correlations between Co and Ni (P<0.01); significant positive correlations between Cu and Fe (P<0.05); between Cu and Zn (P<0.01); between Pb and Zn (P<0.01); highly significant positive correlations between Ni and Cr (P<0.01) and significant positive correlations between Fe and Mn (P<0.05). Sojka et al. (2023) conducted research on the sediments of reservoirs in Poland and reported a strong relationship between Ni and Cr (r= 0.94). Researchers reported that no correlation was detected between heavy metals and pH, EC, and organic matter of the sediments.

Table 5- Correlation matrix for correlations between sediment heavy metals, particle size distribution, pH, and organic matter

	Cd	*Co*	*Cu*	*Cr*	*Fe*	*Mn*	*Ni*	*Pb*
Cd	1							
Co	0.12	1						
Cu	0.15	-0.01	1					
Cr	-0.12	0.74**	0.30	1				
Fe	-0.04	-0.13	0.71**	0.37	1			
Mn	0.27	-0.17	0.68**	0.05	0.64**	1		
Ni	-0.12	0.82**	0.14	0.96**	0.16	-0.10	1	
Pb	-0.14	-0.07	0.43*	0.20	0.33	0.17	0.11	1
	Zn	*C*	*Si*	*F.S*	*C.S*	*pH*	*O.M*	
Zn	1							
C	0.17							
Si	0.19	0. 073**	1					
F.S	-0.20	0.0.90**	0.95**	1				
C.S	0.09	00.12	0.03	-0.19	1			
pH	-0.22	0 0.50*	-0.41	0.49*	-0.17	1		
OM	0.23	0 0.51*	0.75**	0.69**	0.01	-0.69**	1	

Heavy metals (mg kg^{-1}), C: Clay (%), Si: Silt (%), F.S: Fine sand (%), C.S: Coarse sand (%), O.M: Organic matter (%): * P<0.05: ** P<0.01

In PCA, 3 components with eigenvalues values greater than 1 were identified (Table 6). These 3 components explained 74.83% of the total variance. Component 1, which is responsible for 35.62% of the total variance, consists of Cu, Fe and Mn. Of these 3 heavy metals, Fe and Mn are the most abundant in Earth's crust (Kabata-Pendias 2011; Algül & Beyhan 2020). Although Cu was generally of anthropogenic origin (Bhuyan et al. 2023), it showed a significant positive correlation with Fe. Such a case shows that Cu is of natural origin. Low CF values of Cu also support this phenomenon. The second component explained 21.16% of the total variance and showed strong positive loading values for Cr and Ni. The EF and CF values of Cr and the EF, I_{geo} and CF values of Ni, except for AYR, were found to be low (Table 7). In addition, Cr and Ni concentrations in sediments of 4 reservoirs did not exceed PEL values. Varol et al. (2022) indicated that Cr and Ni in sediments of three stagnant water bodies in Northern Türkiye were of natural origin. Furthermore, Tumuklu et al. (2023), working on Gümüşler Reservoir sediments, indicated the origin of Cr and Ni as the weathering of mafic rocks in the research area. Component 3, which showed strong positive loading values for Co and Zn, explained 18.06% of the total variance (Table 7). In sediments of all reservoirs, EF, I_{geo} and CF values were determined to be low for Co and only CF values were determined to be low for Zn (Table 7). Additionally, Zn concentrations of the sediments did not exceed PEL values (Table 4). Therefore, it was concluded that Co and Zn were of geogenic origin. Cüce et al. (2022) indicated the terrestrial origin of Co & Bhuyan et al. (2023) stated that Zn is naturally present in Earth's crust. Additionally, Canpolat et al. (2022) reported that heavy metals (Co, Cr, Cu, Fe, Mn, Ni, and Zn) in Keban reservoir (Türkiye) sediments originate from lithological units of the study area.

Table 6- Varimax rotated component matrix for heavy metals*

	Component 1	*Component 2*	*Component 3*
Cd	0.310	-0.551	0.520
Co	0.037	0.318	**0.699**
Cu	**0.895**	0.171	0.071
Cr	0.289	**0.895**	0.235
Fe	**0.807**	0.336	-0.072
Mn	**0.882**	-0.178	0.094
Ni	0.097	**0.921**	0.260
Pb	0.427	0.386	-0.504
Zn	0.035	0.155	**0.759**
Eigenvalues	3.206	1.904	1.625
% of variance	35.622	21.158	18.052
% cumulative variance	35.622	56.780	74.832
Kaiser- Meyer- Olkin measure of sampling adequacy			0.667
Bartlett's test of sphericity			0.000

*: Bold values are factor loadings of the principal components

Table 7- Enrichment factor (EF), geoaccumulation index (I_{geo}), contamination factor (CF), potential ecological risk factor (Er'), pollution load index (PLI), potential ecological risk index (RI), and toxic risk index (TRI) values of sediment heavy metals

	Cd	*Co*	*Cu*	*Cr*	*Mn*	*Ni*	*Pb*	*Zn*
EF values								
AYR	0.65	0.33	4.22	1.41	22.42	3.25	4.64	3.19
BDR	2.44	0.10	6.68	0.19	74.98	0.001	4.43	6.19
BCR	0.79	0.03	6.45	0.57	35.99	0.55	2.55	3.02
UBR	0.98	0.06	10.15	0.50	56.36	0.55	7.46	11.91
I_{geo} values								
AYR	-3.41	0.16	1.06	1.54	2.67	1.30	0.75	1.49
BDR	-3.27	0.11	0.97	0.60	2.67	-1.23	0.07	1.44
BCR	-3.36	0.10	1.30	1.50	2.67	0.99	0.49	1.57
UBR	-3.42	0.15	1.25	1.31	2.64	-0.02	0.80	1.65
CF values								
AYR	0.05	0.16	0.31	0.09	1.77	0.24	0.35	0.24
BDR	0.08	0.003	0.23	0.008	2.60	0.0001	0.14	0.21
BCR	0.07	0.004	0.77	0.07	4.25	0.09	0.34	0.36
UBR	0.06	0.004	0.62	0.025	3.32	0.03	0.45	0.74
E_r values								
AYR	1.4	-	1.56	0.19	-	1.23	0.71	0.24
BDR	2.5	-	1.14	0.02	-	0.0002	0.28	0.21
BCR	2.3	-	3.85	0.14	-	0.45	0.68	0.36
UBR	1.8	-	3.08	0.05	-	0.15	0.91	0.74
		AYR		BDR		BCR		UBR
PLI values		2.49		3.04		5.38		4.78
RI values		47.94		16.47		46.73		13.44
TRI values		10.95		1.09		6.37		1.83

3.4. Contamination degree and ecological risk indices of heavy metals

The average EF, I_{geo}, CF, PLI, Er, RI and TRI values calculated for the sediments of each reservoir are given in Table 7. EF and I_{geo} indices were used to detect natural and anthropogenic sources of sediment heavy metals (Aykır et al. 2023; Kükrer et al. 2020). While the EF values of Cd, Co and Cr in AYR were below 2 (‘minimal enrichment’ level), the EF values of Cu, Ni, Pb and Zn varied between 2–5 (‘moderate enrichment’ level), and the EF value of Mn was determined at a very high enrichment level. The EF values of Co, Cr and Ni in BDR were determined at 'minimal enrichment' level, the EF values of Cd and Pb were between 2-5 (‘moderate enrichment’ level) and the EF of Mn was determined ‘very high enrichment’ level. While the EF values of Cd, Co, Cr and Ni in BCR and UBR were <2 (‘minimal enrichment’ level), the EF values of Cu were between 5-20 (‘significant enrichment’ level) and the EF values of Mn were determined ‘very high enrichment’ level. In four reservoirs, the I_{geo} values of Cd were found to be negative (‘uncontaminated’ level), the I_{geo} values of Co and Pb were determined to be between 0-1 (‘uncontaminated to moderate contaminated’ level), and the I_{geo} values of Mn were determined to be between 2–3 (‘moderately to heavily contaminated’ level). The I_{geo} values of Cu, Cr, Ni and Zn in AYR were determined to be between 1–2 (‘moderate contaminated’ level). The I_{geo} values of Cr and Pb in BDR were between 0–1 ('uncontaminated to moderately contaminated' level), the I_{geo} value of Ni was <0 (‘uncontaminated’ level) and the I_{geo} values of Zn were between 1–2 (‘moderate contaminated’ level). In BCR and UBR, the I_{geo} values of Cu, Cr and Zn were determined to be between 1–2 (‘moderate contaminated’ level), while the I_{geo} values of Pb were between 0–1 (‘uncontaminated to moderate contaminated’ level). CF was used to determine the degree of contamination of each heavy metal of the sediment (Hakanson 1980). Except for Mn, CF values of other heavy metals were found to be <1 (‘low contamination’ level). The CF values of Mn varied between 1–3 (‘moderate contamination’ level), in AYR and BDR and between 3–6 (‘considerable contamination’ level) in BCR and UBR. Yüksel et al. (2024) determined EF, CF, and Igeo values in the sediments of Almus Dam Lake (Turkey) as low to no contamination.

The Er value is used to determine the potential ecological risk of each heavy metal in the sediment (Hakanson, 1980). Er values varied between 1.56 (Cu) and 0.19 (Cr) in AYR, between 2.5 (Cd) and 0.0002 (Ni) in BDR, between 3.85 (Cu) and 0.14 (Cr) in BCR, between 3.08 (Cu) and 0.05 (Cr) in UBR. Er values are determined in the low-risk class for heavy metals. PLI provides an overall assessment of heavy metal contamination in sediment (Zoidou & Sylaious 2021). PLI values were >1 in all reservoirs, indicating that they were 'contaminated'. Potential ecological risk index (RI) was used to determine the total ecological risk of heavy metals (Aykır et al. 2023). RI values were determined as 47.94 for AYR, 16.47 for BDR, 46.73 for BCR, and 13.44 for UBR. Such a case showed that four reservoirs were in the low-risk class. Toxic risk index (TRI) is used to determine the toxic effects of elements (Zhang et al. 2016). TRI values were determined as 1.09 for BDR and 1.83 for UBR. In BDR and UBR, heavy metals were detected in the "no toxic risk" class. TRI values were determined as 10.95 (moderate toxic risk) in AYR 6.37 in BCR (low toxic risk). Fural et al. (2020) reported TRI values in the sediments of İkizcetepeler Dam Lake (Turkey) as between 4 - 6.6. The researchers reported that TRI in sediment samples were in no toxic risk and low toxic risk class.

4. Conclusions

In this study, which is important in terms of determining and risk assessment of heavy metal-containing sediments, which are one of the risks of ecological diversity, sediment sources, sediment heavy metal concentrations, physicochemical characteristics, contamination status of heavy metals and potential ecological risks were investigated in four reservoirs used for drinking and irrigation water supply in northwest Türkiye. There were significant differences in sediment Al, Cr, Cu, Fe, Mn, and Zn concentrations, particle size distribution, and EC values. Except for Mn, sediment heavy metal concentrations were generally lower than the other water bodies. Multivariate statistical analyses (correlation analysis and PCA) were also performed in this study. Correlation analyses revealed significant and highly significant (1% and 5% level) correlations between majority of the heavy metals. Such a case indicated that contamination sources were coming from the same and similar transport processes. PCA revealed the natural sources of heavy metals. Except for Mn, EF values of heavy metals were within the range of minimal enrichment and moderate enrichment, I_{geo} values of heavy metals were within the range of uncontaminated ile moderate uncontaminated, and CF values were determined to be low contamination. Er, RI and TRI values of reservoir sediments indicate low ecological risk; PLI values were determined as "contaminated degree". For better risk assessment of heavy metals in similar ecosystems, chemical fractionation of sediments should be performed. For sustainable management of reservoirs, sediment heavy metal concentrations should be monitored regularly.

References

Aissaoui A, Ahmed D S A, Cherchar N & Gherib A (2017). Assessment and biomonitoring of aquatic pollution by heavy metals (Cd, Cr, Cu, Pb and Zn) in Hammam Grouz Dam of Mila (Algeria). *International Journal of Environmental Studies* 74(3): 428-442. http://doi.org/10.1080/00207233.2017.1294423

Alghamdi B A, Mannoubi I E & Zabin S A (2019). Heavy metals' contamination in sediments of Wadi Al-Aqiq water reservoir dam at Al-Baha region, KSA: Their identification and assessment. *Human and Ecological Risk Assessment: An International Journal* 25(4): 793-818. http:// doi.org/10.1080/10807039.2018.1451746

Algül F & Beyhan M (2020). Concentrations and sources of heavy metals in shallow sediments in Lake Bafa, Turkey. Scientific Reports 10: 11782. http://doi.org/10.1038/s41598-020-68833-2

Alla Y M K & Liu L (2021). Impacts of dams on the environment: A review. *International Journal of Environment, Agriculture and Biotechnology* 6(1): 64-74. doi.org/10.22161/ijeab.61.9

Asomba H C, Ezewudo B I, Okeke C J & Islam M S (2023). Grain size analysis and ecological risk assessment of metals in the sediments of Konsin River and Igboho dam reservoir, Oyo State, Nigeria, under agricultural disturbances. Environmental Monitoring and Assessment 195: 378. http:// doi.org/10.1007/s10661-023-11009-y

Atabey E (2015). Elements and their effects on health. Hacettepe University Mesothelioma and Medical Geology Research and Application Center Publications No: 1, (in Turkish) Ankara.

Aykır D, Fural Ş, Kükrer S & Mutlu Y E (2023). Havran Lagünü'nde (Balıkesir) ekolojik risk seviyesinin zamansal değişimi. *Coğrafya Dergisi* 46: 123-135. http:// doi.org/10.26650/JGEOG2023-1180818

Ayyanar A & Thatikonda S (2020). Distribution and ecological risks of heavy metals in Lake Hussain Sagar, India. Acta Geochimica 39(2): 255–270. http://doi.org/10.1007/s11631-019-00360-y

Bhuyan M S, Haider S M B, Meraj G, Bakar M A, Islam M T, Kunda M, Siddique MA B, Ali M M, Mustary S, Mojumder I A & Bhat M A (2023). Assessment of heavy metal contamination in beach sediments of Eastern St. Martin's Island, Bangladesh: implications for environmental and human health risks. Water 15: 2494. http://doi.org/10.3390/w15132494

Buccione R, Fortunato E, Paternoster M, Rizzo G, Sinisi R, Summa V & Mongelli G (2021). Mineralogy and heavy metal assessment of the Pietra del Pertusillo reservoir sediments (Southern Italy). Environmental Science and Pollution Research 28: 4857–4878. http://doi.org/10.1007/s11356-020-10829-6

Canpolat Ö, Varol M, Öztekin OÖ & Eriş KK (2022). Sediment contamination by trace elements and the associated ecological and health risk assessment: A case study from a large reservoir (Turkey). Environmental Research 204: 112145. http://doi.org/10.1016/j.envres.2021.112145

Cüce H, Kalıp E, Ustaoğlu F, Dereli MA & Türkmen A (2022). Integrated spatial distribution and multivariate statistical analysis for assessment of ecotoxicological and health risks of sediment metal contamination, Ömerli Dam (Istanbul, Turkey). Water Air and Soil Pollution 233: 199. http://doi.org/10.1007/s11270-022-05670-1

Dong S L & Li L (2023). Sediment and Remediation of Aquaculture Ponds. In: Dong SL, Tian XL, Gao QF & Dong YW. (Eds) Aquaculture Ecology. Springer, Singapore. https://doi.org/10.1007/978-981-19-5486-3_8

DSİ (2023). DSI has built 20 dams and 8 ponds in Çanakkale in the last 18 years. [cited 2023 Sep 28] Available from:http://www.dsi.gov.tr/Galeri/ResimgaleriDetay/1236 (in Turkish).

El-Radaideh N, Al-Taani AA & Al Khateeb W A (2017). Characteristics and quality of reservoir sediments, Mujib Dam, Central Jordan, as a case study. Environmental Monitoring and Assessment 189: 143. http://doi.org/10.1007/s10661-017-5836-3

Erdoğan Ş, Başaran Kankılıç G, Seyfe M, Tavşanoğlu ÜN & Akın Ş (2023). Assessment of heavy metal pollution with different indices in Süreyyabey dam lake in Turkey. Chemistry and Ecology 39: 153-172. http://doi.org/10.1080/02757540.2022.2162045

Fang T, Yang K, Lu W, Cui K, Li J, Liang Y, Hou G, Zhao X & Li H (2019). An overview of heavy metal pollution in Chaohu Lake, China: enrichment, distribution, speciation, and associated risk under natural and anthropogenic changes. Environmental Science and Pollution Research 26: 29585–29596. http://doi.org/10.1007/s11356-019-06210-x

Fikirdeşici Ergen Ş, Tekatlı Ç, Gürbüzer P, Üçüncü Tunca E, Türe H, Biltekin D, Kurtuluş B & Tunca E (2021). Elemental accumulation in the surficial sediment of Kesikköprü, Çubuk II and Asartepe Dam Lakes (Ankara) and potential sediment toxicity. Chemistry and Ecology 37: 552-572. http://doi.org/10.1080/02757540.2021.1902509

Fural Ş, Kükrer S & Cürebal İ (2020). Geographical information systems based ecological risk analysis of metal accumulation in sediments of İkizcetepeler Dam Lake (Turkey). Ecological Indicators 119: 106784. https://doi.org/10.1016/j.ecolind.2020.106784

Fural Ş, Kükrer S, Cürebal İ & Aykır D (2021). Spatial distribution, environmental risk assessment, and source identification of potentially toxic metals in Atikhisar dam, Turkey. Environmental Monitoring and Assessment 193: 208. http://doi.org/10.1007/s10661-021-09062-6

Gee GW & Or D (2002). Particle-size analysis. In: Dane JH, Topp GC, editors. Methods of soil analysis. Part 4, Physical methods. SSSA Book Series 5. Soil Science Society of America, Madison, Wisconsin, USA; 2002. p. 255–293.

Githaiga K B, Njuguna S M, Gituru R W & Yan X (2021). Water quality assessment, multivariate analysis and human health risks of heavy metals in eight major lakes in Kenya. *Journal of Environmental Management* 297: 113410. https://doi.org/10.1016/j.jenvman.2021.113410

Goher M E, Farhat H I, Abdo M H & Salem S G (2014). Metal pollution assessment in the surface sediment of Lake Nasser, Egypt. *Egyptian Journal of Aquatic Research* 40: 213–224. http:// doi.org/10.1016/j.ejar.2014.09.004

Hakanson L (1980). An ecological risk index for aquatic pollution control: A sedimentological approach. Water Research 14: 975-1001.

Jia Z, Li S & Wang L (2018). Assessment of soil heavy metals for eco-environment and human health in a rapidly urbanization area of the upper Yangtze Basin. Scientific Reports 8: 3256. http:// doi.org/10.1038/s41598-018-21569-6

Kabata-Pendias A (2011). Trace elements of soils and plants (4th ed). CRC Press, Taylor & Francis Group, LLC.

Kang G, Chen H, Hu C, Wang F & Qi Z (2024). Spatiotemporal Distribution Characteristics and Influencing Factors of Dissolved Potentially Toxic Elements along Guangdong Coastal Water, South China. *Journal of Marine Science Engineering* 12: 896. https://doi.org/10.3390/ jmse12060896

Kankılıç G B, Tüzün İ & Kadıoğlu Y K (2013). Assessment of heavy metal levels in sediment samples of Kapulukaya Dam Lake (Kırıkkale) and lower catchment area. Environmental Monitoring and Assessment 185: 6739–6750. http:// doi.org/10.1007/s10661-013-3061-2

Kırmızı Erdal C (2019). Agriculture geography of Umurbey River Basin (Çanakkale). (master's thesis) Canakkale Onsekiz Mart University, Social Sciences Institute, Geography Department, (in Turkish); 2019

Koç T (2007). Geomorphology of the northern part of Kaz Mountain (Bayramiç-Çanakkale). Journal of Geographical Sciences 5(2): 27-53, (in Turkish).

Kükrer S, Erginal A E, Kılıç Ş, Bay Ö, Akarsu T & Öztura E (2020). Ecological risk assessment of surface sediments of Çardak Lagoon along a human disturbance gradient. Environmental Monitoring Assessment 192: 359. http://doi.org/10.1007/s10661-020-08336-9

Li J, Yang S, Wana F,Gao M, He L, Zhoa G, Ye S, LiuY & Hu K (2024). Ecological risk assessment of heavy metal(loid)s in riverine sediments along the East China Sea: A large-scale integrated analysis. Marine Pollution Bulletin 203: 116382. https://doi.org/10.1016/j.marpolbul.2024.116382

Ma H, Zhang Y, Liu Z, Chen Y& Lv G (2023). Pollution characteristics of heavy metals in surface sediments of the Shuimo River in Urumqi, China. Metals 13:1578. http:// doi.org/10.3390/ met13091578

Mac Donald D D, Ingersoll C G & Berger (2020). Development and evaluation of consensus-based sediment quality guidelines for freshwater ecosystems. Archieves of Environmental Contamination and Toxicology 39: 20-31. http://doi:10.1007/s002440010075

Muhammad S (2023). Evaluation of heavy metals in water and sediments, pollution, and risk indices of Naltar Lakes, Pakistan. Environmental Science and Pollution Research 30: 28217–28226. doi.org/10.1007/s11356-022-24160-9

Nelson R E & Sommers L E (1996). Total carbon, organic carbon and organic matter. In: Sparks DL, editors. Methods of soil analysis. Part 3. Chemical methods. American Society of Agronomy, Madison, Wisconsin, USA. p. 961-1010.

Parlak M, Everest T & Tunçay T (2021). Investigation of sediments of Uluköy and Alemşah irrigation ponds (Çanakkale-Türkiye) in terms of heavy metal pollution. KSÜ Agriculture and Nature Journal 2021;24 (2): 372-378 (in Turkish). doi.org/10.18016/ksutarimdoga.vi.752777

Parlak M, Taş İ, Görgişen C & Gökalp Z (2023). Spatial distribution and health risk assessment for heavy metals of the soils around coal-fired power plants of northwest Turkey. *International Journal of Environmental Analytical Chemistry* doi.org/10.1080/03067319.2023.2243231

Pobi K K, Satpati S, Dutta S, Nayek S, Saha R N & Gupta S (2019). Sources evaluation and ecological risk assessment of heavy metals accumulated within a natural stream of Durgapur industrial zone, India, by using multivariate analysis and pollution indices. Applied Water Science 9: 58. http://doi.org/10.1007/s13201-019-0946-4

Post J E (1999). Manganese oxide minerals: Crystal structures and economic and environmental significance. Proceedings of the National Academy of Sciences, 96(7): 3447-3454.

Proshad R, Kormoker T, Al M A, Islam M S, Khadka S & Idris A M (2022). Receptor model-based source apportionment and ecological risk of metals in sediments of an urban river in Bangladesh. *Journal of Hazardous Materials* 423:127030. http://doi.org/10.1016/j.jhazmat.2021.127030

Rezapour S, Asadzadeh F, Nouri A, Khodaverdiloo H & Heidari M (2022). Distribution, source apportionment, and risk analysis of heavy metals in river sediments of the Urmia Lake basin. Scientific Reports 12: 17455. http://doi.org/10.1038/s41598-022-21752-w

Rhoades J D (1996). Salinity: Electrical conductivity and total dissolved solids. In: Sparks DL, editors. Methods of soil analysis. Part 3. Chemical methods. American Society of Agronomy, Madison, Wisconsin, USA; 1996. p. 417-436.

Salomons W & Förstner U (1984). Metals in the hydrocycle. Springer-Verlag Berlin Heidelberg; 1984.

Sojka M, Jaskuła J & Siepak M (2018). Heavy metals in bottom sediments of reservoirs in the Lowland Area of Western Poland: Concentrations, distribution, sources and ecological risk. Water 11: 56. http://doi:10.3390/w11010056

Sojka M, Ptak M, Jaskuła J & Krasniqi V (2023). Ecological and health risk assessments of heavy metals contained in sediments of Polish Dam Reservoirs. *International Journal of Environmental Research and Public Health* 20: 324. https://doi.org/10.3390/ijerph20010324

Şavran G & Küçük F (2022). Heavy metal accumulation in aquatic organisms and its effects. Akademia Journal of Nature and Human Sciences 8: 65-78, (in Turkish) dergipark.org.tr/tr/pub/adibd/ issue/68882/1165848

Taş İ, Büyükgaga H İ, Tekiner M, Çamoğlu G, Yavuz M Y, Kızıl Ü, Yıldırım M, Erken O, Aksu S, İnalpulat M, Mucan U & Nar H (2023). Water storage structures and water potentials of Çanakkale province. In: Şeker M, Kahrıman F, Sungur A, Polat B (eds.). Çanakkale Agriculture. Free Publishing Distribution Co. Ltd. Gaziantep, Turkiye, ss. 839-908 (in Turkish).

Thomas G W (1996). Soil pH and soil acidity. In: Sparks DL, editors. Methods of soil analysis. Part 3. Chemical methods. American Society of Agronomy, Madison, Wisconsin, USA. pp. 475-490

Toller S, Funari V, Zannoni D, Vasumini I & Dinelli E (2022). Sediment quality of the Ridracoli freshwater reservoir in Italy: insights from aqua regia digestion and sequential extractions. Science of the Total Environment 826:154167. https://doi.org/10.1016/j.scitotenv.2022.154167

Tumuklu A, Sunkari E D, Yalcin F & Atakoğlu O O (2023). Data analysis of the Gumusler Dam Lake Reservoir soils using multivariate statistical methods (Nigde, Türkiye). *International Journal Environmental Science and Technology* 20: 5391-5404. http://doi.org/10.1007/s13762-022-04519-8

Turekian K K & Wedepohl K H (1961). Distribution of the elements in some major units of the earth's crust. Geological Society of America Bulletin 72: 175–192.

USEPA (United States Environmental Protection Agency) (1996). Method 3050B: Acid Digestion of Sediments, Sludges, and Soils (Revision 2) In: United States Environmental Protection Agency, Washington DC.

Ustaoğlu F, İslam Md S & Tokatlı C (2022). Ecological and probabilistic human health hazard assessment of heavy metals in Sera Lake Nature Park sediments (Trabzon, Turkey). *Arabian Journal of Geosciences* 15: 597. https://doi.org/10.1007/s12517-022-09838-1

Varol M, Ustaoğlu F & Tokatlı C (2022). Ecological risk assessment of metals in sediments from three stagnant water bodies in Northern Turkey. Current Pollution Reports 8: 409–421. http:// doi.org/10.1007/s40726-022-00239-2

Wang L, Wan X, Chen H, Wang Z & Jia X (2022). Oyster arsenic, cadmium, copper, mercury, lead and zinc levels in the northern South China Sea: Long-term spatiotemporal distributions, combined effects, and risk assessment to human health. Environmental Science and Pollution Research 29: 12706–12719. https://doi.org/10.1007/s11356-021-18150-6

Xu Y, Wu Y, Han J & Li P (2017). The current status of heavy metal in lake sediments from China: Pollution and ecological risk assessment. Ecology and Evoluation 7: 5454–5466. https://doi.org/10.1002/ece3.3124

Yağcı B (1995). The geotechnical studies of Ayvacık Dam. M. Sc. Thesis. Balıkesir University, Institute of Science, Department of Civil Engineering, p. 99, (in Turkish).

Yüksel B, Ustaoğlu F, Aydın H, Tokatlı C, Toplademir H, İslam Md S & Muhammed S (2024). Appraisal of metallic accumulation in the surface sediment of a fish breeding dam in Türkiye: A stochastical approach to ecotoxicological risk assessment. Marine Pollution Bulletin 203: 116488. http:// doi.org/10.1016/j.marpolbul.2024.116488

Zhang G, Bai J, Zhao Q, Lu Q, Jia J & Wen X (2016). Heavy metals in wetland soils along a wetland-forming chronosequence in the Yellow River Delta of China: Levels, sources and toxic risks. Ecol Indic 69: 331-339. http://doi.org/10.1016/j.ecolind.2016.04.042

Zoidou M & Sylaious G (2021). Ecological risk assessment of heavy metals in the sediments of a Mediterranean lagoon complex. *Journal of Environmental Health Science and Engineering* 19:1835–1849. doi.org/10.1007/s40201-021-00739-1

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Application of the Machine Learning Methods to Assess the Impact of Physico-chemical Characteristics of Water on Feed Consumption in Fish Farms

Nedim Özdemir[a], Mustafa Çakir[b], Mesut Yilmaz[c], Hava Şimşek[a], Mükerrem Atalay Oral[d], Okan Oral[e]*

[a]*Faculty of Fisheries, Muğla Sıtkı Koçman University, Muğla, TURKEY*

[b]*Iskenderun Vocational School of Higher Education, İskenderun Technical University, İskenderun, TURKEY*

[c]*Faculty of Fisheries, Akdeniz University, Antalya, TURKEY*

[d]*Faculty of Maritime, Akdeniz University, Antalya, TURKEY*

[e]*Faculty of Engineering, Akdeniz University, Antalya, TURKEY*

ARTICLE INFO

Research Article

Corresponding Author: Okan Oral, E-mail: okan@akdeniz.edu.tr

Received: 17 April 2023 / Revised: 25 July 2024 / Accepted: 31 July 2024 / Online: 14 January 2025

Cite this article

Özdemir N, Çakir M, Yilmaz M, Şimşek H, Oral M A, Oral O (2025). Application of the Machine Learning Methods to Assess the Impact of Physico-chemical Characteristics of Water on Feed Consumption in Fish Farms. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):71-79. DOI: 10.15832/ankutbd.1470111

ABSTRACT

Machine learning methods, which are one of the subfields of artificial intelligence and have gained popularity in applications in recent years, play an important role in solving many challenges in aquaculture. In this study, the relationship between changes in the physico-chemical characteristics of water and feed consumption was evaluated using machine learning methods. Eleven physico-chemical characteristics (temperature, pH, dissolved oxygen, electrical conductivity, salinity, nitrite nitrogen, nitrate nitrogen, ammonium nitrogen, total phosphorus, total suspended solids, and biological oxygen demand) of water were assessed. Among all the measured physico-chemical characteristics of water, temperature was determined to be the most important parameter to be evaluated in fish feeding. Moreover, pH2, EC2, TP2, TSS2, S2 and NO_2-N parameters detected in the outlet water are more important than those detected in the inlet water in terms of feed consumption. Through regression analysis carried out using machine learning methods, the models developed with Random Forest, Gradient Boosting Machine and eXtreme Gradient Boosting algorithms exhibited higher success rates in predicting feed consumption compared to the other models. The present study highlights the pivotal role of machine learning methods in enhancing our understanding of fish feeding dynamics based on physico-chemical characteristics of water, thus contributing significantly to aquaculture management practices.

Keywords: Aquaculture, Feed intake, Artificial intelligence, Rainbow trout, Sustainability

1. Introduction

Fish farms release various amounts of waste into the aquatic environment, which increases the necessity of examining and analysing the impact of aquaculture on the environment (Ahmad et al. 2022). The polluted environment primarily harms biodiversity, disrupts ecological balances and prevents sustainability by affecting production (Leaf & Weber 1998; Sharma & Birman 2024).

The total trout production of Turkey was 145649 tons in 2022, and Muğla province contributed to this production with 18.2% of the total amount (Çöteli 2023). The Eşen River is at the heart of intensive trout production in Muğla province (Sezgin et al. 2023; Koçer et al. 2010; Pulatsü & Yıldırım 2011).

For sustainability, it is essential to ensure that natural resources are used effectively and efficiently, in a balanced way, and in harmony with nature (Qin et al. 2024; Moldan et al. 2012). It is necessary to know the potential and structure of natural resources well and to observe the changes that occur in these resources. Research and monitoring activities in aquaculture are necessary to manage resources (Subasinghe et al. 2009; Mandal & Ghosh 2024).

New application areas and algorithms are constantly being developed with artificial intelligence (Kaya et al. 2023; Akgül et al. 2023; Kaya 2023). Many successful activities have been reported in fish farms, such as planning production with computer support, monitoring environmental conditions and fish health (Yilmaz et al. 2022; Yilmaz et al. 2023; Cakir et al. 2023), and aeration tools, growth statistics, intensive data analysis, production of feed (Dikel & Öz 2022). Feed expenses constitute the largest part of the production cost in aquaculture farms (Li et al. 2020). Of course, the balance established between the amount of harvested product and the amount of feed consumed represents successful production (Pahlow et al. 2015). However, the

concept of production solely for commercial concerns negatively affects sustainability (Folke & Kautsky 1992). Changes in water quality are the critical factor that directly affect fish production (Muir 2005). Evaluating the sensitive relationship between the physico-chemical characteristics of water and feed consumption with the help of computer-aided methods can help both optimize the amount of feed used and predict the amount of feed to be consumed on the farm (Zhao et al. 2021). Thus, while reducing production costs, predictable feed consumption will allow businesses to make more accurate future.

Machine learning (ML) methods, which are one of the subfields of artificial intelligence (AI) and have gained popularity in applications in recent years, play an important role in solving many challenges in aquaculture (Zhao et al. 2021).

In the current study, the relationship between changes in the physico-chemical characteristics of water and feed consumption was evaluated using ML methods. At the same time, it was aimed to determine the most successful ML method for predicting the effect of environmental conditions on rainbow trout production.

2. Material and Methods

2.1. Study area

The Eşen River (Figure 1) originates at approximately 2000 meters altitude in the southwest of Turkey. This river is used for different purposes such as drinking water supply, electricity generation, agricultural irrigation, and Rainbow Trout breeding (*Oncorhynchus mykiss* Walbaum 1792).

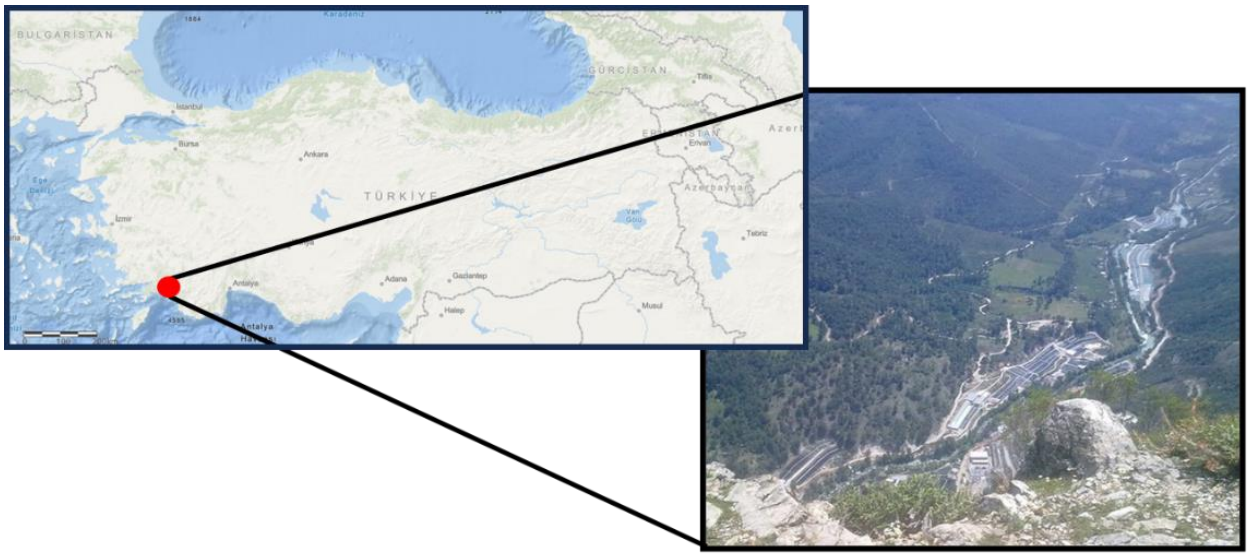


Figure 1- A view from Eşen River (Original)

Seven stations, including inlet and outlet waters of 3 farms and reference point, were monitored for 1 year. Moreover, feed consumption rates of farms were recorded. While determining the stations, care was taken to ensure that they were representative of the fish farms.

The Reference point is located 115 m upstream of the farms, which are located on Eşen River. This station, which is not exposed to pollutants, also serves as a reference for the rest of the river.

2.2. Water sampling and analysis

Water samples obtained from a total of 7 stations were filled into 2-liter polyethylene bottles and transported to Muğla Sıtkı Koçman University Laboratories in an ice-cooled box.

From the water samples taken from the stations, water temperature (T), pH, dissolved oxygen (DO), electrical conductivity (EC), and salinity (S) parameters were determined on-site in the field with a YSI multiparameter (MPS 556). Nitrite nitrogen (NO_2-N), nitrate nitrogen (NO_3-N), ammonium nitrogen (NH_4), total phosphorus (TP), total suspended solids (TSS) and biological oxygen demand (BOD) analyses were carried out in Muğla Sıtkı Koçman University Research Laboratories and Central Environmental Laboratory according to APHA (2012) methods.

2.3. ML methods

Regression methods are statistical techniques used to predict one variable based on one or more other variables. These methods model the relationship between data and make predictions using these models. There are many algorithms used in literature to obtain these predictions. However, some algorithms are more preferred in academic studies because they produce more successful predictions compared to others or because the models can be easily interpreted. These popular algorithms include Artificial Neural Network (ANN), Decision Tree (DT), Generalized Linear Model (GLM), Gradient Boosting Machine (GBM), K-Nearest Neighbour (KNN), Random Forest (RF), Support Vector Machine (SVM) and eXtreme Gradient Boosting (XGBoost). The success of these algorithms may vary according to the structure and complexity of the problem.

The regression models aim to estimate the relationship between at least one input variable and one output variable. Input variables can be defined as independent variables and output variables as dependent variables. The input variables used in this study are the physico-chemical characteristics of water, while the output variable is Monthly Feed Consumption per kg (MFC). Input variables are inlet and outlet water temperatures (T1, T2), pH values (pH1, pH2), dissolved oxygen amounts (DO1, DO2), electrical conductivities (EC1, EC2), salinity values (S1, S2), NH_4, NO_2 and NO_3 values (NH_41, NH_42, NO_21, NO_22, NO_31, NO_32), total suspended solids amount (TSS1, TSS2), Total phosphorus values (TP1, TP2) and biological oxygen needs (BOD1, BOD2). Since the stocking density for portion fish was the same in all farms included in the study (25 kg/m^3), stocking density was not included in the study as a variable.

In the regression analysis, ML techniques were used to model the relationship between the above-mentioned input variables and the output variable. The dataset contains 22 input variables and one output variable for a total of 180 observations. A significant portion of the dataset (80%) was randomly divided to be used for model training and the remaining portion (20%) for testing. The 10-fold cross-validation method was applied for the validation of the model to be created during the training phase. In addition, hyperparameter optimization of the regression algorithms was achieved to obtain the best model in a limited solution space.

To measure the performance of regression models, three different evaluation metrics are commonly used in studies, which are Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and R^2 respectively.

MAE is the average of the absolute values of the differences between the predicted values ($\hat{y}$) and the actual values (y) and is expressed as in Equation 1. N represents the number of observations (James et al. 2013).

$$MAE = (1/\mathrm{N}) * \sum|\mathrm{y} - \hat{\mathrm{y}}| \quad (1)$$

RMSE is another evaluation metric used to measure the performance of a regression model, and it is the square root of the mean square of the differences between the predicted values and the actual values (Géron 2019).

Mathematically, for a data set with N observations, the RMSE is calculated as shown in Equation 2.

$$RMSE = \sqrt{(1/\mathrm{N}) * \sum|\mathrm{y} - \hat{\mathrm{y}}|} \quad (2)$$

Since RMSE is calculated using the squares of the error quantities, large errors have a more significant impact, while small errors have a smaller impact. This makes RMSE a more sensitive evaluation metric than MAE, making it more suitable for datasets with outliers. In cases where outliers lead to large errors, the RMSE will be higher. However, RMSE is preferred because it is more sensitive than MAE, because of the use of the squares of the errors.

R^2 or coefficient of determination is an evaluation metric used to measure how well a regression model is fitted. This metric expresses how much of the variance of the actual data is explained by the model (Draper & Smith, 1998).

Mathematically, R^2 is calculated by the formula in Equation 3:

$$R^2 = 1 - (SSres / SStot) \quad (3)$$

Where; SSres represents the sum of the error squares, which measures the deviation of the model's predictions from the actual data, and SStot represents the total variance of the actual data.

The R^2 value ranges from 0 to 1, with a higher value indicating a better fit to the data. An R^2 value of 1 means that the model explains the actual data perfectly while a value of 0 means that the model does not explain the data at all. In a successful regression model, the MAE and RMSE error metrics are expected to be as small as possible, approaching zero, while the R^2 metric is expected to be close to 1.

3. Results and Discussions

3.1. Evaluation metrics of regression models

In this section, we present the results of the regression models where 4/5 of the dataset consisting of 180 observations was used for training the model, and 1/5 of the dataset was used for testing. Three different metrics were used to evaluate the ML regression models. Among these metrics, MAE and RMSE are error metrics, and these metric values can take values ranging from zero to infinity. In models with good performance, MAE and RMSE values are expected to be close to zero. These metrics are obtained both in the training phase and in the testing phase. Since 10-fold cross-validation is applied in the training phase, the error metrics consist of ten different values instead of a single value.

In the fish farms, where the study was conducted, a certain proportion of biomass is given depending on the water temperature and the amount of dissolved oxygen in line with the recommendations of the feed manufacturer. However, fish behavior during feeding is still evaluated by technical personnel. Feeding is stopped when a situation affecting feed intake (such as a decrease in the fish's feed intake or reluctance to feed) is observed. Therefore, the amounts of feed consumption in farms are affected by instantaneous changes. Possible reasons that may reduce the fish's demand for food cannot be monitored instantly. The obtained results can explain the possible reasons for the parameters affecting feed consumption. The order of importance of physico-chemical characteristics of water affecting feed consumption was evaluated with BORUTA analysis (Figure 2).

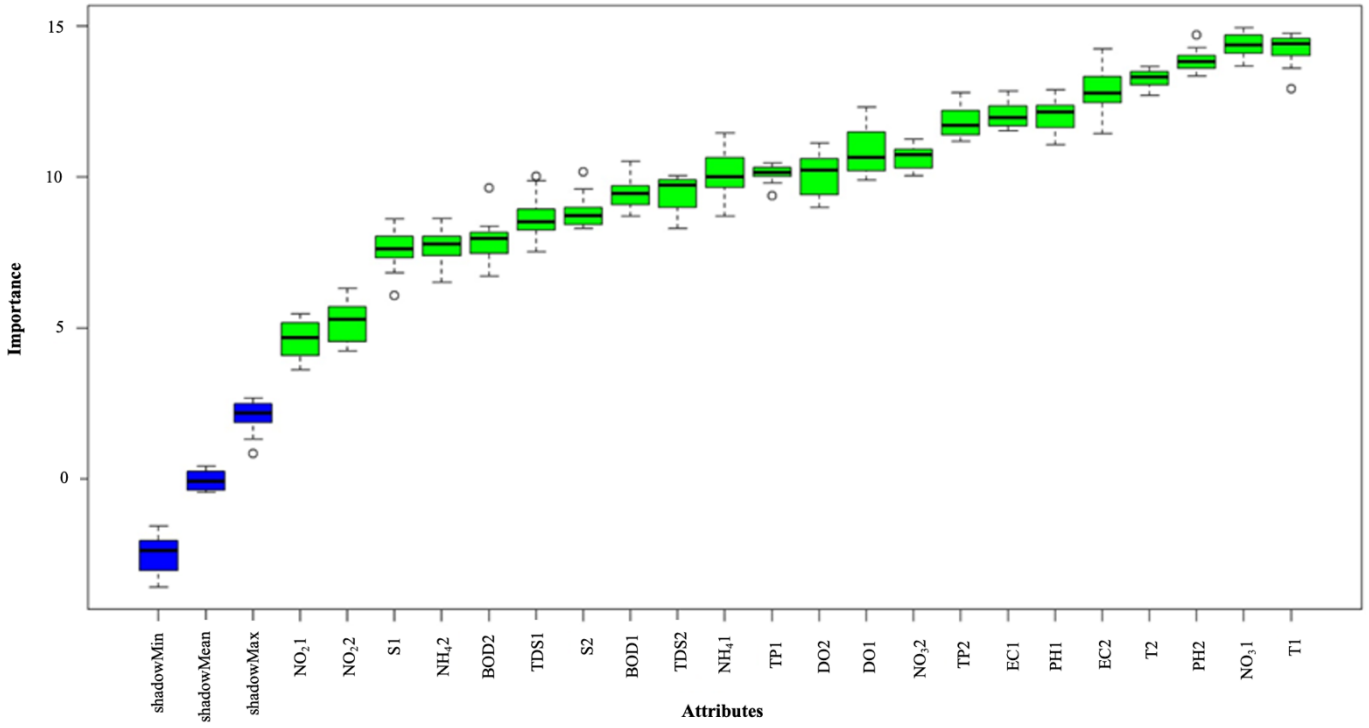


Figure 2- The order of importance of physico-chemical characteristics of water affecting feed consumption

These parameters of the wastes (caused by feed, fish and metabolic wastes) released into the water from the farm environment can be seen from the analysis results of the farm inlet and outlet waters that are directly affected by the physico-chemical characteristics of water (pH, TP, NH_4, BOD, NO_2-N, NO_3-N, EC, TSS, S, DO and T) evaluated in the current study in the farm environment (Table 1).

The pH2, EC2, TP2, TSS2, S2 and NO_2 parameters detected in the outlet water are more important than those detected in the inlet water in terms of feed consumption (Figure 2). Surprisingly, NO_3, pH, EC, TP2 are ranked as the most important and more important than DO, which is evaluated first (along with water temperature) in fish farms. In trout farms, the DO rate is desired to be above 6 mg/L for feed consumption (Pedersen 1987; McDaniel et al. 2005). In the farms where the study was carried out, DO mean values were measured in the range of 6.43-8.40 mg/L.

Among all the measured physico-chemical characteristics of water, temperature was determined to be the most important parameter to be evaluated in fish feeding (Figure 2). In trout farming, temperatures in the range of 1-25 $^{\circ}$C (with an optimal 16 $^{\circ}$C) are needed to grow fish (Woynarovich et al. 2011). In the current study, water temperature was measured between 11.30 and 14.88 $^{\circ}$C. This value is within the desired range for trout farming and is close to the optimum value (16 $^{\circ}$C).

Table 1- Measured physico-chemical characteristics of water (min-max / mean)

Physico-chemical Characteristics	***Reference***	***Farm A***		***Farm B***		***Farm C***	
		Inlet	*Outlet*	*Inlet*	*Outlet*	*Inlet*	*Outlet*
T (°C)	11.32-13.75	11.37-13.77	11.31-13.78	12.50-13.68	12.30-14.29	11.30-14.75	11.30-14.88
	13.04	13.09	13.05	13.08	13.27	13.12	13.12
pH	7.53-9.31	7.70-8.73	7.44-8.70	7.67-9.12	7.46-8.75	7.50-9.12	7.66-8.87
	8.15	8.14	8.01	8.15	7.97	8.08	8.05
DO (mgL^{-1})	7.43-10.51	7.3-10.43	5.49-9.00	5.08-10.3	4.97-9.21	5.42-10.26	5.1-9.00
	8.57	8.40	6.85	7.09	6.43	7.14	6.67
EC (μScm^{-1})	254.80-400.00	253.80-364.00	257.30-401.00	254.20-404.00	258.60-405.00	254.40-390.00	253.30-390.00
	336.46	337.70	344.55	337.45	348.40	344.18	338.97
S (‰)	0.160-0.220	0.160-0.230	0.160-0.230	0.160-0.220	0.160-0.230	0.160-0.220	0.16-0.220
	0.185	0.190	0.190	0.186	0.190	0.187	0.188
NH_4 (mgL^{-1})	BDL-0.10	BDL-0.13	BDL-0.54	BDL-0.39	0.23-2.92	BDL-0.53	BDL-0.82
	0.01	0.02	0.31	0.27	0.78	0.34	0.57
NO_2-N (mgL^{-1})	BDL-0.05	BDL-0.02	BDL-BDL	BDL-0.25	BDL-0.20	BDL-0.50	BDL-0.10
	0.01	0.01	BDL	0.07	0.06	0.07	0.03
NO_3-N (mgL^{-1})	BDL-10.37	0.91-9.59	1.19-6.56	1.30-17.51	1.20-28.39	1.16-20.56	1.04-22.59
	2.84	2.89	2.29	3.76	4.92	4.33	4.18
TP ($mg\ L^{-1}$)	BDL-0.285	BDL-0.005	BDL-0.046	BDL-0.043	0.013- 0.108	BDL-0.084	0.011- 0.914
	0.026	0.001	0.024	0.030	0.073	0.042	0.137
BOD_5 (mgL^{-1})	0.07-4.26	0.21-3.14	1.16-4.28	0.41-4.85	2.33-5.70	0.76-4.62	1.02-5.37
	1.94	2.16	2.7	2.71	3.97	2.77	3.21
TSS (mgL^{-1})	BDL-1.00	BDL-0.40	0.10-10.90	0.10-6.50	0.30-3.90	0.40-4.30	0.50-8.20
	0.23	0.11	1.91	1.65	2.23	1.58	2.53
Total Feed Consumption (ton)		649.5		1156.2		748.0	

*BDL: Below Detection Limit

According to the European Union directives, 2006/44/EC, (EU, 2006) on the quality of fresh waters that need to be protected or improved to support fish life, the pH, TP, NH_4 and BOD parameters of the water for salmonids should be 6-9, 0.2, ≤1 and ≤3 mg/L, respectively. During the study, pH was found to be 7.44-9.12, TP BDL was found to be 0.914, NH_4 BDL-2.92 and BOD were found to be between 0.21 and 5.70 in the farms. Although the measured TP, NH_4 and BOD values exceeded the values required for the growth of salmonids, pH2 measured within the recommended value range is the 3rd most important parameter affecting feed consumption among all the parameters. The fact that temperature and pH have an impact on all chemical reactions, metabolism, and toxicity (Jana & Sarkar 2005) may explain why they are at the top of the list of importance.

According to BORUTA analysis, although NO_2-N is more toxic in terms of feed consumption, it is ranked lower than NO_3-N in terms of importance. Inlet water NO_3-N value is listed as the most important parameter, right after temperature. NO_3 is the final product of the two-stage oxidation of ammonia. The intermediate product, NO_2, is oxidized with the help of bacteria to produce nitrate (Hargreaves 1998). Therefore, it can be said that the NO_2 concentration in the environment is kept at levels that will not affect feed consumption through nitrification.

High values of total dissolved solids, which describe inorganic salts and dissolved materials in water (Devi et al. 2017; Firooz et al. 2012), mean that they are unsuitable for fish health (Ahmed et al. 2019). It has been reported that total dissolved solids are among the top physico-chemical characteristics of water that cause fish disease outbreaks (Yılmaz et al. 2022). However, TSS, which was revealed in the current study to be an important parameter in terms of feed consumption, comes after other physico-chemical characteristics of water in importance. Similarly, S is one of the parameters that affects the fish's feed consumption less.

EC has greater importance on feed intake than TSS and S, comparable to pH and less important than temperature. In fact, it is affected by these 4 physico-chemical characteristics of water and the presence of inorganic dissolved solids such as ions carrying a negative charge (nitrate and phosphate anions) and ions carrying a positive charge (sodium, calcium, iron, etc.). With this feature, EC can be considered a critical indicator in evaluating feed consumption in fish farms where ML methods are used.

It should be noted that the data in this study are limited to the current measured values of physico-chemical characteristics of water measured in the farms where the study was carried out. In other words, it can be said that the order of importance of physico-chemical characteristics of water in terms of feed consumption may vary in each farm's own dynamics. It was revealed in the current study that a critical parameter such as DO may be placed behind other physico-chemical characteristics of water in terms of importance in farm environments where it remains at values that will not negatively affect fish feed consumption throughout the season (Figure 2). This shows that the farm environment should be evaluated with all its variables for effective production.

Table 2 shows the minimum, 1st quantile, median, mean, 3rd quantile, and maximum values of the MAE-based comparative error values for each regression model during the training process. As shown in the table, the DT algorithm has the highest average MAE error among the regression models. The regression models with the lowest mean MAE error values are XGBoost, GBM and RF. It can be observed that the minimum and maximum error ranges are narrower for the GLM, RF and GBM models, but wider for the ANN and DT models in terms of MAE error values.

Table 2- MAE-based error table for training data

	MAE					
	Min.	***1st Qu.***	***Median***	***Mean***	***3rd Qu.***	***Max.***
DT	0.0162611160	0.0218714429	0.024323865	0.025098375	0.028672426	0.034461268
ANN	0.0112890995	0.0123856034	0.013982100	0.016852160	0.017949030	0.037744668
SVM	0.0109737158	0.0118979884	0.012412511	0.014065447	0.016024698	0.019771007
GLM	0.0111900914	0.0121773948	0.013181225	0.013256358	0.014288051	0.015159422
KNN	0.0062285714	0.0086330769	0.012843077	0.011462479	0.014234856	0.015253333
RF	0.0039281359	0.0046053656	0.004942887	0.005486680	0.006135318	0.008105229
GBM	0.0032916726	0.0043113957	0.005101034	0.005223984	0.005706219	0.008745502
XGBoost	0.0002086571	0.0009246325	0.003543157	0.003952663	0.006798657	0.008891349

Table 3 shows the minimum, 1st quantile, median, mean, 3rd quantile and maximum values of the RMSE-based comparative error values of each regression model during the training process. As can be seen from the table, the regression model with the highest mean RMSE error is the DT algorithm. The regression models with the lowest RMSE mean error values are GBM, XGBoost and RF. Regarding RMSE error values, it is understood that the minimum and maximum error ranges are narrower for GLM, RF and GBM models but wider for ANN, DT and XGBoost.

Table 3- RMSE-based error table for training data

	RMSE					
	Min.	***1st Qu.***	***Median***	***Mean***	***3rd Qu.***	***Max.***
DT	0.021817266	0.025538575	0.029346811	0.030055109	0.032247176	0.04257985
ANN	0.014761440	0.015181156	0.018514751	0.020932702	0.021602019	0.04369090
SVM	0.014338495	0.015829987	0.017048988	0.018649991	0.021480378	0.02544829
GLM	0.013903197	0.015474243	0.016332362	0.016662840	0.017832495	0.01985353
KNN	0.008120521	0.011424263	0.016634153	0.015426489	0.018643893	0.02231042
RF	0.005425699	0.006055001	0.007568518	0.007929979	0.008630886	0.01313262
XGBoost	0.000285502	0.001776021	0.007803664	0.007383174	0.012750269	0.01427064
GBM	0.004772173	0.005297180	0.007186077	0.007276941	0.007980441	0.01348817

The minimum, 1st quantile, median, mean, 3rd quantile and maximum values of the R^2 based comparative model explanatory power values of each regression model during the training process are shown in Table 4. As can be seen from the Table 4, the regression models with the highest mean R^2 values are RF, GLM and XGBoost algorithms. The regression model with the lowest mean R^2 model explanatory power is DT. It is understood that the XGBoost model produced the maximum R^2 model explanatory power values.

Table 4- R^2-based performance table for training data

	R^2					
	Min.	***1st Qu.***	***Median***	***Mean***	***3rd Qu.***	***Max.***
RF	0.8925463	0.9705421	0.9745838	**0.9701575**	0.9862443	0.9956343
GLM	0.9231790	0.9671374	0.9761152	**0.9698487**	0.9793847	0.9914711
XGBoost	0.8680894	0.9167459	0.9710129	**0.9536456**	0.9981806	0.9999758
KNN	0.7948448	0.8324811	0.9086410	0.8809891	0.9142515	0.9495288
GLM	0.8147133	0.8313405	0.8613456	0.8624089	0.8809775	0.9267126
SVM	0.7267628	0.7996369	0.8334013	0.8262501	0.8627225	0.8870258
ANN	0.2662195	0.7661713	0.8500663	0.7779701	0.8693918	0.8865680
DT	0.1951602	0.4227606	0.5267219	0.5051544	0.6157219	0.7078576

The highest mean R^2 values were indicated in bold

After analysing the model evaluation metrics obtained for the training data, it is necessary to measure the performance of the regression models with test data that the regression models have never seen in the training phase. The results of the evaluation metrics of the performance test with the test data are given in Table 5.

Table 5- Model evaluation values with test data

		Evaluation Metrics		
		MAE	***RMSE***	***R^2***
Regression Models	**DT**	0.022643467	0.027030900	0.598961001
	ANN	0.016318823	0.019655126	0.787960261
	SVM	0.014680428	0.019378929	0.793877621
	GLM	0.014104027	0.017393552	0.833948736
	KNN	0.011044444	0.014378379	0.886528852
	RF	0.004423242	0.006448258	0.977178175
	XGBoost	**0.003074641**	0.006178471	0.979047897
	GBM	0.004402947	**0.005872574**	**0.981071218**

The lowest MAE, RMSE and the highest R^2 values were indicated in bold

Based on the evaluation metrics in Table 5, XGBoost has the lowest MAE error value, followed closely by GBM and RF. Therefore, XGBoost, GBM and RF are the most successful models according to the MAE metric.

According to the RMSE metric (Table 5), the lowest error value belongs to the GBM model with 0.005872574. Similarly, XGBoost and RF models are also close to the GBM model with the lowest error value.

When the results are analysed according to the R^2 model explanatory power metric (Table 5), the GBM model can explain the test data in the best way. Close R^2 values are also observed for XGBoost and RF models.

In the regression analysis carried out using ML techniques, the models developed with RF, GBM and XGBoost algorithms yielded better results for both the training and test datasets compared to models developed with other algorithms. Based on the MAE, RMSE and R^2 metrics used to evaluate the models, it was observed that RF, GBM and XGBoost algorithms generated similar results, although the order of performance varied, indicating that all three models are effective. Hence, the analysis of the eight different models highlighted that RF, GBM and XGBoost are the top three most effective algorithms. On the other hand, the DT algorithm produced the least effective results both in terms of model training and test performance.

In recent years, machine learning methods have been tried to be adapted to aquaculture (Zhao et al. 2021). Since feed cost is the main determinant of production cost (Li et al. 2020), many studies including computer-aided feeding systems (Hu et al. 2022) and the development of smart systems that enable optimization of feeding amount by monitoring fish behaviour (Ubina et al. 2021; Du et al. 2023; Zhou et al. 2019) have been reported. In the current study, RF, GBM, and XGBoost were determined to be the most effective algorithms that enable the estimation of the feeding amount by monitoring physico-chemical characteristics of water in fish farms. Comparable to the current study, it has been reported that methods such as convolution neural network (Ubina et al. 2021) and deep learning techniques using the image processing method (Hu et al.

2022; Zhou et al. 2019) can be successfully used to optimize feeding in aquaculture. In addition, studies investigating the effects of changes in water quality on the stock (Rana et al. 2021), fish length estimation (Li et al. 2020) and biomass detection (Yang et al. 2021) have been successfully carried out using ML methods.

4. Conclusions

In fish farming, feed expenses are the main factor determining the production cost. Feeds that cannot be consumed on farms are directly dumped into the water. These unused feeds deteriorate physico-chemical characteristics of water in the farm environment, reducing production performance and quality. In addition, these unconsumed feeds dissolve in water and provide an environment for the proliferation of undesirable pathogenic microorganisms. The emergence of diseases threatens fish health and causes great losses in production. Moreover, since unused feed are released to nature, they are loaded into the receiving environment as pollutants.

For aquaculture production to be sustainable, appropriate business policies must be developed for the use of natural resources and input costs must be kept at reasonable levels. The development of computer-aided smart production systems has the potential to prevent losses, protect the environment and support healthy and economical production. Sustainable aquaculture can be supported through effective risk management using smart systems.

References

Ahmad A L, Chin J Y, Harun M H Z M & Low S C (2022). Environmental impacts and imperative technologies towards sustainable treatment of aquaculture wastewater: A review. *Journal of Water Process Engineering* 46: 102553

Ahmed M, Rahaman M O, Rahman M & Kashem M A (2019). Analyzing the quality of water and predicting the suitability for fish farming based on iot in the context of bangladesh. In 2019 *International Conference on Sustainable Technologies for Industry* 4.0 (STI) (pp. 1–5). IEEE

Akgül İ, Kaya V & Zencir Tanır Ö (2023). A novel hybrid system for automatic detection of fish quality from eye and gill color characteristics using transfer learning technique. *Plos one* 18(4): e0284804

Baird R B, Eaton A D & Rice E W (2017) Standart Methods for The Examination of Water and Wastewater, 23st Edition, American Public Health Association, Washington.

Cakir M, Yilmaz M, Oral M A, Kazanci H Ö & Oral O (2023). Accuracy assessment of RFerns, NB, SVM, and kNN machine learning classifiers in aquaculture. *Journal of King Saud University-Science* 35(6): 102754

Çöteli F T (2023). Aquaculture Product Report (In Turkish). Tepge publication No: 373, Republic of Turkey Ministry of Agriculture and Forestry, Ankara.

Dikel S & Öz M (2022). Artificial intelligence (AI) application in aquaculture. In ISPEC 10th International Conference on Agriculture, Animal Sciences and Rural Development, July 18-19, Sivas, Turkey.

Devi P A, Padmavathy P, Aanand S & Aruljothi K (2017). Review on water quality parameters in freshwater cage fish culture. *International Journal of Applied Research* 3(5): 114–120

Draper N R & Smith H (1998). Applied regression analysis. John Wiley & Sons.

Du L, Lu Z & Li D (2023). A novel automatic detection method for breeding behavior of broodstock based on improved YOLOv5. *Computers and Electronics in Agriculture* 206: 107639

EU 2006. Directive 2006/44/EC of the European Parliament and of the Council of 6 September 2006 on the quality of fresh waters needing protection or improvement in order to support fish life

Firooz F, Mehdi R, Mostafa F, Alireza M & Gholamhossein N F (2012). Evaluation of physicochemical parameters of waste water from rainbow trout fish farms and their impacts on water quality of Koohrang stream –Iran. *International Journal of Fisheries and Aquaculture* 4(8): 170–177

Folke C & Kautsky N (1992). Aquaculture with its environment: prospects for sustainability. *Ocean & coastal management* 17(1): 5-24

Géron A (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media, Inc.

Hargreaves J A (1998). Nitrogen biogeochemistry of aquaculture ponds. *Aquaculture* 166(3-4): 181-212

Hu W C, Chen L B, Huang B K & Lin H M (2022). A computer vision-based intelligent fish feeding system using deep learning techniques for aquaculture. *IEEE Sensors Journal* 22(7): 7185-7194

James G, Witten D, Hastie T, Tibshirani R & Taylor J (2013). An Introduction to Statistical Learning: with Applications in R. Springer, Switzerland.

Jana B B & Sarkar D (2005). Water quality in aquaculture-Impact and management: A review. *The Indian Journal of Animal Sciences 75*(11)

Kaya V (2023). A Perspective on Transfer Learning in Computer Vision. In: Kılıc G B (Ed.). Advances in Engineering Sciences, Platanus Publishing, Ankara, Turkey

Kaya V, Akgül İ & Tanır Ö Z (2023). IsVoNet8: a proposed deep learning model for classification of some fish species. *Journal of Agricultural Sciences* 29(1): 298-307

Koçer M A, Muhammetoğlu A, Emre Y, Sevgili H, Türkgülü İ, Kanyılmaz M, Yılayaz A, Uysal R, Emre N, Mefut A, Uysal G, Yalım B & Topcuoğlu Ö A (2010). Determination of the Effects of Fish Farming and Basin-Related Pollution on the Eşen Stream Ecosystem by Using Mathematical Modeling Methods and Control of Nutrient Flux to the Mediterranean (in Turkish). (TÜBİTAK Project No: 107Y084).

Leaf A. & Weber P C (1998). Cardiovascular effects ofn- 3 fatty acids. *The New England Journal of Medicine* 318: 549-557

Li D, Wang Z, Wu S, Miao Z, Du L & Duan Y (2020). Automatic recognition methods of fish feeding behavior in aquaculture: A review. *Aquaculture* 528: 735508

Mandal A & Ghosh A R (2024). Role of artificial intelligence (AI) in fish growth and health status monitoring: A review on sustainable aquaculture. *Aquaculture International* 32(3): 2791-2820

McDaniel N K, Sugiura S H, Kehler T, Fletcher J W, Coloso R. M, Weis P & Ferraris R P (2005). Dissolved oxygen and dietary phosphorus modulate utilization and effluent partitioning of phosphorus in rainbow trout (Oncorhynchus mykiss) aquaculture. *Environmental Pollution* 138(2): 350-357

Moldan B, Janoušková S & Hák T (2012). How to understand and measure environmental sustainability: Indicators and targets. *Ecological indicators* 17: 4-13

Muir J (2005). Managing to harvest? Perspectives on the potential of aquaculture. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360(1453): 191-218

Pahlow M, Van Oel P R, Mekonnen M M & Hoekstra A Y (2015). Increasing pressure on freshwater resources due to terrestrial feed ingredients for aquaculture production. *Science of the Total Environment* 536: 847-857

Pedersen C L (1987). Energy budgets for juvenile rainbow trout at various oxygen concentrations. *Aquaculture* 62(3-4): 289-298

Pulatsü S & Yıldırım H B (2011) Evaluation of Outlet Water Characteristics of Land-Based Trout Farms of Different Capacities (Mugla, Fethiye) (in Turkish). Ankara University Scientific Research Project Final Report, Project no: 09B4347009

Qin C, Xue Q, Zhang J, Lu L, Xiong S, Xiao Y & Wang J (2024). A beautiful China initiative towards the harmony between humanity and the nature. *Frontiers of Environmental Science & Engineering* 18(6): 1-9

Rana M, Rahman A, Dabrowski J, Arnold S, McCulloch J & Pais B (2021). Machine learning approach to investigate the influence of water quality on aquatic livestock in freshwater ponds. *Biosystems Engineering* 208: 164-175

Sezgin S S, Yilmaz M, Arslan T & Kubilay A (2023). Current antibiotic sensitivity of Lactococcus garvieae in rainbow trout (Oncorhynchus mykiss) farms from Southwestern Turkey. *Journal of Agricultural Sciences* 29(2): 630-642

Sharma I & Birman S (2024). Biodiversity Loss, Ecosystem Services, and Their Role in Promoting Sustainable Health. In The Climate-Health-Sustainability Nexus: Understanding the Interconnected Impact on Populations and the Environment (pp. 163-188). Cham: Springer Nature, Switzerland.

Subasinghe R, Soto D & Jia J (2009). Global aquaculture and its role in sustainable development. *Reviews in aquaculture* 1(1): 2-9

Ubina N, Cheng S C, Chang C C & Chen H Y (2021). Evaluating fish feeding intensity in aquaculture with convolutional neural networks. *Aquacultural Engineering* 94: 102178

Woynarovich A, Hoitsy G & Moth-Poulsen T (2011). Small-scale rainbow trout farming. FAO fisheries and aquaculture technical paper, (561). Food and agriculture organization of the united nations, Rome

Yang L, Liu Y, Yu H, Fang X, Song L, Li D & Chen Y (2021). Computer vision models in intelligent aquaculture with emphasis on fish detection and behavior analysis: A review. *Archives of Computational Methods in Engineering* 28(4): 2785–2816

Yilmaz M, Çakir M, Oral M A, Kazanci H Ö & Oral O (2023). Evaluation of disease outbreak in terms of physico-chemical characteristics and heavy metal load of water in a fish farm with machine learning techniques. *Saudi Journal of Biological Sciences* 30(4): 103625.

Yilmaz M, Çakir M, Oral O, Oral M A & Arslan T (2022). Using machine learning technique for disease outbreak prediction in rainbow trout (Oncorhynchus mykiss) farms. *Aquaculture Research* 53(18): 6721-6732

Zhao S, Zhang S, Liu J, Wang H, Zhu J, Li D & Zhao R (2021). Application of machine learning in intelligent fish aquaculture: A review. *Aquaculture* 540: 736724

Zhou C, Xu D, Chen L, Zhang S, Sun C, Yang X & Wang Y (2019). Evaluation of fish feeding intensity in aquaculture using a convolutional neural network and machine vision. *Aquaculture* 507: 457-465

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

The Effect of UV-A / UV-B Radiation on Quality Changes of Harvested Curly Lettuce During the Storage

Öznur Cumhur Değirmenci[a*], Alev Akpınar Borazan[a], Emre Devlez[b]

[a]*Biotechnology Application and Research Center, Bilecik Seyh Edebali University, Bilecik, TURKEY*

[b]*Department of Biotechnology, Institute of Graduate Studies, Bilecik Seyh Edebali University, Bilecik, TURKEY*

ARTICLE INFO

Research Article

Corresponding Author: Öznur Cumhur Değirmenci, E-mail: oznur.cumhur@bilecik.edu.tr

Received: 11 May 2024 / Revised: 22 July 2024 / Accepted: 05 August 2024 / Online: 14 January 2025

Cite this article

Değirmenci Ö C, Borazan AA, Devlez E (2025). The Effect of UV-A / UV-B Radiation on Quality Changes of Harvested Curly Lettuce During the Storage. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):80-90. DOI: 10.15832/ankutbd.1482282

ABSTRACT

This study investigated the effects of UV-A and UV-B radiation on curly lettuce quality. Results focused on colour, total phenolic content, antioxidant activity, and ascorbic acid. The findings revealed that the highest phenolic content (46.1 mg GAE/100 g FL) had been observed in lettuce samples treated with high dose UV-B on the 7th day. The lowest phenolic content (13.7 mg GAE/100 g FL) was recorded in those treated with low dose UV-B on the same day of storage. Data showed an increase of 29.7% in antioxidant activity and 53.7% in total phenolic content after 7 days of storage in samples treated with high dose UV-B. High dose UV-A radiation was found to be the most effective in maintaining and enhancing the ascorbic acid content of the lettuce. UV applications did not cause yellowing in the stored lettuce leaves. Further research on different doses and optimization is recommended.

Keywords: Ultraviolet A radiation, Ultraviolet B radiation, Lettuce, Quality characteristics

1. Introduction

The demand for freshly cut or minimally processed food products has increased recently. Processed products have become an essential part of the food market due to changing lifestyles and developing technology. Salads, widely consumed and generally defined as healthy, attract attention for this reason. Lettuce is one of the vegetables frequently used in ready-made salads, either alone or in combination. A leafy lettuce species known to originate from Turkey and the Middle East belongs to *Asteraceae* family (Das & Bhattacharjee 2020). Lettuce, which has health-promoting properties, contains certain bioactive components. These include ascorbic acid, carotenoids, phenolic compounds, and fibre. Lettuce also shows a significant oxygen radical scavenging capacity (Kim et al. 2016; Liu et al. 2007). Lettuce suffers a nutrient loss during harvest to the final consumer and deteriorates rapidly. Therefore, to protect human health, it is imperial to search for effective methods to enhance the biosynthesis of components with significant impact on health. Many researchers have investigated the phenolic structures and antioxidant properties of lettuce under environmental growing conditions (Llorach et al. 2004; Muscolo et al. 2022; Pérez-López et al. 2013) and storage conditions (Collado-González et al. 2022; Wang et al. 2022). Results for lettuce varied depending on the species and conditions. Significant product and nutrient losses occur in fruits and vegetables due to rapid aging and diseases in post-harvest storage. Post-harvest preservation has traditionally relied on refrigeration and chemical preservation technology (Usall et al. 2016). Food science and food industry developments led to new processing technologies enabling practical access to fresh and healthy foodstuffs in all seasons. Many new and developing technologies include aseptic packaging, controlled atmospheric packaging under vacuum, ozonation and recently UV radiation applications (Koutchma 2019). Since lettuce and other leafy greens continue to breathe after harvest, temperature, and humidity control are very important in storage. One of the most important functions of refrigeration is to control the respiration rate of vegetables and fruits. Uncontrolled respiration leads to the decomposition of sugar, fat, and proteins in the product cells and heat release. Loss of this type of stored food through respiration means reduced food nutritional value, loss of flavor, reduced salable weight, and faster spoilage. A food's respiration rate largely determines its transit and post-harvest life. The higher the storage temperature, the higher the respiration rate will be. For refrigeration to be effective in delaying spoilage, keeping the temperature in cold storage as constant as possible is important. Additionally, exposure to fluctuating temperatures can cause moisture accumulation (sweating) on the product surface, accelerating rotting. Storage rooms should be well insulated and adequately cooled, allowing air circulation to prevent temperature change. The best results can be achieved with vacuum or hydro-vac cooling for leafy products with a high surface/volume ratio, such as lettuce. The biggest disadvantage of these applications is the system cost. However, when choosing a current trend alternative, product qualities can be preserved in addition to the installation cost and practicality of UV-A and B

applications (Cefola & Pace 2023; Janghu et al. 2024). In UV applications, when factors such as radiation dosage, application surface and surface area, and application distance are selected per the food, it is possible to ensure the desired level of process efficiency. Otherwise, changes in genomics, lipid destruction, oxidative damage, changes in plant biochemistry and reduced growth, food toxicity and the risk of consumer illness due to the negative effects of radiation are inevitable (Guerrero-Beltr·n et al. 2004; Csapó et al.2019). In addition to numerous UV damage to plants, this radiation can sometimes induce the plants to synthesize useful compounds for humans (Jansen et al. 2008). UV radiation technology reveals two different beneficial effects on fruits and vegetables. The first effect is the inactivation of microorganisms in food products, which leads to elevated shelf life (Koutchma 2019). For this purpose, UV-C is frequently used to produce safe food. As a second beneficial effect, UV radiation technology produces some desirable results in improving the defenses of food products against microorganisms, increasing the content of ingredients with beneficial effects for health, extended shelf life, preservation, and improvement of sensory properties (Koutchma 2019). Azarafshan and colleagues' study (2020) showed that plants' secondary metabolites increased with increasing UV-B radiation. The research indicated that plants produce greater amounts of total phenols, flavonoids, and anthocyanin to counteract the effects of radiation, especially at high radiation intensities. Some pathways containing phenol, flavonoid, and anthocyanin increase activity in response to UV-B rays. For example, many of the enzymes in the phenylpropanoid pathway are activated by UV-B. The buildup of these compounds in the epidermal cell vacuoles blocks UV-B rays from penetrating into vulnerable leaf regions, particularly photosynthetic tissues. Additionally, due to their redox properties, these compounds could act as antioxidants, scavenging single oxygen and effectively inhibiting the oxidation of macromolecules like lipids, proteins, and nucleic acids to minimize plant damage (Azarafshan et al. 2020).

Therefore, UV radiation applications attract the attention of numerous researchers in terms of reducing product losses of fruits, vegetables, protecting, and improving bioactive plant components, nutritional values, and quality characteristics. Lettuce is among the examples, the nutritional and health properties of which could be increased based on the UV radiation source. UV technology application in lettuce mainly focused on post-harvest effects. To the best of our knowledge, there is only one study evaluating the possibility of applying UV-B radiation in packaging lines for fruit, vegetable, and root crops, including freshly cut "Iceberg" lettuce (Du et al. 2014). The aim of this study is to investigate the effects of varying doses of UV-A and UV-B radiation on colour change, total phenolic content, antioxidant activity, and ascorbic acid profile of curly lettuce, a lettuce species widely cultivated worldwide.

2. Methods and Material

2.1. Chemicals and reagents

L(+)-ascorbic acid, oxalic acid dihydrate, Folin–Ciocalteu reagent (FCR), 2,6-dichlorophenol-indophenol sodium salt hydrate, gallic acid, sodium carbonate (anhydrous Na_2CO_3), methanol were purchased from Merck (Darmstadt, Germany). Sodium bicarbonate puriss and 2,2-diphenyl-1-picrylhydrazil (DPPH) and methanol were obtained from Sigma-Aldrich (Darmstadt, Germany).

2.2. Selection of plant materials and UV treatment procedures

Curly lettuce (*Lactuca sativa* var. *crista*) was obtained from a lettuce grower in İnegöl/Bursa, Turkey in June 2021. Locarno / curly lettuce has green, curly leaves on the outside and a crisp white heart. The leaves have fairly mild flavors and a crispy texture. Curly lettuce was mature in full sun and partial shade and ready to harvest in 70 to 80 days. Lettuce harvested early in the morning was brought to the laboratory in a till basket within 2 hours immediately after harvest. The lettuce was stored at 10 ± 2 °C with air conditioning under laboratory conditions until UV treatment. The dented, rotten, and yellowed parts of the lettuce were removed before application. No pre-treatment was applied to the lettuce before UV employment. UV radiation was carried out on harvest day and under laboratory ambient conditions.

Curly lettuce was exposed to UV-A or UV-B radiation in an open front and back radiation chamber, coated with aluminium foil. The design of the chamber was modified based on recent publications (Allende et al. 2006; Salemi et al. 2021). Experiments were performed in a wooden UV radiation chamber (width: 70 mm, depth: 65 mm, height: 100 mm) equipped with two parallel UV-A (PL-L 36W/10/4P, Poland) or UV-B (PL-L 36W/01/4P, Poland) lamps. The distance of the lettuce from the lamps was 65 cm (Azarafshan et al. 2020). UV radiation intensity measurements were recorded via radiometer (Model PCEUV34, PCE Instruments UK Ltd., Deutschland) to determine the UV-A and UV-B intensities at 18 different locations in the application tray without blocking the light in the cabinet. The UV treatment conditions applied to the lettuce samples were varied by varying the exposure time. The UV lamps were turned on before use for 15 minutes to enable stabilization. The conditions applied during UV treatment and the experimental steps are illustrated in Table 1 and Figure 1.

Table 1- UV application procedure to lettuce leaves

UV Lamp	*Exposure Dose Stage*	*Exposure Dose (kJ m^{-2})*	*Exposure Time (min)*
UV-B	Low dose	1.52	4
UV-B	Medium dose	3.04	8
UV-B	High dose	4.56	12
UV-A	Low dose	1.52	4
UV-A	Medium dose	3.03	8
UV-A	High dose	4.55	12
Control	The samples were treated in the same manner without switching the UV lamps on.		

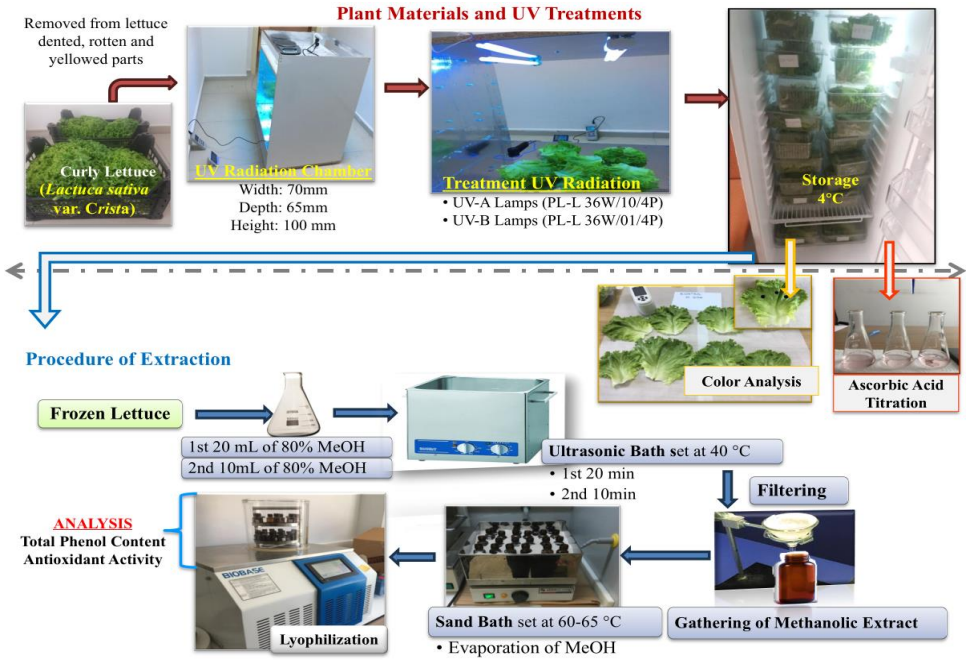


Figure 1- Experimental flow chart.

According to Du et al. (2014), lettuce samples deteriorated on the 10th day after UV application. In the literature survey and considering the fact that lettuce was a perishable product, samples were stored at 4 °C for a maximum of 7 days after UV application. After UV treatment, lettuce leaves (130-150 g) were placed in polyethylene terephthalate containers containing ventilation holes in the surface cover and stored in the refrigerator at 4±1 °C for 1, 3, or 7 days. Ascorbic acid determination and colour analysis were performed on the 1st, 3rd, and 7th days. The remaining lettuce leaf samples were frozen and stored at −82 °C until further analysis.

2.3. Colour analyses

The colour of treated and stored lettuces was measured with the Chroma Meter NR-200 (Shenzhen 3NH Technology Co., Ltd., China). Readings were conducted from three different points of lettuce leaves. This procedure was repeated at least three times. The differences with untreated control samples were calculated in terms of ΔL^*, Δa^*, and Δb^* in the lettuce samples exposed to different doses of UV radiation and the storage period. The total colour difference (AE) for each sample was calculated using the following equation (Pathare et al. 2013):

$$\Delta E = \sqrt{[(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2]}$$

The L^* coordinates represent the lightness-to-darkness scale, ranging from white to black. The a^* and b^* coordinates represent the colour's chromaticity, with the a^* axis varying from green to red and the b^* axis varying from blue to yellow. $\Delta E_{process}$ and $\Delta E_{storage}$ were the two types of colour differences calculated in analyses. In $\Delta E_{process}$, the colour difference values were calculated by comparing the control samples among themselves on the same storage day. On the other hand, $\Delta E_{storage}$ was determined by measuring the colour difference values compared to 0th-day samples within the same UV treatment groups.

2.4. Ascorbic acid determination

The amount of ascorbic acid in curly lettuce was determined via the titrimetric method utilizing 2,6-dichlorophenolindophenol solution (Nadkarni 1965).

2.5. Extraction method

For the determination of antioxidant activity and total phenolic content, 5-6 g of each frozen lettuce sample were extracted with 20 mL 80% aqueous methanol for 20 minutes at 40 °C in an ultrasonic bath and filtered using Whatman No. 1 filter paper. This process was carried out in two stages, and the filtrates were combined in an amber tube. The methanol fraction of the mixture was evaporated in a sand bath (60-65 °C) and then the remaining aqueous extracts in amber bottles were dried via Lyophilization (Biobase, China) at -56 °C under mbar vacuum pressure for 24 h (Llorach et al. 2004). Lyophilized extracts of the lettuce samples in amber bottles were kept in the refrigerator until analysis.

2.6. Total phenolic content

The total phenolic content was determined by the Folin-Ciocalteu method (Tomás-Callejas et al. 2012). 30-35 mg samples were freeze-dried and then homogenized in 5 mL of 80% methanol solution via ultrasonic bath. The lettuce extracts (200 µl) were mixed with 1.5 mL of Folin-Ciocalteu reagent (previously diluted 10-fold with distilled water) and vortexed. 1.2 mL of 7.5% (w/v) Na_2CO_3 solution was added to this solution. The resulting mixture was kept in the dark for 90 min at room temperature. The absorbance was measured at 765 nm with a UV/VIS spectrophotometer (AgileSpec). A standard curve of gallic acid solution (0.005-0.15 mg mL^{-1}) was prepared using a similar procedure and quantification was done concerning this standard curve (R^2= 0.9986). The final quantitative results were expressed as mg gallic acid equivalent (GAE)/g of fresh lettuce (FL).

2.7. Total *antioxidant activity*

The DPPH radical scavenging assay was conducted following the method of Kang & Saltveit (2002) with some modifications. The ability to scavenge the DPPH radical was the main factor in measuring the antioxidant activity of the extract. 1 mL sample from 0.07 mg lettuce mL^{-1} solution containing 80% aqueous MeOH and 2 mL of DPPH solution was added and vortexed. The absorbances of the colours formed as a result of the reaction in the solutions. The aliquots were kept in the dark for 30 minutes at room temperature, absorbances were recorded at 517 nm in a UV/VIS spectrophotometer (AgileSpec) with MeOH used as the blank. The DPPH concentration in the lettuce sample medium was calculated using the calibration curve (R^2= 0.9972) prepared using the DPPH standard (0.008 and 0.04 mg mL^{-1}). The free radical scavenging effects of the lettuce extracts were given as the concentration (EC_{50} µg mL^{-1}) at which 50% of the DPPH was scavenged.

2.8. Statistical analysis

The data were expressed as means ± SD and the statistical evaluation was performed using SPSS 22.0 (SPSS Inc., USA). A comparison was made between UV treatment groups or storage using the one-way analysis of variance (ANOVA). The significance of differences was conducted with the Tukey Honestly Significance Difference test ($P<0.05$).

3. Results and Discussion

3.1. Effects of UV radiation on colour change of curly lettuce

The colour of fruits and vegetables is one of the crucial factors used in evaluating the quality of products. Colour is also the main factor affecting the preferences of consumers. Technically, instrumental colour analysis is determined by spectrophotometric approaches, in which the observed colour can be expressed numerically. The effect of UV treatments and storage on L*, a* and b* values in lettuce leaves were illustrated in Figure 2. The changes of L* and a* values in lettuce leaves treated with UV were statistically not significant ($P>0.05$). However, storage was found to have an altering effect on the total colour change of treated samples. The initial L* value of control samples was 52.1. It was determined that the brightness (L* value) of lettuce samples treated with high dose UV-A and medium dose UV-B had increased on the 3rd and 7th days of storage (Figure 2c).

The a* value of control samples was -9.1 at the beginning of storage. The only statistically significant increase (3rd day: -9.4; 7th day: -10.1) in the green colour of lettuce (a* value) leaves was in the case of high dose UV-A application during storage (Figure 2a). UV radiation and storage did not cause any yellowing change (b value) in the lettuce samples (Figure 2b). Literature survey revealed various result regarding the colour change upon UV- treatment. Kasim & Kasim (2017b) reported the yellowing of spinach. Similar yellowing results were obtained in the work of Aiamla-or et al. (2010) with broccoli. On the other hand, the effect of UV-B application seemed to be variant. Kasim & Kasim (2017a) reported an increase of the red colour of capia peppers. Consequently, the results obtained for UV-B radiation seemed to be independent of the type of vegetable investigated in the studies.

On the other hand, further literature investigation revealed results like those obtained in the present study. UV employment preserves the colour of green-coloured (such as lettuce) fruits and vegetables due to the suppression of chlorophyll degradation, the most common abiotic stress observed by UV treatment (Aiamla-or et al. 2010). This was the case in the present study and similar results were obtained in the work of Aiamla-or et al. (2009). They reported that UV-B application had delayed the yellowing of broccoli flowers, while yellowing had been inevitable in the presence of UV-A application.

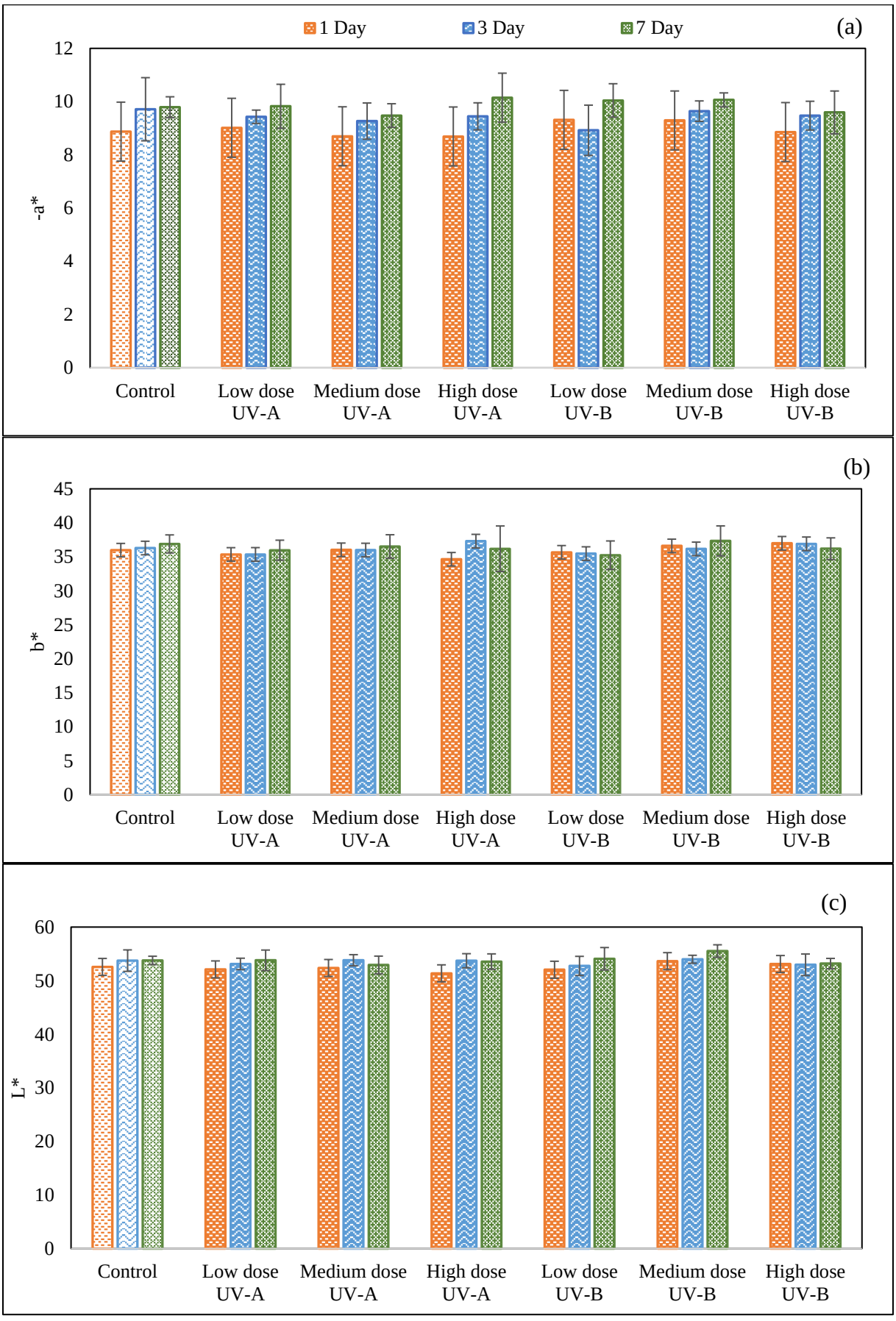


Figure 2- Effect of UV-A or UV-B treatment on lettuce during storage: Changes in (a) red to green colour, a*; (b) yellow to blue colour, b* and (c) lightness to darkness, L* values.

The ΔE values given in Tables 2 and 3 represented the colour changes relative to the initial values. Following UV radiation exposure, the ΔE values in lettuce leaves ranged from 1.17 to 3.00 (Table 2). Throughout the storage period, colour changes were determined within the range of 1.33 to 4.11 (Table 3). However, these colour changes were not statistically significant (P>0.05), indicating that the UV treatments had not been effective in inducing noticeable alterations in colour. Figure 3 illustrates the lettuce samples after 7 days of UV-A and UV-B rays exposure.

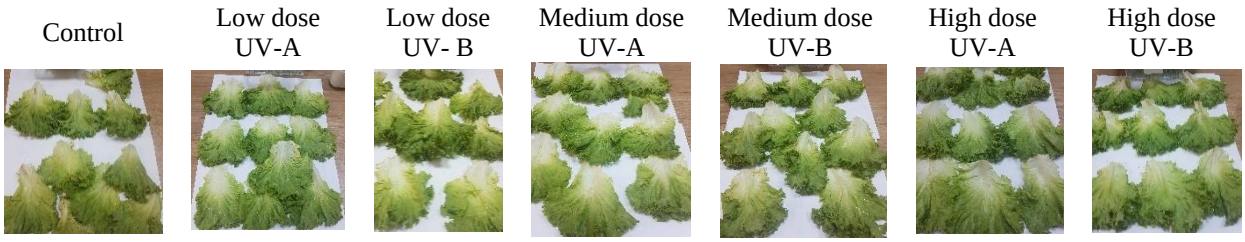


Figure 3- Effects of different doses UV-A and UV-B on the colour of lettuce at the end of the 7th.

Table 2- Effect of UV treatments on the total colour difference of lettuce, ($\Delta E_{process}$).

$\Delta E_{process}$	*Low dose*		*Medium dose*		*High dose*	
	UV-A	*UV-B*	*UV-A*	*UV-B*	*UV-A*	*UV-B*
1st Day	2.14 ± 0.46	2.53 ± 1.46	1.17 ± 0.63	2.16 ± 0.96	3.00 ± 1.32	2.37 ± 1.54
3rd Day	2.54 ± 0.62	2.93 ± 1.27	1.80 ± 1.54	2.16 ± 0.91	2.61 ± 2.7	1.82 ± 1.19
7th Day	2.03 ± 0.98	2.72 ± 1.07	2.22 ± 1.13	2.55 ± 1.65	1.75 ± 0.62	2.36 ± 1.16

Table 3- Effect of storage on total colour difference of applied lettuce, ($\Delta E_{storage}$).

Applied dose	*UV lamp*	$\Delta E_{storage}$		
		1st Day	*3rd Day*	*7th Day*
Low	UV-A	2.00 ± 0.51	2.14 ± 1.54	2.60 ± 0.92
	UV-B	2.30 ± 1.69	2.61 ± 1.25	2.88 ± 1.53
Medium	UV-A	1.33 ± 0.57	2.29 ± 1.79	2.13 ± 1.56
	UV-B	2.43 ± 1.12	2.73 ± 1.48	4.11 ± 2.15
High	UV-A	2.84 ± 1.05	3.30 ± 2.75	2.28 ± 1.06
	UV-B	2.66 ± 1.59	2.19 ± 1.15	2.47 ± 1.29
Control		2.65 ± 0.86	2.14 ± 1.05	2.26 ± 1.25

3.2. Effects UV radiation on ascorbic acid in curly lettuce

Lettuce is an important source of ascorbic acid, and the increased radiation intensity during lettuce production accelerates the synthesis of vitamin C by promoting the activity of enzymes involved in vitamin C metabolism in plant leaves (Martínez-ispizua et al. 2022). Ascorbic acid amounts of lettuce samples are given in Figure 4. Results indicated that UV-A and UV-B application doses and storage had affected the amount of ascorbic acid (P<0.05). The average initial amount of ascorbic acid in the control samples was 32.5 mg/100 g FL. The amount of ascorbic acid in untreated lettuce leaves decreased to 31.2 mg/100 g FL after 7 days of storage at 4 °C. Post-harvest storage causes deterioration of ascorbic acid in lettuce samples over time. However, the vitamin C content increased on the 7th day of storage in all lettuce samples with UV treatment, except for the moderate UV-B application.

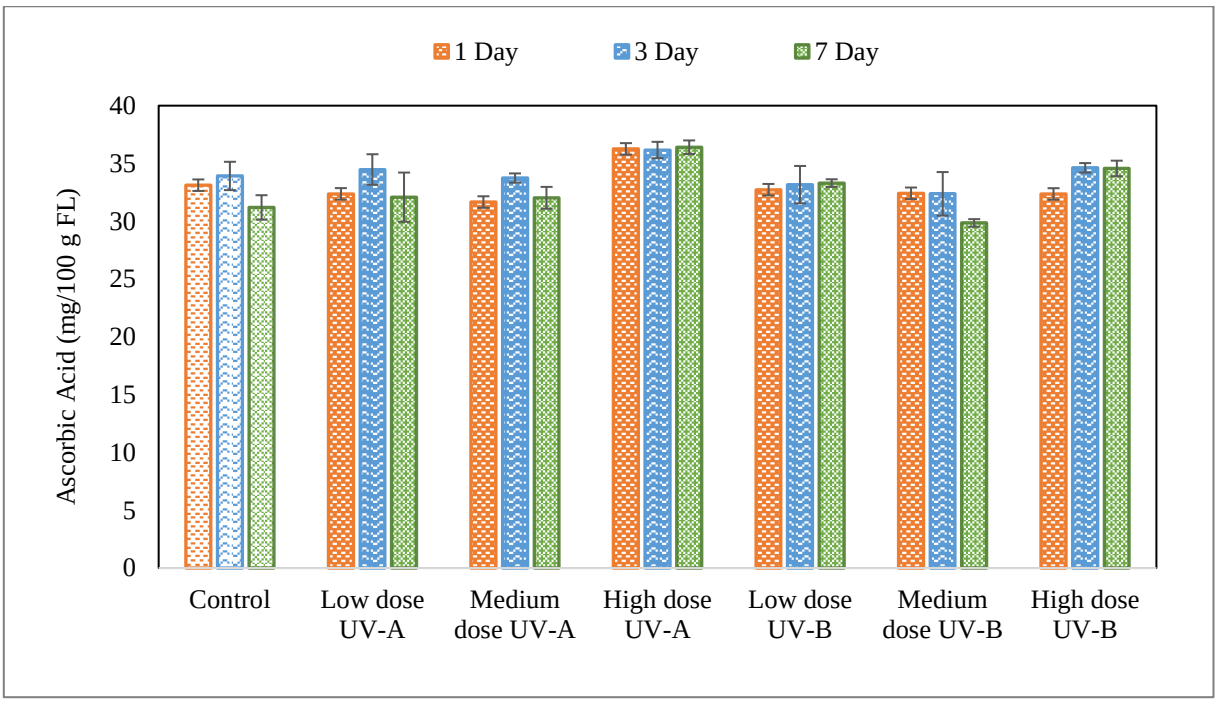


Figure 4- Effect of UV-A or UV-B treatment on total ascorbic acid content of lettuce during storage at 4 °C. Ascorbic acid content of control samples on day 0: 32.5 ±0.5 mg/100 g FL.

The amount of ascorbic acid in lettuce varies according to species, pre-harvest and post-harvest factors (Nicolle et al. 2004). Mampholo et al. (2016) reported that the ascorbic acid content in curly lettuce species had varied between 13.39 and 26.5 mg/100 g FL. The amount of ascorbic acid was higher than the control samples of the present study. Ascorbic acid has high UV absorption in the 200-280 nm germicide wavelength range. Photochemical reactions at these wavelengths may lead to photodegradation (Koutchma 2009; Tikekar et al. 2011). In this study, when the effects of UV-A and UV-B applications were evaluated, high rates of losses in ascorbic acid amounts should not be expected since ascorbic acid does not absorb light significantly at wavelengths above 300 nm (Koutchma 2009).

The amount of ascorbic acid in the UV-A treated lettuces was significantly higher than the control samples at the end of 7 storage days. The amount of ascorbic acid in lettuce leaves exposed to high dose UV-A was the highest at 36.4 mg/100 g FL, indicating a 12% increase. The amount of ascorbic acid was higher in the lettuce samples treated with low and high UV-B levels than the control samples on the 7th day of storage. Similarly, it was determined that the ascorbic acid content increased in products such as apples (Hagen et al. 2007) and tomatoes (Castagna et al. 2013) after UV-B applications. In contrast, Liu et al. (2011) reported that UV-B radiation negatively affected the ascorbic acid content in green ripe tomatoes after harvest. Literature results reveal that each study should be evaluated on a product basis.

3.3. Effects of UV radiation on total phenolic content and antioxidant activity in curly lettuce

Lettuce is a vegetable that contains significant antioxidant activity and phenolic content. The main compounds are phenolic acids found especially in green leafy lettuce (Das & Bhattacharjee 2020). Short-term exposure to environmental factors, agricultural practices, and stress factors could also be employed to improve important nutritional values and antioxidant capacity. On the other hand, the total phenolic content and antioxidant activities vary depending on the species and colour of lettuce (Martínez-Ispizua et al. 2022; Nicolle et al. 2004). The change of total phenolic content of the lettuce expressed as mg GAE with UV-A and UV-B dose and storage was illustrated in Figure 5a. Control lettuce samples had a mean total phenolic content of 14.2 ± 0.7 mg GAE/100 g FL. The total amount of phenolic content in control samples reached the highest level (41.8 ± 1.4 mg GAE/100 g FL) on the 1st day of storage. At the end of the storage, the total phenolic content ranged between 27.0 and 33.0 mg GAE/ 100 g FL.

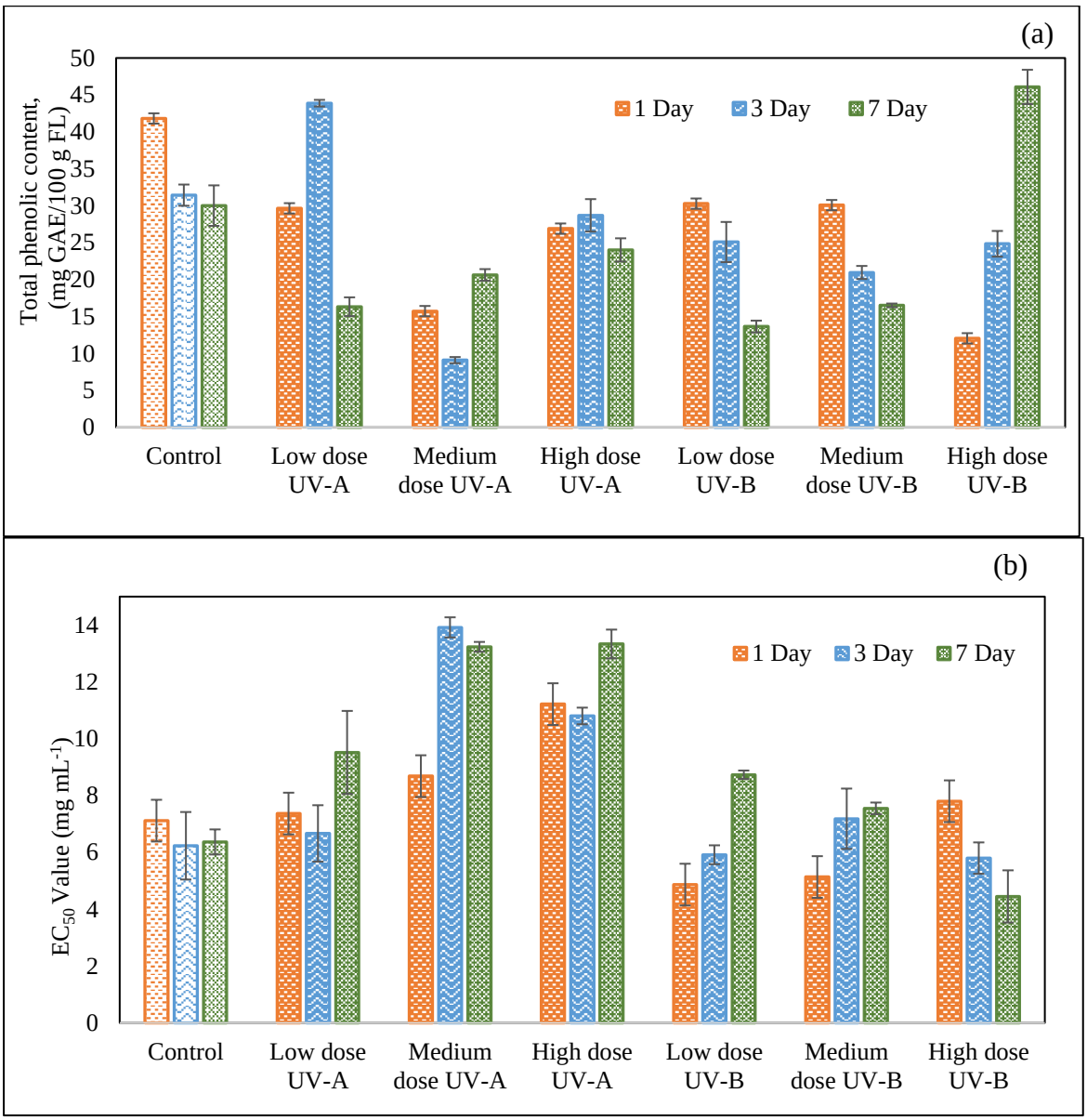


Figure 5- Effect of UV-A or UV-B treatment on lettuce during storage at 4 °C. Changes in (a) total phenolic content of lettuce, (b) EC_{50} value of lettuce. Phenolic content of the control samples on day 0: 14.2 ±0.7 mg GAE/100 g FL. EC_{50} value of the control samples on day 0: 12.2 ±0.7 mg mL^{-1}.

The total phenolic amount was statistically affected by post-harvest UV-A and UV-B applications. The total phenolic content in fruits and vegetables should increase depending on the light-induced stress conditions. The effect of UV radiation on phenolic content in fruits and vegetables depends on the product type, UV type, radiation dose, incubation time, and environmental factors (Du et al. 2014; Nguyen et al. 2014; Scattino et al. 2016). The responses obtained in the biochemical properties of the products might have varied positively or negatively depending on UV radiation, which was an abiotic stress factor (Jansen et al. 2008). Therefore, it was important to evaluate the persistence of the response effects and the parameters of the product depending on various factors.

The phenolic content in lettuce samples treated with all doses of UV-A and low-moderate doses of UV-B decreased during storage. On the other hand, the total phenolic content in the lettuce leaves treated with high dose UV-B increased by 53.7% after 7 storage days. The results obtained in the present study were similar to the literature. In the study conducted by Du et al. (2014), an increase of approximately 66% in total phenolic content was found on the 6th day in the lettuce samples related with low dose UV-B. Results also indicated an approximately 150% increase in lettuce's total phenol amount in the case of medium-high dose UV-B on the 6th day (Du et al. 2014). Hence it was implied that high total phenolic content requires further storage periods.

In this study, the total phenolic content of lettuce leaves treated with high dose UV-B was better compared to control samples. In addition, a decrease in the total amount of phenol was detected in low and moderate UV-B and UV-A treatments on the 7th day of storage compared to the control samples in this study. Although the total phenolic content of lettuce varied quantitatively and qualitatively according to the lettuce species, it was thought that such different effects might have occurred as the effects in

UV applications had been related to the absorbed doses. Interdonato et al. (2011) stated in their study that different types of plants had exhibited irregular behavior and therefore had reacted differently to increased UV radiation due to variant UV absorbers in plant structure. Similarly, Scattino et al. (2016) obtained variant results with UV-B treatment. The total phenol amounts of peaches both increased and decreased which was related to varying genotypes.

Antioxidant activities in lettuce samples were calculated using the EC_{50} value, an indicator of the amount of antioxidants required to scavenge 50% of DPPH concentration. Results indicated higher antioxidant activity of the sample in the case of a lower calculated EC_{50} value. UV treatments and storage time caused a statistically significant difference (0.05) in antioxidant activities in lettuce samples. The EC_{50} value of the control lettuce samples was 12.2 mg mL^{-1}. EC_{50} values decreased in the control samples on the 1st day of storage. Since the low EC_{50} value was an indicator of high antioxidant activity, it was determined that the antioxidant compound accumulation had increased on the 1st day of storage and had been preserved afterwards (Figure 5b).

High dose UV-B application increased antioxidant activity, and the EC_{50} value was the lowest (4.5 mg mL^{-1}) on the 7th storage day. High dose UV-B application increased the antioxidant activity by 29.7% compared to the control samples on the 7th storage day. Low and medium doses of UV-B applications increased the antioxidant capacity temporarily on the 1st day, however the values decreased on the 3rd and 7th storage days. Similar to our results, Formica-Oliveira et al. (2017) detected increases in antioxidant capacities at different levels of UV-B exposure applied to broccoli leaves, stems, and crown parts. It was determined that the lettuce samples, treated with medium and high doses of UV-A, had the lowest DPPH radical scavenging effects on the 7th day with the highest EC_{50} values of 13.2 mg mL^{-1} and 13.3 mg mL^{-1}, respectively.

Increases and decreases in the amount of total antioxidant substance during storage were similar to the changes in total phenolic content. However, a linear correlation was not detected between the results. The differences were related to the changing characteristics of the plant. The total amount of phenol in green lettuce constitutes more than 60% of the total antioxidant capacity, while the amount of ascorbic acid was responsible for 3.2%-24.5% of the antioxidant capacity (Nicolle et al. 2004).

As scientists have declared, exposing plants to UV may cause abiotic stress. Plants respond to stress by revving up their production of their own natural enzymes. As the production of enzymes increases, levels of phenolic compounds and antioxidants synthesized by the enzymes also increase. Despite this and other knowledge about plants' responses to stress and UV-B, the idea of using UV-B to quickly, safely, and conveniently enrich the antioxidant heft of fresh produce has not been extensively studied (Paul et al. 2012; Espinosa-Leal et al. 2022).

4. Conclusions

The demand for freshly cut or minimally processed products has increased in the world food markets. This demand leads to an increase in fruit and vegetable production. However, the increase in production has brought many problems regarding the maintenance of product quality throughout the processes reaching the consumer, and the problems have necessitated a new approach to the subject. In today's world, where modern technologies are applied, there is a struggle to minimize the loss of nutrients and food properties with the use of ultraviolet light applications; one of the most recent technologies in post-harvest storage of fruits and vegetables.

Accordingly, lettuce, one of the most commonly used vegetables in ready-to-eat salads, was treated with UV-A and UV-B at determined intensities and doses to clarify their effects on quality characteristics during storage. It was statistically shown that UV-A and UV-B application doses and storage periods had been effective on quality characteristics such as colour, ascorbic acid, antioxidant activity, and total phenolic matter. However, UV dose applications showed variant trends in each quality value. The ascorbic acid loss of UV-treated lettuce was negligible. The high-dose UV-A application was the most effective application in terms of protection and improvement of the ascorbic acid amount. As a result of high-dose UV-A, the amount of ascorbic acid increased by 16.7% compared to the control sample. On the other hand, high dose UV-B enabled ascorbic acid to increase by 10.9% compared to the control sample. In terms of total phenolic content, the highest values were obtained in lettuce treated with high dose UV-B after 7 storage days. There was a 53.7% increase in the total phenolic content in high-dose UV-B treated lettuce samples compared to control. Similarly, it was determined that a high dose of UV-B was effective in increasing the antioxidant activity of lettuce, and antioxidant activity increased by 29.7% on the 7th storage day compared to control.

As a result, it was shown in this study that lettuce samples exposed to UV-A and UV-B light might have the potential to be used to maintain or improve some quality features. Among the three doses of UV-A and UV-B treatments, high-dose treatments had been more effective. However, more studies are needed to illuminate the effects of UV-A or UV-B on food product contents in general. In future studies, it is thought that optimization of different doses of UV-A and UV-B would have been beneficial in the context of process conditions and UV-treated food products. Innovative application solutions create value in lettuce within a storage period that otherwise might be lost or wasted. Good UV-A and B applications are key to maintaining quality, preserving nutrient content, gaining higher market prices, and reducing losses of fresh produce. Considering that almost half of the vegetables are wasted before they can be sold, interventions related to post-harvest handling are also economically important.

Acknowledgments

The authors appreciate the support from the Biotechnology Application and Research Center at Bilecik Şeyh Edebali University. The study was conducted as part of a research project supported by the Research Fund of Bilecik Şeyh Edebali University (BAP project-2020-02.BŞEÜ.25-04).

References

Aiamla-or S, Kaewsuksaeng S, Shigyo M & Yamauchi N (2010). Impact of UV-B irradiation on chlorophyll degradation and chlorophyll-degrading enzyme activities in stored broccoli (*Brassica oleracea* L. Italica Group) florets. *Food Chemistry* 120(3): 645–651. https://doi.org/10.1016/J.FOODCHEM.2009.10.056

Aiamla-or S, Yamauchi N, Takino S & Shigyo M (2009). Effect of UV-A and UV-B irradiation on broccoli (*Brassica oleracea* L. Italica Group) floret yellowing during storage. *Postharvest Biology and Technology* 54(3): 177–179. https://doi.org/10.1016/J.POSTHARVBIO.2009.07.006

Allende A, McEvoy JL, Luo Y, Artes F & Wang CY (2006). Effectiveness of two-sided UV-C treatments in inhibiting natural microflora and extending the shelf-life of minimally processed "Red Oak Leaf" lettuce. *Food Microbiology* 23(3): 241–249. https://doi.org/10.1016/J.FM.2005.04.009

Azarafshan M, Peyvandi M, Abbaspour H, Noormohammadi Z & Majd A (2020). The effects of UV-B radiation on genetic and biochemical changes of *Pelargonium graveolens* L'Her. *Physiology and Molecular Biology of Plants: An International Journal of Functional Plant Biology* 26(3): 605–616. https://doi.org/10.1007/s12298-020-00758-6

Castagna A, Chiavaro E, Dall'Asta C, Rinaldi M, Galaverna G & Ranieri A (2013). Effect of postharvest UV-B irradiation on nutraceutical quality and physical properties of tomato fruits. *Food Chemistry* 137(1–4): 151–158. https://doi.org/10.1016/J.FOODCHEM.2012.09.095

Cefola M & Pace B (2023). Advances postharvest preservation technology. *Foods* 12(8): 1664. https://doi.org/10.3390/foods12081664

Collado-González J, Piñero MC, Otalora G, Lopez-Marín J & del Amor FM (2022). Unraveling the nutritional and bioactive constituents in baby-leaf lettuce for challenging climate conditions. *Food Chemistry* 384: 132506. https://doi.org/10.1016/J.FOODCHEM.2022.132506

Csapó J, Prokisch J, Albert CS & Sipos P (2019). Effect of UV light on food quality and safety. *Acta Universitatis Sapientiae, Alimentaria* 12(2019): 21-41. https://doi.org/10.2478/ausal-2019-0002

Das R & Bhattacharjee C (2020). Lettuce. In A. K. Jaiswal (Ed.). *Nutritional composition and antioxidant properties of fruits and vegetable*s. United Kingdom: Academic Press. p. 143–157.

Du WX, Avena-Bustillos RJ, Breksa AP & McHugh TH (2014). UV-B light as a factor affecting total soluble phenolic contents of various whole and fresh-cut specialty crops. *Postharvest Biology and Technology* 93: 72–82. https://doi.org/10.1016/J.POSTHARVBIO.2014.02.004

Espinosa-Leal CA, Mora-Vásquez S, Puente-Garza CA & Alvarez-Sosa DS (2022). Recent advances on the use of abiotic stress (water, UV radiation, atmospheric gases, and temperature stress) for the enhanced production of secondary metabolites on in vitro plant tissue culture. *Plant Growth Regulation* 97: 1–20. https://doi.org/10.1007/s10725-022-00810-3

Formica-Oliveira AC, Martínez-Hernández GB, Díaz-López V, Artés F & Artés-Hernández F (2017). Use of postharvest UV-B and UV-C radiation treatments to revalorize broccoli byproducts and edible florets. *Innovative Food Science & Emerging Technologies* 43: 77–83. https://doi.org/10.1016/J.IFSET.2017.07.036

Guerrero-Beltr·n JA & Barbosa-C·novas GV (2004). Advantages and limitations on processing foods by UV light. *Food Science and Technology International* 10(3): 137–147. https://doi:10.1177/1082013204044359

Hagen SF, Borge GIA, Bengtsson GB, Bilger W, Berge A, Haffner K & Solhaug KA (2007). Phenolic contents and other health and sensory related properties of apple fruit (*Malus domestica* Borkh., cv. Aroma), Effect of postharvest UV-B irradiation. *Postharvest Biology and Technology* 45(1): 1–10. https://doi.org/10.1016/J.POSTHARVBIO.2007.02.002

Interdonato R, Rosa M, Nieva CB, González JA, Hilal M & Prado FE (2011). Effects of low UV-B doses on the accumulation of UV-B absorbing compounds and total phenolics and carbohydrate metabolism in the peel of harvested lemons. *Environmental and Experimental Botany* 70(2–3): 204–211. https://doi.org/10.1016/J.ENVEXPBOT.2010.09.006

Janghu S, Kumar V & Yadav AK (2024). Post-harvest management of fruits and vegetables. *Current Perspectives in Agriculture and Food Science* 7: 125–148. https://doi.org/10.9734/bpi/cpafs/v7/7984E

Jansen MAK, Hectors K, O'Brien NM, Guisez Y & Potters G (2008). Plant stress and human health: Do human consumers benefit from UV-B acclimated crops? *Plant Science* 175(4): 449–458. https://doi.org/10.1016/J.PLANTSCI.2008.04.010

Kang HM & Saltveit ME (2002). Antioxidant capacity of lettuce leaf tissue increases after wounding. *Journal of Agricultural and Food Chemistry* 50(26): 7536–7541. https://doi.org/10.1021/JF020721C

Kasim MU & Kasim R (2017a). The effects of ultraviolet B (UV-B) irradiation on color quality and decay rate of Capia pepper during postharvest storage. *Food Science and Technology International* 38(2): 363–368. https://doi.org/10.1590/1678-457X.05817

Kasim MU & Kasim R (2017b). Yellowing of fresh-cut spinach (*Spinacia oleracea* L.) leaves delayed by UV-B applications. *Information Processing in Agriculture* 4(3): 214–219. https://doi.org/10.1016/J.INPA.2017.05.006

Kim MJ, Moon Y, Tou JC, Mou B & Waterland NL (2016). Nutritional value, bioactive compounds and health benefits of lettuce (*Lactuca sativa* L.). *Journal of Food Composition and Analysis* 49: 19–34. https://doi.org/10.1016/J.JFCA.2016.03.004

Koutchma T (2009). Advances in ultraviolet light technology for non-thermal processing of liquid foods. *Food and Bioprocess Technology* 2(2): 138–155. https://doi.org/10.1007/S11947-008-0178-3/FIGURES/3

Koutchma T (2019). *Ultraviolet light in food technology: Principles and applications*. New york: CRC Press. 337 p.

Liu C, Han X, Cai L, Lu X, Ying T & Jia Z (2011). Postharvest UV-B irradiation maintains sensory qualities and enhances antioxidant capacity in tomato fruit during storage. *Postharvest Biology and Technology* 59(3): 232–237. https://doi.org/10.1016/J.POSTHARVBIO.2010.09.003

Liu X, Ardo S, Bunning M, Parry J, Zhou K, Stushnoff C, Stoniker F, Yu L & Kendall P (2007). Total phenolic content and DPPH radical scavenging activity of lettuce (*Lactuca sativa* L.) grown in Colorado. *LWT-Food Science and Technology* 40(3): 552–557. https://doi.org/10.1016/J.LWT.2005.09.007

Llorach R, Tomás-Barberán FA & Ferreres F (2004). Lettuce and chicory by products as a source of antioxidant phenolic extracts. *Journal of Agricultural and Food Chemistry* 52(16): 5109–5116. https://doi.org/10.1021/JF040055A

Mampholo BM, Maboko MM, Soundy P & Sivakumar D (2016). Phytochemicals and overall quality of leafy lettuce (*Lactuca sativa* L.) varieties grown in closed hydroponic system. *Journal of Food Quality* 39(6): 805–815. https://doi.org/10.1111/JFQ.12234

Martínez-Ispizua E, Calatayud Á, Marsal JI, Basile F, Cannata C, Abdelkhalik A, Soler S, Valcárcel JV & Martínez-Cuenca MR (2022). Postharvest changes in the nutritional properties of commercial and traditional lettuce varieties in relation with overall visual quality. *Agronomy* 12(2): 403. https://doi.org/10.3390/AGRONOMY12020403/S1

Muscolo A, Marra F, Canino F, Maffia A, Mallamaci C & Russo M (2022). Growth, nutritional quality and antioxidant capacity of lettuce grown on two different soils with sulphur-based fertilizer, organic and chemical fertilizers. *Scientia Horticulturae* 305: 111421. https://doi.org/10.1016/J.SCIENTA.2022.111421

Nadkarni BY (1965). Determination of ascorbic acid in coloured extracts: A new modification of the indophenol technique. *Mikrochimica Acta* 53(1): 21–27. https://doi.org/10.1007/BF01218730/METRICS

Nguyen CTT, Kim J, Yoo KS, Lim S & Lee EJ (2014). Effect of prestorage UV-A, -B, and -C radiation on fruit quality and anthocyanin of "Duke" blueberries during cold storage. *Journal of Agricultural and Food Chemistry* 62(50): 12144–12151. https://doi.org/10.1021/JF504366X/ASSET/IMAGES/MEDIUM/JF-2014-04366X_0005.GIF

Nicolle C, Carnat A, Fraisse D, Lamaison JL, Rock E, Michel H, Amoureux P & Remesy C (2004). Characterisation and variation of antioxidant micronutrients in lettuce (*Lactuca sativa* folium). *Journal of the Science of Food and Agriculture* 84(15): 2061–2069. https://doi.org/10.1002/JSFA.1916

Pathare PB, Opara UL & Al-Said FAj (2013). Colour measurement and analysis in fresh and processed foods: A review. *Food Bioprocess Technology* 6: 36–60. https://doi.org/10.1007/s11947-012-0867-9

Paul ND, Moore JP, McPherson M, Lambourne C, Croft P, Heaton JC & Wargent JJ (2012). Ecological responses to UV radiation: Interactions between the biological effects of UV on plants and on associated organisms. *Physiologia Plantarum* 145(4): 565–581. https://doi:10.1111/j.1399-3054.2011.01553.x

Pérez-López U, Miranda-Apodaca J, Muñoz-Rueda A & Mena-Petite A (2013). Lettuce production and antioxidant capacity are differentially modified by salt stress and light intensity under ambient and elevated CO_2. *Journal of Plant Physiology* 170(17): 1517–1525. https://doi.org/10.1016/J.JPLPH.2013.06.004

Salemi S, Saedisomeolia A, Azimi F, Zolfigol S, Mohajerani E, Mohammadi M & Yaseri M (2021). Optimizing the production of vitamin D in white button mushrooms (*Agaricus bisporus*) using ultraviolet radiation and measurement of its stability. *Lebensmittel-Wissenschaft + [i.e. Und] Technologie. Food Science + Technology. Science + Technologie Alimentaire* 137: 110401. https://doi.org/10.1016/J.LWT.2020.110401

Scattino C, Negrini N, Morgutti S, Cocucci M, Crisosto CH, Tonutti P, Castagna A & Ranieri A (2016). Cell wall metabolism of peaches and nectarines treated with UV-B radiation: A biochemical and molecular approach. *Journal of the Science of Food and Agriculture* 96(3): 939–947. https://doi.org/10.1002/JSFA.7168

Tikekar RV, Anantheswaran RC, Elias RJ & Laborde LF (2011). Ultraviolet-induced oxidation of ascorbic acid in a model juice system: Identification of degradation products. *Journal of Agricultural and Food Chemistry* 59(15): 8244–8248. https://doi.org/10.1021/JF201000X

Tomás-Callejas A, Otón M, Artés F & Artés-Hernández F (2012). Combined effect of UV-C pretreatment and high oxygen packaging for keeping the quality of fresh-cut Tatsoi baby leaves. *Innovative Food Science & Emerging Technologies* 14: 115–121. https://doi.org/10.1016/J.IFSET.2011.11.007

Usall J, Ippolito A, Sisquella M & Neri F (2016). Physical treatments to control postharvest diseases of fresh fruits and vegetables. *Postharvest Biology and Technology* 122: 30–40. https://doi.org/10.1016/J.POSTHARVBIO.2016.05.002

Wang W, Zhang C, Shang M, Lv H, Liang B, Li J & Zhou W (2022). Hydrogen peroxide regulates the biosynthesis of phenolic compounds and antioxidant quality enhancement in lettuce under low nitrogen condition. *Food Chemistry: X* 16: 100481. https://doi.org/10.1016/J.FOCHX.2022.100481

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Application of the Different Machine Learning Algorithms to Predict Dry Matter Intake in Feedlot Cattle

Özgür Koşkan[a], Malik Ergin[a*], Hayati Köknaroğlu[a]

[a]*Department of Animal Science, Faculty of Agriculture, Isparta University of Applied Sciences, 32000, Isparta, TÜRKIYE*

ARTICLE INFO
Research Article
Corresponding Author: Malik Ergin, E-mail: malikergin@isparta.edu.tr
Received: 13 October 2023 / Revised: 06 August 2024 / Accepted: 07 August 2024 / Online: 14 January 2025

Cite this article
Koşkan Ö, Ergin M, Köknaroğlu H (2025). Application of the Different Machine Learning Algorithms to Predict Dry Matter Intake in Feedlot Cattle. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):91-99. DOI: 10.15832/ankutbd.1375383

ABSTRACT
Due to the development of computing technology and different machine learning models, big data sets have gained importance in animal science as well as in many disciplines. The main objective of this study was to compare different machine learning algorithms to predict daily dry matter intake (DMI) in feedlot cattle. The data consisted of 2660 cattle pens placed on feed between January 1988 and December 1997. Machine learning methods were compared in heifers and steers, with 718 in pens of heifers and 1942 in pens of steers. Initial body weight, days on feed, and average proportion of dietary concentrate were used as independent variables to predict DMI in steers and heifers separately. The multivariate linear regression (LR), random forest (RF), gradient boosting regressor (GBR), and light gradient boosting machine (LGBR) algorithms were compared in terms of several performance metrics (MAE, MAPE, MSE, and RMSE). Results showed that the determination coefficient alone is not a good single criterion. It is recommended that the interpretation of model consistency should also consider MAE, MAPE, MSE, and RMSE values. In the current study, all machine learning algorithms yielded similar and lower performance metrics. However, the LGBR and GBR algorithms, were found to perform slightly better than the other algorithms, especially in heifers. Increasing the number of animals and using different independent variables that are related to the DMI can affect the accuracy of DMI prediction.

Keywords: Bigdata, Feedlot cattle, Machine learning algorithms

1. Introduction

With the advancement of computer and internet technologies in recent years, the amount of data has reached a huge size. Thus, "big data" and "data science" have become the most important subjects in science. Fuzzy logic, artificial neural networks, and machine learning methods have been widely used as computer algorithms that model the dataset of classification or estimation problems (Atalay & Çelik 2017). Especially big data, which is seen as the most valuable information of the future, and data mining, which is the technique of processing this data, have become the most important subject of science. This situation has drawn attention in various fields and emphasized using existing data mining methods. Data mining has also been used in animal husbandry to develop prediction models using artificial intelligence. Asadzadeh et al. (2021) compared seven different machine learning methods to predict the live weight of camels by applying different body measurements. Mikail et al. (2014) predicted daily milk yield in Holstein cattle by using support vector machines and artificial neural network models. Huma & Iqbal (2019) used regression trees, support vector machines, and random forest models to predict live weight in Balochi rams, a Pakistani sheep breed. Mammadova & Keskin (2013) detected subclinical and clinical mastitis using support vector machines in cattle. In this study, we used multivariate linear regression (MLR), random forests (RF), gradient boosting regressor (GBR), and light gradient boosting machine algorithms (LGBM) to predict dry matter intake (DMI) in beef cattle. Since DMI is the basis for the calculation and prediction of nutrient requirements, gain, and profit, DMI must be estimated accurately (Hicks et al. 1990). Combined data from cattle fed high-energy diets and initial weight on feed could be used to predict DMI of cattle (NASEM 2016). Koknaroglu et al. (2017) predicted dry matter intake of steers and heifers in the feedlot by using initial weight. Koskan et al. (2014) predicted dry matter intake of steers and heifers in the feedlot by using categorical and continuous variables. The purpose of this study was to predict dry matter intake of feedlot cattle by using machine learning. In the present study, RF, GBR and LGBR algorithms are tree-based algorithms. These tree-based algorithms have a similar but slightly different mathematical background. In addition, when nonlinear relationships between variables appear, these algorithms can be beneficial for accurate prediction. Therefore, a comparative analysis of these algorithms can be useful for future research.

2. Material and Methods

2.1. Material

Closeout information, which was gathered through the Iowa State University Animal Science Extension Program from Iowa cattle producers using the Iowa State University Feedlot Performance and Cost Monitoring Program, was used to derive data for this study. The following information related to animal performance and management were provided and received on mailed-in data sheets: starting date on feed, end of feeding period date, number of cattle in the pen, sex (1= steer, 2= heifer), facility code (1= confinement, 2= partially open lot, 3= open lot), days on feed, initial pay weight, final pay weight, feed efficiency (FE), average percent concentrate, average daily gain (ADG), percent death loss. To obtain a detailed information about the material, study conducted by Koknaroglu et al. (2005), that examined the factors affecting the performance and profitability of beef cattle should be examined. Since DMI was not provided in the close-out sheets, DMI was generated by computer by using the equation DMI = ADG × FE.

From 1988 through 1997, a total of 405 573 animals were represented in the 2759 pens with an average of 150 cattle per pen. No information was available concerning the age or background of the animals. Average body weight at starting time was 322 kg, and animals were fed for an average of 172 d. The 2759 pens consisted of 2032 pens of steers and 727 pens of heifers. Average percentage concentrate ratio was 81%. Each observation (pen of cattle within close-outs) was accepted as an independent observation, even though some observations were obtained from the same farm.

The study used days on feed (DF), initial weight (IW), and concentrate ratio (PRC) as independent variables to predict DMI. The frequency distribution of DMI is shown in Figure 1. In addition, the descriptive statistics of continuous variables are presented in Table 1. Several machine learning methods used in heifers and steers were compared, with 718 in pens of heifers and 1942 in pens of steers.

Table 1- Descriptive statistics for numerical variables

Variable	*Mean*[a]	*SD*[b]	*CV*[c]
Days on feed (d)	172.49	51.24	29.70
Initial weight (kg)	321.86	60.21	18.71
Concentrate ratio (%)	81.41	7.67	9.43
Dry matter intake (kg)	9.77	1.26	12.92

a, b, c: Arithmetic mean, Standard deviation, Coefficient of variation

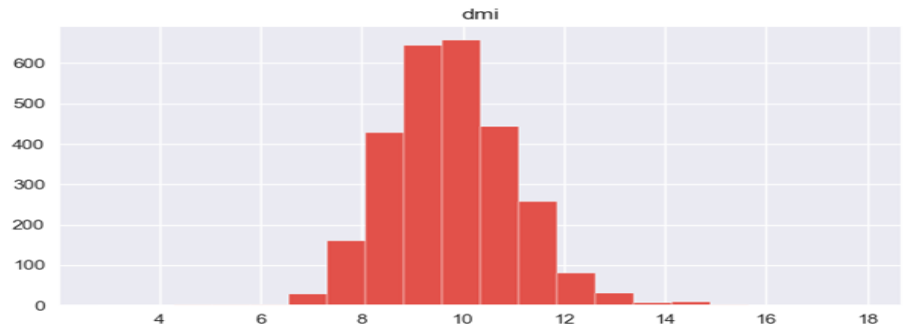


Figure 1- The frequency distribution of dry matter intake (DMI)

2.2. The machine learning algorithms used in the study

2.2.1. Multivariate linear regression (MLR)

Multivariate linear regression is a frequently used and functional algorithm among machine learning algorithms. The multivariate linear regression algorithm best explains the relationship between independent variables and a dependent variable in a linear form. The assumptions of this model include the normal distribution of the data and the elimination of multicollinearity among independent variables. Unlike a simple linear regression model that involves single independent and single dependent variable, an increase in the number of independent variables in a multivariate linear regression model can lead to multicollinearity among these independent variables Equation (1). In general, in a dataset, the desired situation is for each independent variable to have a high correlation with the dependent variable and for the independent variables to have a low correlation with each other (Ray, 2019).

$$DMI_i = \beta_0 + \beta_1 \text{IBW} + \beta_2 \text{DF} + \beta_1 \text{PRCCONC} \ldots\ldots + e_i \quad (1)$$

Where; β_0, Constant; $\beta_1X_1 + \ldots + \beta_nX_n$: coefficient of regression, i: refers to the pen of cattle, IBW, DF, and PRCCONC states the value for i_{th} observation, e: random error term

In the mathematical model given above, β1 means that, the increase in the unit of the dependent variable DMI is equal to the increase in independent variable's (IBW, DF, or PRCCONC) unit.

2.2.2. Random forest (RF)

RF is a supervised machine learning algorithm used in both regression and classification problems when the dependent variable is continuous or categorical, respectively. Its theoretical background is similar to the classification and regression tree (CART). In CART algorithm, single decision tree is used for classification or regression. Therefore, it can cause the overfitting problem on the training dataset. The RF algorithm fixes the major disadvantage of decision trees called overfitting. This is achieved through the combination of multiple classification and regression trees (CARTs), where each tree provides its own prediction, adding diversity to the model (Breiman 2001; Müller & Guido 2016). The basic principle of the RF method is based on the minimization of a function. Random vectors containing observations in dependent and independent variables are defined as *X* and *Y* respectively. It is assumed that the joint distribution, which is a probability distribution of the relationship between these two random vectors, is defined by $P_{XY}(X, Y)$. The main purpose of all these assumptions is to determine an independent function *f(X)* related to the observation in the *X* vector to predict the observation value in the dependent variable. Therefore, this prediction function is determined by a loss function *L(Y, f(X))*, which needs to be minimized. Through this loss function, a penalty technique is applied to measure the distance between *f(X)* and the *Y* vector, penalizing *f(X)* values that are far from *Y*. In the current study, as in the RF method, the least squares method presented in Equation 2 is used to apply the loss function to regression problems:

$$L\big(Y, f(X)\big) = (Y - f(X))^2 \tag{2}$$

To minimize the loss function, the sum of the k basic learners, denoted by $b = [h_1(X), h_2(X), h_3(X),..., h_k(X)]$ in Equation 3, defines the ensemble predictor *f(X)*. This function gives the best prediction of *Y* (Cutler et al. 2012; Bovo et al. 2021).

$$f(X) = \frac{1}{n}\sum_{i=1}^{n} h_i(X) \tag{3}$$

The most advantageous aspect of the RF algorithm is that it can be effectively and easily used in cases of nonlinear relationships between variables. In addition, the RF algorithm is fast for predictions and can handle overfitting problems (Breiman 2015; Çelik & Yılmaz 2023).

In the present study, IW, DF, and PRC variables were considered as independent variables (inputs) in the training dataset. The pen DMI values were considered as dependent (output) in the training dataset. The RF algorithm tries to establish a relationship between the inputs and output variables to predict the pen DMI.

2.2.3. Gradient boosting regressor (GBR)

The gradient boosting algorithm is a tree-based ensemble method developed to enhance predictive performance with respect to the dependent variable in both regression and classification problems. In the boosting method, a series of simple models called weak learners is constructed. These simple models are utilized to correct the errors made by previous models. Similar to RF, it is also formed by combining decision trees, but each one was trained by adjusting the amount of error made by the previous one. When considering the base unit as a decision tree, the final ensemble is indicated as boosted tree (Di Persio & Fraccarolo, 2023). When there are *n* observations in Equation 4, and it is assumed that each observation value of the independent variable x corresponds to a value in the dependent variable y, the GBR algorithm aims to find an estimate $\hat{f}(x)$ which approximates the function $f^*(x)$ that maps observations to the dependent variable.

$$S = \{(x_i, y_i)\}_{i=1}^{n} \tag{4}$$

To achieve this, the algorithm minimizes the expected value of the loss function $L(y, f(x))$. Then, as seen in Equation 5, the additive prediction of the $f^*(x)$ is generated by weighting all obtained functions.

$$f_k(x) = f_{k-1}(x) + p_k h_k(x) \tag{5}$$

Where: p_k, weight of the t^{th} base learner (k = 1, 2, ..., K); h_k, a base learner.

Suppose $L(y_i, a)$ is differentiable loss function, the prediction of the $f^*(x)$ is calculated by an iterative process in Equation 6. Here, in each new iteration, a new tree is constructed that corrects the errors remaining from the predictions of the previous tree.

$$f_0(x) = argmin_a \sum_{i=1}^{n} L(y_i, a) \quad (6)$$

The base learners seek to minimize the expected value of the loss function $L(y_i, a)$ by Equation 7.

$$\left(p_k h_k(x)\right) = argmin_{p,h} \sum_{i=1}^{n} L(y_i, f_{k-1}(x_i) + ph(x_i)) \quad (7)$$

Subsequently, the pseudo-residuals of each observation, which represents the error remaining from the prediction of the previous tree, are calculated according to Equation 8 (Sibindi et al., 2022; Otchere et al., 2022).

$$r_{ti} = \left[\frac{\partial L(y_i, f(x))}{\partial f(x)}\right]_{f(x)=f_{k-1}(x)} \quad (8)$$

There are several advantages such as robustness against non-linear relationships among variables, handling of outliers in the dataset, automatic feature selection for predicting the dependent variable, ability to work with independent variables that have high linear correlations with each other (multicollinearity) and support for various loss functions (Hastie et al. 2009; Ogutu et al. 2011; Hong 2015). The GBR algorithm creates a series of decision trees to estimate the pen DMI in steers and heifers. Each tree contains rules that determine the value of the pen DMI value based on the observations in the input variables (IW, DF, and PRC). The individual predictions of each tree are weighted and combined to optimize the overall prediction accuracy of the model.

2.2.4. Light gradient boosting regressor (LGBR)

Similar to GBR algorithm, the LGBR algorithm uses decision trees in classification and prediction problems. This algorithm has faster training speed and higher performance than many other algorithms when creating models (Chen et al. 2019). Contrary to other tree-based algorithms such as GBR, and XBGR, LGBR algorithms typically grow the tree vertically which is one of the most effective aspects of LGBR for handling large-scale data and variables (Sun et al. 2018). The mathematical background of the LGBR algorithm is similar to GBM but differs in some aspects. A detailed explanation of the calculations for LGBR was given according to Sun et al. (2018). When the training data set is assumed to be as in equation 4, the expected value of the loss function is calculated with Equation 9.

$$\hat{f} = \arg\min_f E_{y,X}\, L\big(y, f(x)\big) \quad (9)$$

The LGBR algorithm integrates *T* regression trees to make its final model using Equation 10.

$$f_T(X) = \sum_{t=1}^{T} f_t(X) \quad (10)$$

Regression trees are characterized by the number of leaves *J* and an index *q* representing the rules of the tree, where the example weight $w_{q(x)}$ applies to the q^{-th} leaf of the regression tree, q ∈{1,2, …, J}. Therefore, in Equation 11, LGBR is additively trained over t steps.

$$\Gamma_t = \sum_{i=1}^{n} L\big(y_i, F_{t-1}(x_i) + f_t(x_i)\big) \quad (11)$$

The objective function is quickly approximated with Newton's approach. After some simplification steps, Equation 11 will be replaced by Equation 12.

$$\Gamma_t \cong \sum_{i=1}^{n} \left(g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)\right) \quad (12)$$

In Equation (X) g_i and h_i states the 1st and 2nd order gradient statistics of the loss function. When *Ij* represents the sample set of leaf *j*, Equation 12 can be explained as Equation X.

$$\Gamma_t = \sum_{j=1}^{j} \left(\left(\sum_{i\in I_j} g_i\right) w_j + \frac{1}{2}\left(\sum_{i\in I_j} h_i + \lambda\right) w_j^2\right) \quad (13)$$

In the case of $q(x)$ a tree structure, the optimal scores of the leaf weight for leaf nodes w_j^* and extreme values of Γ_k would be expressed as in Equations 14 and 15.

$$w_j^* = -\frac{\sum_{i\in I_j} g_i}{\sum_{i\in I_j} h_i + \lambda} \quad (14)$$

$$\Gamma_T^* = -\frac{1}{2}\sum_{j=1}^{J} \frac{\left(\sum_{i\in I_j} g_i\right)^2}{\sum_{i\in I_j} h_i + \lambda} \quad (15)$$

The objective function is finally calculated by integrating the split in Equation 16.

$$G = \frac{1}{2}\left(\frac{\left(\sum_{i\in I_L} g_i\right)^2}{\sum_{i\in I_L} h_i+\lambda} + \frac{\left(\sum_{i\in I_R} g_i\right)^2}{\sum_{i\in I_R} h_i+\lambda} - \frac{(\sum_{i\in I} g_i)^2}{\sum_{i\in I} h_i+\lambda}\right) \quad (16)$$

Where; I_L and I_R state the left and right branches, respectively.

2.3. Evaluation metrics of prediction models

A few performances scores that are frequently used in the literature were used to compare algorithms for predicting capability of pen DMI. The main objective of using these evaluation criteria was to compare the performance of the machine learning models we have used. In this study, evaluation criteria such as mean absolute error (MAD), mean squared error (MSE), root mean squared error (RMSE), mean absolute percentage error (MAPE) were considered (Çelik & Yılmaz 2017). Equations for performance scores of R^2, MAD, MAPE, MSE and RMSE are presented in Equations (3-7).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - y_{ip})^2}{\sum_{i=1}^{n}(y_{ip} - \bar{y}_{ip})^2} \quad (3)$$

$$MAD = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - y_{ip}\right| \quad (4)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - y_{ip}}{y_i}\right| * 100 \quad (5)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - y_{ip})^2 \quad (6)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - y_{ip})^2} \quad (7)$$

Where: n is the number of animals, p is the number of dependent variables for predicting; DMI, y_i is the actual observation value of DMI of cattle, y_{ip} is the predicted DMI.

K-fold cross validation was applied to all models for optimizing the models. Tested and selected hyperparameters for all algorithms were presented in Table 2. In fact, cross validation is a resampling method. In this method, the dataset is divided into sub-samples that are different from each other and have an equal number of observations. Observations from the training dataset are selected into sub-samples randomly and without replacement. The model in question is trained with k-1 sub-grouped samples. In this approach, during each k iterations, a different data-sample is held out for testing while the remaining k-1 sub-sample is used for training. This process is repeated k times, ensuring that each of the k folds is used exactly once for validation. Ultimately, the results obtained from the k iterations are combined to make a prediction (Refaeilzadeh et al. 2016). In the present study, all statistical analyses were performed using R software version 4.4 (R Core Team, 2024). The RF, GBR and MLR algorithms were executed using the *caret* package (version 6.0.94) that consists of several regression algorithms (Kuhn, 2008). The LGBR algorithm was evaluated using the *LightGBM* package (version 4.3.0) in Python (Ke et al., 2017). Prior to conducting the analyses, all data were divided into training and test sets, and then hyperparameters were randomly searched. For steers, the training and test dataset ratios were determined as 60% and 40%, respectively. In heifers, 70% and 30% of observations were randomly split for the training and test datasets, respectively.

In all training processes, days on feed (DF), initial weight (IW), and the average proportion of dietary concentrate (PRC) were considered as independent variables to predict dry matter intake (DMI). The importance of the predictors was assessed using the *varImp* function from the caret package in R. One of the main objectives of machine learning algorithms is to determine which variable is most important in explaining the variation of the dependent variable (dry matter intake, in the current study). Variable importance uses specific coefficients (gain, weight, cover etc.) to evaluate the relationship between the dependent and independent variables. For instance, in multivariate linear regression, each independent variable is ordered based on correlation coefficients to determine its significance in predicting the DMI variable. This process aids in dimensionality reduction and feature selection, which enhance the model's predictive capability. Determining the key independent variables that account for most of the variance in the predictor variable is essential for developing highly predictive models.

Table 2- Summary of the tested and best hyperparameters of all algorithms

	Steers		*Heifers*	
Algorithms	***Tested Parameters***	***Best Parameters***	***Tested Parameters***	***Best Parameters***
MLR	No need	No need	No need	No need
RF	ntreeTry = 1000 mtry = 1:5	ntreeTry = 1000 mtry = 1	ntreeTry = 500 mtry = 1:5	ntreeTry = 500 mtry = 2
GBM	n.trees = 100:250:500:1000 interaction.depth = 1:3 shrinkage = 0.01:0.05:0.1 n.minobsinnode = 20	n.trees = 500 interaction.depth = 3 shrinkage = 0.01 n.minobsinnode = 20	n.trees = 100:250:500:1000 interaction.depth = 1:3 shrinkage = 0.01:0.05:0.1 n.minobsinnode = 20	n.trees = 200 interaction.depth = 3 shrinkage = 0.01 n.minobsinnode = 10
LGBR	learning_rate = 0.05:0.1 boosting_type = gbdt num_leaves = 2:10 max_depth = 2:8 bagging_freq = 2:6 bagging_fraction = 0.7:0.8 iterations = 25:1000	learning_rate = 0.08 boosting_type = gbdt num_leaves = 3 max_depth = 2 bagging_freq = 6 bagging_fraction = 0.75 iterations = 75	learning_rate = 0.05:0.1 boosting_type = gbdt num_leaves = 2:10 max_depth = 2:8 bagging_freq = 2:6 bagging_fraction = 0.7:0.8 iterations = 25:1000	learning_rate = 0.1 boosting_type = gbdt num_leaves = 2 max_depth = 2 bagging_freq = 2 bagging_fraction = 0.4 iterations = 50

GBR: Gradient Boosting Regressor, LGBR: Light Gradient Boosting Machine, MLR: Multivariate Linear Regression and RF: Random Forests

3. Results and Discussion

The box plot of all variables in the models is provided in Figure 2. When examining the box plot graphic, it is observed that initial weight, which is one of the independent variables, fits the normality, the rest deviate from the normality. Therefore, it is decided that non-linear machine learning algorithms are suitable for this dataset.

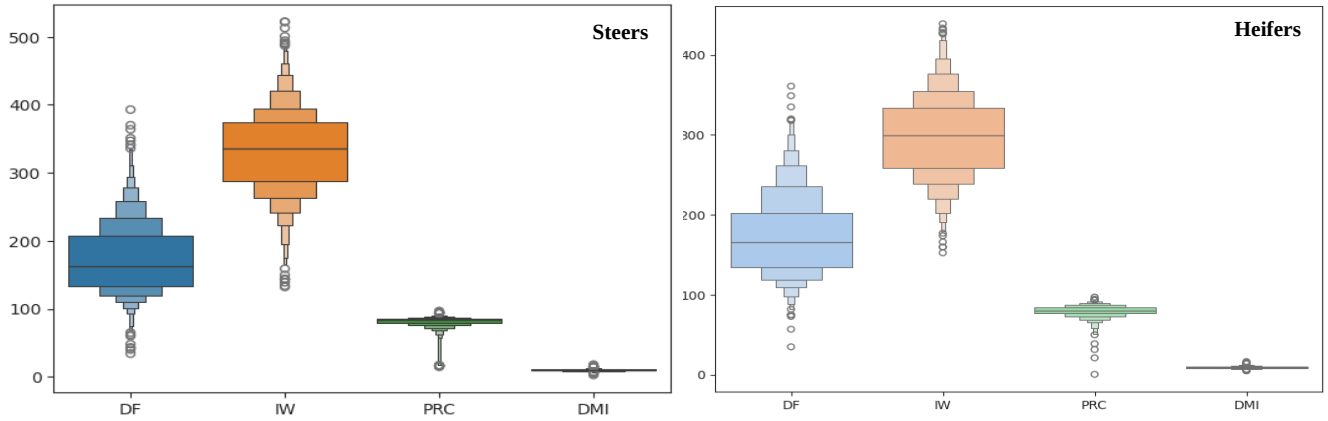


Figure 2- The box plots of the dependent and independent variables in both steers and heifers, DF: days on feed, IW: initial weight (kg), PRC: average proportion of dietary concentrate; DMI: dry matter intake (pen)

In Table 3, the prediction performances of all algorithms on the train and test dataset for predicting DMI in steers and heifers were presented. Various popular performance scores such as R^2, MAD, MAPE, MSE, and RMSE were compared. When looking at the average determination of coefficients (R^2) of the four models, it was seen that GBR and LGBR algorithms explained the variation better than other models. In the test dataset, MAD ranged between 0.64 (RF) and 0.71 (MLR). In addition, MAPE changed between 6.4 (RF) and 7.3 (LR). When assessing the models based on MSE and RMSE, the RF model had the lowest values with 0.72 and 0.85, respectively. According to all performance scores, all algorithms except from MLR, yielded similar predictions on DMI. LGBR and GBR algorithms were considered superior in terms of R^2, while RF performed best in terms of error scores. When training and test datasets were examined simultaneously, it can be said that both LGBR and GBR algorithms resulted quite consistent values in both datasets. The minimal differences between training and test scores suggests that overfitting did not occur in the LGBR and GBR algorithms. Furthermore, it is believed that the LR model, used in this study has a lower impact on predicting daily dry matter intake in cattle than non-linear models.

A test dataset containing 214 observations was used for predicting DMI while 504 observations were used for training in the heifers. For test dataset, the highest average coefficients of determination were found in GBR (R^2 = 0.43) and MLR (R^2 = 0.44). In addition, MAD ranged between 0.69 (RF) and 0.76 (LGBR) in test dataset. MAPE ranged between 7.5 (RF) and 8.4 (LGBR). Regarding MSE and RMSE criteria, the highest values were calculated in LGBR (MSE = 1.15, RMSE = 1.32). The lowest MSE and RMSE values were found as 0.86 and 0.93 for the GBR model, respectively. In all algorithms, it can be observed that the performance scores of all algorithms were not exposed to overfitting and remained consistent based on the relationship between the training and test datasets. Considering all performance scores, the GBR algorithm showed superior results in predicting DMI

in heifers. Statistically, the observed differences between the results of steers and heifers are thought to arise from differences in sample sizes and data splitting.

Table 3- Performances of machine learning algorithms on the training and testing datasets for DMI in steers and heifers

Algorithm	***Steers***									
	Training (n=1225)					***Testing (n=817)***				
	R^2	***MAD***	***MAPE***	***MSE***	***RMSE***	***R^2***	***MAD***	***MAPE***	***MSE***	***RMSE***
RF	0.46	0.65	6.67	0.83	0.91	0.42	0.64	6.4	0.72	0.85
LGBR	0.47	0.64	6.4	0.72	0.85	0.45	0.67	7.0	0.81	0.90
GBR	0.49	0.62	6.35	0.67	0.82	0.45	0.68	6.79	0.86	0.93
MLR	0.40	0.66	6.72	0.83	0.91	0.41	0.71	7.23	0.86	0.93
Algorithm	***Heifers***									
	Training (n=504)					***Testing (n=214)***				
	R^2	***MAD***	***MAPE***	***MSE***	***RMSE***	***R^2***	***MAD***	***MAPE***	***MSE***	***RMSE***
RF	0.44	0.74	7.9	1.01	1.03	0.41	0.69	7.5	0.88	0.94
LGBR	0.43	0.72	7.7	0.99	0.98	0.40	0.76	8.4	1.15	1.32
GBR	0.48	0.71	7.6	1.0	1.0	0.43	0.70	7.73	0.86	0.93
MLR	0.44	0.73	7.7	1.0	1.0	0.44	0.73	7.8	1.0	1.0

n: number of cattle, GBR: Gradient Boosting Regressor, LGBR: Light Gradient Boosting Machine, MLR: Multivariate Linear Regression and RF: Random Forests, R^2: Coefficient of Determination, MAD: Mean Absolute Deviation, MAPE: Mean Absolute Percentage Error, MSE: Mean Squared Error, RMSE: Root Mean Squared Error

The variable importances in steers and heifers were illustrated in Figure 3. Variable importances were achieved using training datasets. Consequently, variable importance plots were generated based on the best algorithm for heifers and steers. GBR was used for heifers, whereas LGBR was used for steers to calculate importances of independent variables. The IW is the most contributed variable for variation in DMI in both steers and heifers. Furthermore, in steers, DF variable is the second most important variable, while in heifers, the PRC variable is the second most important variable.

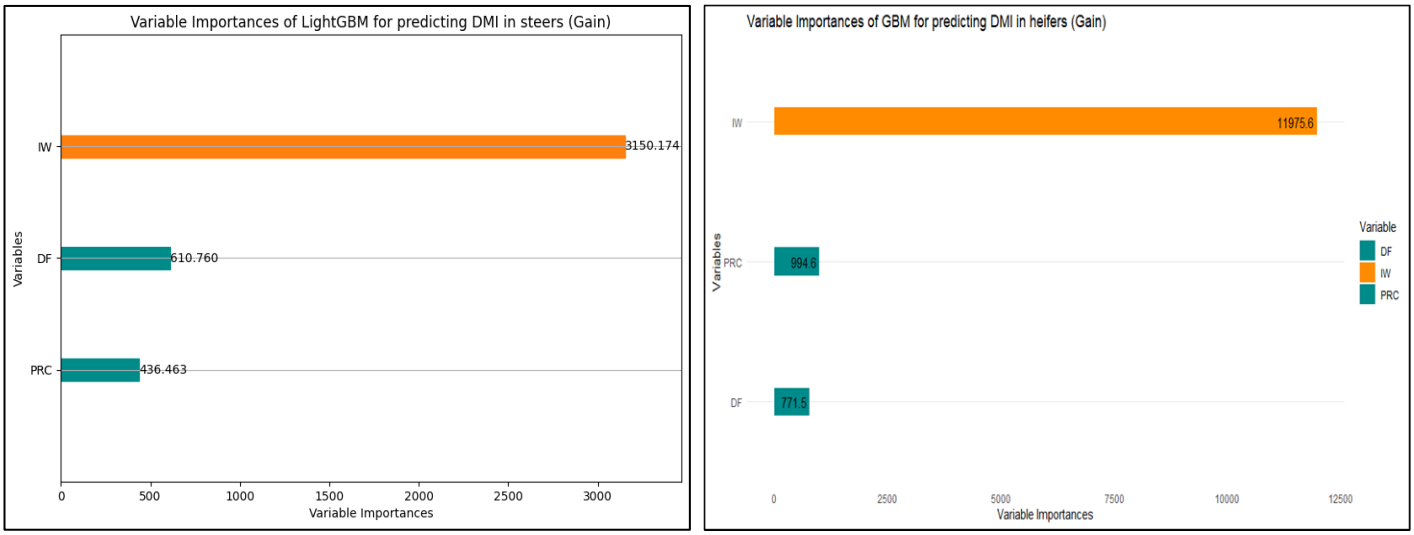


Figure 3- Variable importances of independent variables in steers and heifers for training datasets, DF: days on feed, IW: initial weight (kg), PRC: average proportion of dietary concentrate; DMI: dry matter intake (pen)

In the literature, several DMI predictions were evaluated using different independent variables. Blake et al. (2023) studied repeated measurements ANOVA, repeated measurements random forest regression, and classical random forest regression to predict DMI in 125 bulls and 53 steers. Researchers used different measured variables such as age, sex, full body weight, average daily gain, and climate factors. At the end of the study, the authors stated that the best prediction was performed by the repeated measurements random forest regression model with R2= 0.65 and MSE= 1.09. In the study, the random forest regression model was reported with R2= 0.45 and MSE= 1.71 for predicting DMI. Similarly, the RF model resulted in coefficients of determination of 0.44 and 0.46 in steers and heifers, respectively, in the present study. Authors did not separate results for both bulls and steers. In addition, they used various independent variables for predicting DMI, such as climate factors. These are the main differences between the two studies.

Various technologies have been used for predicting DMI in cattle. For instance, Shadpour et al. (2022) examined different ANN structures to accurately predict weekly DMI in Canadian Holstein cows. Authors used mid-infrared reflectance

spectroscopy (MIRS), weekly average DMI, test-day milk yield, fat yield, protein yield, metabolic body weight, calving traits, country, and herd data. Although we did not use ANN in our study, the results are similar to the findings of the researchers. Additionally, the use of independent variables and a technology such as MIRS also has a significant impact on the results. In addition, Salleh et al. (2023) investigated the use of machine learning algorithms for predicting DMI. They employed partial least squares regression (PLS), support vector machine regression (SVM), and random forest regression (RF) algorithms using MIRS (milk mid-infrared spectra) values. The authors reported that the determination coefficients (R^2) ranged from 0.52 to 0.65. The best determination coefficient was observed in the PLS regression approach with 0.65, followed by 0.62 in RF regression and 0.55 in SVM regression approach. Furthermore, mid-infrared reflectance spectroscopy (MIRS) analysis of milk and near-infrared reflectance spectroscopy (NIRS) analysis of feces from cows were compared in terms of predicting DMI (Lahart et al. 2019). In this comparative study, authors used traditional linear regression and partial least squares regression on the data from 457 cows. In the study where various combinations of the MIRS and NIRS wavelengths with known animal energy sinks and status traits used resulted in the equation with R2= 0.68 and RMSE= 1.52 kg. When compared with the present study, results could change due to independent variables and utilized technologies. In the literature, high determination coefficients are often found in predicting animal body weight, indicating a supportive relationship. Similarly, in predicting DMI, determination coefficients ranged from 0.40 to 0.70, which is consistent with our study.

While a high correlation may seem important, it is crucial to consider the performance of the applied models in revealing the actual relationship. If the amount of variance explained by the independent variables is inherently low, the consistency of the models in revealing the accuracy becomes important. Therefore, the determination coefficient alone should not be the single criteria. The interpretation of model consistency should also consider MAE, MAPE, MSE, and RMSE values. In our study, these values were found to be low and similar.

There is variation in the R^2 for each sex (steers and heifers) of the different machine learning models presented in Table 1 and Table 2. While the best model in steers was RF, the best model in heifers was GBR. The reason for this is that the train and test dataset in machine learning models contains a certain number of individuals. As we mentioned in the method section, to establish a model with the training data set, 70% of the total data set was used and the model was tested for the remaining 30%. For the steers and heifers, 1225 and 504 animals were used for the training data sets, while 817 and 214 animals were used for the test data sets. Therefore, the R^2 of the created model also changes in this direction.

4. Conclusions

The RF, GBR, LGBR, and LR algorithms were used to evaluate model performance in predicting pen DMI of feedlot cattle. In steers, GBR and LGBR algorithms outperformed others with a coefficient of determination of 0.45 in terms of predicting pen DMI, but our results were found lower or similar when compared with other related studies. Compared to other studies conducted on predicting DMI, the model performances in our study differed. The reason for this could be the high variation in the animals' ages, herd management, and body morphological characteristics. Furthermore, the technology such as MIRS could be essential factor in accurately prediction for DMI. In addition, it should be noted that the statistical processes performed on the data may be different for each study. The initial body weight was determined as a key feature for predicting pen DMI in both steers and heifers. Furthermore, while days on feed trait most contributed feature in steers, average proportion of dietary concentrate was significantly important to predict pen DMI in heifers. Therefore, days on feed and average proportion of dietary concentrate variables may contribute to herd management strategies at farm level. This study suggests that ensemble learning techniques can be tried to improve the model performance of DMI predictions. Increasing number of animals and using different independent variables that related to the DMI can affect the accuracy of DMI prediction.

References

Asadzadeh N, Bitaraf D E, Shams H J, Zare M, Khojestehkey S, Abbaasi S & Shafie N (2021). Body weight prediction of dromedary camels using the machine learning models. *Iranian Journal of Applied Animal Science* 11(3): 605-614

Atalay M & Çelik E (2017) Artificial intelligence and machine learning applications in big data analysis. *Mehmet Akif Ersoy Üniversitesi Sosyal Bilimler Enstitüsü Dergisi* 9(22): 155-172

Blake N E, Walker M, Plum S, Hubbart J A, Hatton J, Mata-Padrino D, Holásková I & Wilson M E (2023). Predicting dry matter intake in beef cattle. *Journal of Animal Science* 101: skad269

Bovo M, Agrusti M, Benni S, Torreggiani D & Tassinari P (2021). Random forest modelling of milk yield of dairy cows under heat stress conditions. Animals 11(5): 1305

Breiman L & Cutler A (2015). Random forest. Retrieved June 23, 2015, from https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm

Breiman L. (2001). Random forests. In: Blockeel H & Leuven K U (Eds.), Machine Learning, Scientific Research Publishing, New York, pp. 5-32

Celik S & Yılmaz O (2017). Comparison of different data mining algorithms for prediction of body weight from several morphological measurements in dogs. *Journal of Animal and Plant Sciences* 27(1): 57-64

Çelik Ş & Yılmaz O (2023). Investigation of the Relationships between Coat Colour, Sex, and Morphological Characteristics in Donkeys Using Data Mining Algorithms. *Animals* 13(14): 2366. https://doi.org/10.3390/ani13142366

Chen T, Xu J, Ying H, Chen X, Feng R, Fang X, Gao H & Wu J (2019). Prediction of extubation failure for intensive care unit patients using light gradient boosting machine. Institute of Electrical and Electronics Engineers 7: 960-968

Cutler A, Cutler D R & Stevens J R (2012). Ensemble Machine Learning: Methods and Applications. In Zhang C. & Ma Y. (Eds.), Random forests (pp. 157–175) Springer

Defalque G, Santos R, Bungenstab D, Echeverria D, Dias A & Defalque C (2024). Machine learning models for dry matter and biomass estimates on cattle grazing systems. Computers and Electronics in Agriculture 216: 108520

Di Persio L & Fraccarolo N (2023). Energy consumption forecasts by gradient boosting regression trees. Mathematics 11(5): 1068

Hastie T, Tibshirani R & Friedman J H (2009). The elements of statistical learning: Data mining, inference, and prediction (2nd ed., pp. 1-758). Springer

Hicks R B, Owens F N, Gill D R, Oltjen J W & Lake R P (1990). Daily dry matter intake by feedlot cattle: influence of breed and gender. *Journal of Animal Science* 68(1): 245-253

Hong W (2015). Wavelet Gradient Boosting Regression Method Study in Short-Term Load Forecasting. Smart Grid 5: 189–196

Huma Z E & Iqbal F (2019). Predicting the body weight of Balochi sheep using a machine learning approach. *Turkish Journal of Veterinary & Animal Sciences* 43(4): 500-506

Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Ye Q & Liu T Y (2017). Lightgbm: A highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems 30: 3146–3154

Koknaroglu H, Demircan V & Yilmaz H (2017). Effect of initial weight on beef cattle performance and profitability. Agronecio 13(1): 26-38

Koknaroglu H, Loy D D, Wilson D E, Hoffman M P & Lawrence J D (2005). Factors affecting beef cattle performance and profitability. The Professional Animal Scientist 21(4): 286-296

Koşkan O, Koknaroglu H, Loy D D & Hoffman M P (2014). Predicting dry matter intake of steers and heifers in the feedlot by using categorical and continuous variables. In: American Society of Animal Science Annual Meeting, 20 – 24 July, Kansas City, Missouri, USA, pp. 721-721

Lahart B, McParland S, Kennedy E, Boland T M, Condon T, Williams M, Galvin N, McCarthy B & Buckley F (2019). Predicting the dry matter intake of grazing dairy cows using infrared reflectance spectroscopy analysis. *Journal of dairy science* 102(10): 8907-8918

Mammadova N & Keskin I (2013). Application of the support vector machine to predict subclinical mastitis in dairy cattle. *The Scientific World Journal* 2013: 897-906

Mikail N, Keskin I & Altay Y (2014). The use of artificial neural networks and support vector machines methods in milk yield prediction of holstein cows. In: Proceedings of the International Mesopotamia Agriculture Congress, 22 – 25 September, Diyarbakir, 1137 pp

Müller A C & Guido S (2016). Introduction to Machine Learning with Python: A Guide For Data Scientists. O'Reilly Media, USA.

National Academies of Sciences, Engineering, and Medicine (NASEM) (2016). Nutrient Requirements of Beef Cattle, 8th revised edn. Washington, DC: The National Academies Press

Ogutu J O, Piepho H P & Schulz-Streeck T (2011). A comparison of random forests, boosting and support vector machines for genomic selection. *BMC* proceedings 5: 1-5

Otchere D A, Ganat T O A, Ojero J O, Tackie-Otoo B N & Taki M Y (2022). Application of gradient boosting regression model for the evaluation of feature selection techniques in improving reservoir characterisation predictions. *Journal of Petroleum Science and Engineering* 109: 244-254

PyCaret (2020). An Open Source, Low-Code Machine Learning Library in Python. Retrieved in August, 23, 2023 from https://pycaret.org/

R Core Team (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ray S (2019). A quick review of machine learning algorithms. In: International conference on machine learning, big data, cloud, and parallel computing, 14 – 16 February, Faridabad, India, pp. 35-39

Refaeilzadeh P, Tang L & Liu H (2016). Cross – Validation. Springer, New York.

Salleh S M, Danielsson R & Kronqvist C (2023). Using machine learning methods to predict dry matter intake from milk mid-infrared spectroscopy data on Swedish dairy cattle. *Journal of Dairy Research* 90(1): 5-8

Shadpour S, Chud T C, Hailemariam D, Oliveira H R, Plastow G, Stothard P, Lassen J, Baldwin R, Miglior F, Baes C F Tulpan D & Schenkel F S (2022). Predicting dry matter intake in Canadian Holstein dairy cattle using milk mid-infrared reflectance spectroscopy and other commonly available predictors via artificial neural networks. *Journal of dairy science* 105(10): 8257-8271

Sibindi R, Mwangi R W & Waititu A G (2023). A boosting ensemble learning based hybrid light gradient boosting machine and extreme gradient boosting model for predicting house prices. Engineering Reports 5(4): e12599.

Sun X, Liu M & Sima Z (2020). A novel cryptocurrency price trend forecasting model based on LightGBM. Finance Research Letters 32: 101084

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Determination of Factors Affecting "Level of Dependency on Social Aid" of Household Living in Rural Area: Iğdır Province Rural Area Example, Türkiye

Osman Doğan Bulut[a], **Cengiz Sayın**[b*]

[a]*Igdir University, Faculty of Agriculture, Department of Agricultural Economics, Igdir, TÜRKİYE*

[b]*Akdeniz University, Faculty of Agriculture, Department of Agricultural Economics, Antalya, TÜRKİYE*

ARTICLE INFO

Research Article

Corresponding Author: Cengiz Sayın, E-mail: csayin@akdeniz.edu.tr

Received: 27 May 2024 / Revised: 07 August 2024 / Accepted: 15 August 2024 / Online: 14 January 2025

Cite this article

Bulut O D, Sayın C (2025). Determination of Factors Affecting "Level of Dependency on Social Aid" of Household Living in Rural Area: Iğdır Province Rural Area Example, Türkiye. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):100-109. DOI: 10.15832/ankutbd.1490688

ABSTRACT

Fighting against poverty has become an increasing topic both at international and national levels. In this context, social policy programmes and particularly the implementation and delivery of social aid programs are one of key policy tools widely used in many countries to alleviate poverty and reduce hunger. Assessing the Aid Dependency Rate (ADR) of the beneficiaries is of great importance to achieve the goal of designed and delivered social assistance programs. Therefore, this study determines beneficiaries' level of dependency on social aid and the underlying factors. Primary data were collected through face-to-face survey from 210 households actively beneficiaries of public social aid selected by use of snowball sampling. Additionally, ordinal logistic regression analysis was conducted to determine the factors affecting the level of need of households for public social aid. The findings showed that 46.7% of households were in low level dependency on social aid, 28.6% in intermediate level, and 24.8% in high level. Besides, the results of the ordinary logistic regression analysis revealed that the marital status, employment status of the household and the group of delivered social aid were significant factors affecting level of need for social aid ($P<0.05$). Also, it shows that the odd ratio of household of being in high level increases 264.25 times if there is not any working individual compared to households with two or more working individuals ($P<0.01$) whereas this ratio decreases to 3.71 ($P<0.05$) in households by only one individual working. The study concludes that the presence of even one working individual is of great importance in order to prevent the household's dependence on social aid in high level. Consequently, designing social aid programs that consider the mentioned factors would help to fight against poverty.

Keywords: Social aid, Aid dependency rate, Ordinal logistic regression, Iğdır, Türkiye

1. Introduction

Poverty has become a serious worldwide problem that poses a huge social challenge for policymakers and international institutions. This is due that fighting against poverty is one of sustainable development goal (SGD1) of the United Nations. Determination of the causes of poverty plays an important role in assessing and fighting against poverty. Increasing poverty over the globe is correlated to unfair distribution of income between and inter the countries. Ravallion (2004) notes that income inequality hampers the efforts of poverty reduction. Furthermore, other reasons of poverty include low-income level, marital status, the size of family, low education level, rapid population growth, migration and the difference of development among regions and countries (Ajakaiye & Adeyeye 2001; Dewilde 2004; Kim et al. 2010; Nándori 2011).

On the other hand, many approaches are used to assess poverty and apprehend the dimensional aspects of poverty. While some scholars employ household income as proxy in evaluating the poverty status others use non income proxy such as consumption expenditure, and often human welfare as proxies to estimate individuals' poverty status and well-being (Sen 2000; Wagle 2002; Lister 2004; Jansen et al. 2015). Further, Hoddinott & Quisumbing (2003) highlight that the term vulnerability is often interchangeably used as poverty in economics literature and vulnerability, utility and exposure are three econometric measures used to refer to expected poverty, expected utility and risk measures in economic models.

Measuring poverty at the local level is straightforward; at the national level it is hard but manageable; and at the level of the world as a whole it is extremely difficult, so much so that some people argue that it is not worth the effort. Because there is no world political authority that can set a poverty line use it in antipoverty policies (Deaton 2006). In fact, each country designs its own social policies according to the importance of poverty alleviation in the country policy. Tabor (2002) mentions that social assistance is mostly delivered in cash or in-kind assistance whereas Haushofer & Fehr (2014) stress that the primary goal of social policy is alleviating poverty. Moreover, Midgley & Tang (2010) and Dama (2016) indicate that social assistance benefits

are either delivered by the public authorities, private companies or non-government organizations (NGOs) according to the political system, economic development and social structures of the countries.

In Türkiye, the General Directorate of Family and Social Services (2010) mentions that many small municipalities lack of well-defined method in determining adequate level of need of the households for social assistance. Commonly, the Turkish Statistical Institute (2021) defines a poor as any individual whose income is below a certain threshold and Ministry of Family and Social Services (2021) uses absolute poverty as main criteria and other criterion based on the types of social assistance programs. The neediness threshold for social assistance in Türkiye was considered as a monthly income per capita smaller than one-third (1/3) of the national minimum wage. For this reason, many studies were conducted in Türkiye to examine both the social policy and programmes in different regions as well as the determinants of the success of social assistance in combating poverty. Arı (2003) stated that self-targeting approach of the poor does not reflect the poverty status of about 70% of the applicants. Furthermore, Çetinkaya (2012) indicated that the determination of neediness is still not clear. Daşlı (2016) underscores that identifying the neediness level in implementing social assistance is challenging in Turkey because a lack of appropriate scientific methods. In this context, Taşcı (2019) indicates that the level of need for social assistance is one of widely used tool used in providing social assistance to the poor and vulnerable groups. Abdoul-Azize & Sayın (2022) added that the determination of key factors affecting the level of need of households for public social assistance contributes in designing and delivering effective social assistance benefits that would reduce their risk of being dependent (less needy, needy, very needy) to social assistance.

From the above-mentioned literature, scholars have given a particular attention to designing appropriate criteria to deliver sufficient social protection benefits to the targeting recipients as the use of income level of the beneficiaries is not consistent to apprehend the poverty status of the households. In this view, the classification of the level of dependency on social aid and the determination of factors affecting this level might be of great importance. Accordingly, this study investigates the socioeconomic characteristics and the dependency level of household beneficiaries for social aid and the factors affecting the level of dependency on social aid.

2. Material and Methods

2.1. Study area

This study was carried out in the province of Iğdır and it focused on the households who are beneficiaries of public social aid. The province of Iğdır consists of four regions namely; Iğdır Center, Tuzluca district, Karakoyunlu district and Aralık district. The population of the province of Iğdır is estimated at 199,442 inhabitants of which 56% resides in the city and district centers and 44% others reside in towns and villages.

2.2. Sampling method and data collection

A snowball sampling technique was used to select the households actively beneficiaries of public social aid. The choice of snowball sampling is due to its advantage to gather rich information that reflect a variety of situation on the research topic (Morgan & Morgan 2008; Vogt et al. 2012). Also, such sampling technique is commonly used when the access to the units of the study is difficult and the information on the population is not clear (Patton 2005). Creswell (2013) emphasizes that this technique focuses on people from whom rich data can be obtained while the population of study is reached by following the individuals of the population. Kerlinger & Lee (1999) added the data collection phase of the research is completed that when data saturation is reached.

The study data consisted of primary data collected through face-to-face survey from 210 households actively beneficiaries of public social aid. Since the population of Iğdır center accounts for approximately 70% of the total population, and the characteristics of the districts are similar, 120 households were surveyed from the city center and 30 households from each county as shown in Table 1.

Table 1- Distribution of the number of surveys by region

Region	***Number of Survey***	***Rate (%)***
Iğdır center	120	57.1
Tuzluca district	30	14.3
Aralık district	30	14.3
Karakoyunlu district	30	14.3
Total	210	100.0

2.3. Determination of household's assistance income ratio

Assistance Income Ratio (AIR) is an index indicating the level of need of households below the poverty line for social assistance. This index was developed in the doctoral thesis of Abdoul-Azize (2020) under the supervision of Prof. Dr. Cengiz Sayın. The determination of the value of the AIR index includes all monetary value of both in-kind and cash assistances received for the beneficiary households by converting all in-kind assistance benefits into current values at the market in national currency. Additionally, the monthly amount of social assistance of a given household includes all the monetary values of social assistance delivered the individuals residing in the same household and the household monthly income represents the total of income earned per month by all individuals residing in the same household. According to Abdoul-Azize (2020) the value of AIR of the household is determined by dividing the household monthly amount of social assistance by its monthly income. In this study, because this index shows how dependent the household is on social aid for its livelihood, it was interpreted as the Aid Dependency Rate depicted in Equation 1:

$$\text{ADR} = \frac{\text{Total social assistance for household}}{\text{Total income for household}} \quad (1)$$

2.4. Determination of households' level for dependency on social aid

The ADR is not only an index developed to determine the level of need of households for social aid but also to classify the households below the poverty line and recipients of social aid. Although ADR index in not an eligibility criterion for social aid, its expresses the dependency of households on social aid. Accordingly, the households were ranged into 3 different categories based on their ADR index, which are less level for dependency on social aid, intermediate level for dependency on social aid and high level for dependency on social aid (Table 2).

Table 2- Household's levels on social aid dependency by ADR

Level	***ADR***	***Description***
Low level for dependency on social aid	$0 < \text{ADR} < 1$	These values of ADR index indicates that the share of monthly amount of social aid received by the households are less than the household monthly income. Household's survival depends largely on household income.
Intermediate level for dependency on social aid	$1 \leq \text{ADR} < 2$	These values of ADR indicate that the share of the monthly amount of social aid received by the household equals at least its monthly income or more. Household's survival is largely dependent on social aid.
High level for dependency on social aid	$2 \leq \text{ADR}$	These values of ADR index indicates that the share of monthly amount of social aid of the household is at less two times the monthly income of the recipient household. Therefore, household's survival is highly dependent on social aid.

Graphically the level of need of recipient household for public social aid is presented in Figure 1. The x axis represents the monthly income and the y axis the total monthly amount of aid received by the household recipient of social aid. The level of need of the household is separated by y=x and y= 2x lines whereas the line y=x is the points where the monthly amount of aid (y) of the household equals its monthly income (x). The y=2x line represents the points where the monthly amount of social aid (y) equals two times the monthly income of the household(x). Points on the +y axis represent the households that have no income and are fully dependent on aid. Since this methods is applied to needy household that receive social aid, there is no point located on the +x.

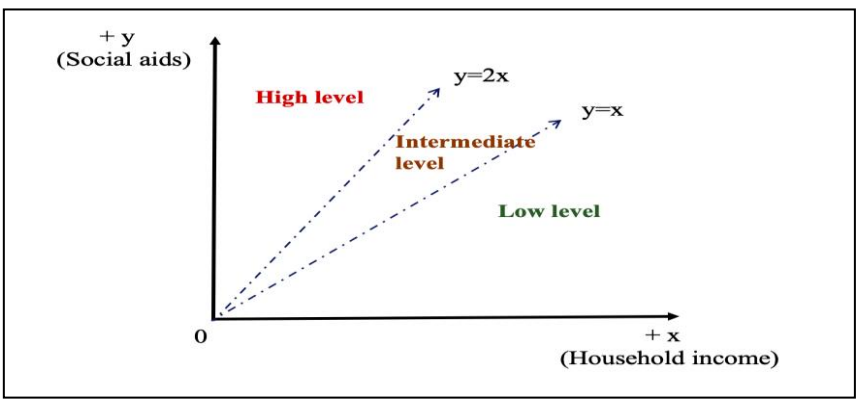


Figure 1- Household's level for dependency on social aid

2.5. Ordinal logistic regression analysis

This study uses ordinal logistic regression model to determine the factors affecting the level of need of the households for public social aid. Ordinal logistic regression model is a model which dependent variable include more than two sortable categories (Chen & Hughes 2004). Commonly it is used for a better comprehension of data and for a strong inference about the characteristics of a population. In all fields of study, ordinal scales arise when the values of continuous variables are measured or summarized by researchers by narrowing them into a set of categories.

In order to reduce subjectivity in social sciences, it is useful to guide ordinal scales about what categories represent (Agresti 2010). The application of linear and logistic regression models largely depends on the dependent variable and the provision of model assumptions. Despite the prevalence of linear regression, binary logistic regression and multivariable logistic regression techniques, ordinal logistic regression analysis is the only alternative technique in studies where the dependent variable has a clear ordering of the category levels (Klaeboe et al. 2003). The model developed by McCullagh (1980) is based on the assumption that there is an unobservable latent variable under an observable categorical variable. In this model, it is assumed that there is an unobservable latent variable (Y^*) that can take values between $-\infty$ and $+\infty$ under the observable ordered categorical dependent variable (Y). A generalized ordinal logistic regression model is used if the different values of the independent variables are higher in different categories of the dependent variable. This regression model is valid when the ordinal dependent variable has three or more categories (McCullagh 1980; Ishwaran & Gatsonis 2000). Five basic link functions are used to obtain the ordinal logistic regression model. The most used functions; logit, probit and cloglog functions (Long 1997). In the ordinal logistic regression model, there is often no clear consideration which link function to choose. In this study, the logit connection function is used for the model.

2.6. Structure of ordinal logistic regression model

The effects of some selected variables on level of dependency on social aid were tested by ordinal logistic regression (OLR) model, structure of which is presented in Figure 2. Independent variables in the model are age, gender, marital status, working status, social assistance group, residential region and residential area kind.

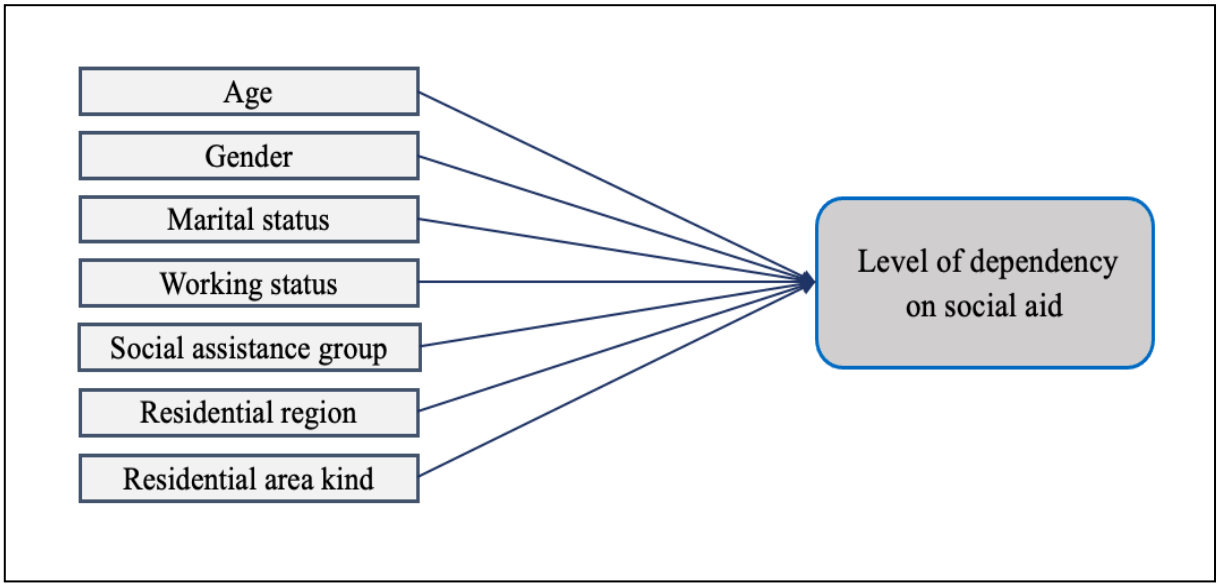


Figure 2- Structure of ordinal logistic regression model used in the study

2.7. Variables used in the ordinal logistic regression model

Dependent and explanatory variables used in the ordinal logistic regression model and their explanations are shown in Table 3.

Table 3- Description of variables in the ordinal logistic regression model

Variable	*Type*	*Explanation*
A. Dependent variable		
·Level of dependency on social aid	Dummy	1: Low 2: Intermediate 3: High*
B. Explanatory variable		
·Age of the head of the household	Continuous	
·Gender of the head of the household	Dummy	1: Male 2: Female*
·Marital status of the head of the household	Dummy	1: Married 2: Single or widowed*
·Working status of individuals residing in the household	Dummy	1: No working person 2: One working person 3: Two or more working people*
·Group of social assistance	Dummy	1: Disabled, elderly or home care payment 2: Family pension for children's education/health or widow payment 3: One-time aid (food, pandemic or home renovation)*
·Residential area kind	Dummy	1: Rural 2: Urban*
·Residential region	Dummy	1: Iğdır center 2: Tuzluca county 3: Aralık county 4: Karakoyunlu county*

*: Reference category

3. Results and Discussion

3.1. Socioeconomic characteristic of the households

Within the scope of the research, the survey was conducted with head of households receiving at least one social assistance (Table 4). Most households are headed by men (81.0%) aged is 48.6 years averagely. Also, 38.6% of them were illiterate, most of them (57.1%) had primary education. The most common type of social assistance (45.7%) is family pensions, which include children's education/health assistance or widow assistances, flowed by disabled, elderly or home care pension (39.0%) and lastly one-time aid (15.2%), consisting of food, pandemic or home renovation. While 52.4% of the needy individuals live in rural areas, 47.6% live in urban areas.

Table 4- Important statistics of respondents

Variables	*Categories*	*Frequency*	*Rate (%)*
Head of the household			
Gender	Male	170	81.0
	Female	40	19.0
Age (Mean: 47.7 years)	≤36	45	21.4
	37-52	87	41.4
	53≤	78	37.2
Marital status	Single (including divorced people)	45	21.4
	Married	165	78.6
Education level	Illiterate	81	38.6
	Primary school	120	57.1
	Secondary school	8	3.8
	University	1	0.5
Individuals of the household			
Working status of household (Mean: 0.7 person)	None	76	36.2
	One person	97	46.2
	Two or more people	37	17.6
Group of social assistance of household	Disabled, elderly or home care pension	82	39.0
	Family pension for children's education/health or widows	96	45.7
	One-time aid (food, pandemic or home renovation)	32	15.2
Type of residential areas	Rural	110	52.4
	Urban	100	47.6

3.2. Level of dependency on social aid

According to the research data most beneficiaries (46.7%) of public social assistance in the province of Iğdır were in low level of dependency on social aid whereas 28.6% in intermediate level and 24.8% in high level. (Table 5). Zengin et al. (2012) and Ikizoğlu (2002) state that priority should be given to those who need assistance the most by taking into account the degree of need in order to avoid the negative effects of social assistance. Standard assistance programs that cover large segments of society and prevent the effective use of resources should be avoided.

Table 5- Level of dependency on social aid by ADR

ADR (Mean)	*Level of dependency on social aid*	*Frequency*	*Rate (%)*
0.27	Low	98	46.7
1.25	Intermediate	60	28.6
3.17	High	52	24.8

3.3. Factors affecting the level of dependency on social aid

In the study, ordinal logistic regression analysis was used to determine the factors affecting the level of dependency on social aid. Some assumptions such as test of parallel lines and multicollinearity were tested and Pseudo R square values was determined to examine the model fitting. Parallel regression assumption or the proportional odds assumption is a necessity while using ordinal logistic regression model for an ordered categorical variable unless a multinomial model was preferred. P value greater than 0.05 implies the failure of rejecting the null hypothesis underlining the parallel regression assumption holds (Liang et al. 2020). In this analysis, the Pearson chi-square test [$\chi^2(11)= 19.267$, $p=0.056$] was non-significant which satisfies the held assumption.

Model fit was assessed by comparing the −2 log likelihood for the intercept-only model and the full model, and Chi-square statistic was used to examine the significance of the model with p values less than 0.05 (Petrucci 2009). In this analysis, it has been seen a significant improvement in fit of the final model over the null model [$\chi^2(11)=204.272$, $P<0.001$]. This situation reveals the existence of the relationship between the dependent variable and the independent variables.

The goodness of fit of the model was examined through R^2. The R^2 values shows what percentage of the dependent variable is explained by the independent variables. But there is no strong guidance in the literature on how these should be used or interpreted (Lomax & Hahs-Vaugn 2012; Smith & McKenna 2013; Osborne 2015; Pituch & Stevens 2016). Pseudo R square values are used by some to assess model fit by determining the effect size of the model. For this analysis, Pseudo R square statistics were as follows: Cox and Snell: 0,622; Nagelkerke: 0,707; McFadden: 0,459.

To examine a multicollinearity problem between independent variables of the ordinal logistic regression model, tolerance and variance inflation factors (VIF) were determined. Hair et al. (1995) note that a VIF value of 10 is a maximum threshold in examining multicollinearity and Ringle et al. (2015) indicate that 5 is a threshold value of VIF while examining multicollinearity issues. The results of multicollinearity estimation of the model used in this study is shown in Table 6.

Table 6- Variance inflation factor results

Variables	*Tolerance*	*VIF*
Age	0.861	1.162
Gender	0.574	1.743
Marital status	0.621	1.611
Working status	0.861	1.161
Social assistance group	0.793	1.262
Residential area kind	0.970	1.031
Residential region	0.955	1.047
VIF average		1.288

The estimated values of the odds ratio were determined to interpret the parameters of the ordinal logistic regression analysis and to determine the reference categories. The interpretations were conducted with the odds ratio of the determined reference categories. The reference category helps in interpreting other categories according to one of the categories of a variable. This method of examining the significance of the parameter refers to odds ratio interpretation. Accordingly, if the odds value is greater than 1, the resulting increase rate is mentioned, and if the odds value is less than 1, the resulting decrease rate is mentioned (Clayton 1974; McCullagh 1980; Garson 2012; Koletsi & Pandis 2018)

Table 7 presents the ordinal logistic regression results. In the model, the third category (High level) of the dependent was determined as the reference category. According to the significance levels of the model parameters; marital status of head of households, working status of households and assistance group of households were found to have a significant effect on level of need for public social assistance (P<0.05). Of the independent variables in the model; age of head of households, gender of head of households, residential area status of households and residential region of households were found non-significant. This finding is not entirely in alignment with previous studies. Barros et al. (1997), Maitra (2002), Demissie et al. (2017), Mdluli-Maziya & Dunga (2022) and Abdoul-Azize & Sayın (2022) found age of the head of the household affected significantly the level of need of the household for social assistance. Barros et al. (1997) and Maitra (2002) found that gender to be an important factor of poverty and they argued female headed households are more likely to be poor compared to male headed households. Boxill & Quarless (2005) found that the poverty is mostly experienced by individuals living in rural areas. Although similar relation was found in study model, it was determined the effect of residential area kind on level of need for public social assistance is non-significant. Residential region has not significant effect on level of need for public social assistance. It was deduced this is associated with the similarity of the development characteristics of the regions in Iğdır.

Table 7- Results of the ordinal logistic regression model

Variable	***Estimate (β)***	***Std. Error***	***Wald***	***Sig. (p)***	***Odds Ratio (e^{β})***
Age	0,020	0.015	1.837	0.175	1.020
Gender (1)	0.235	0.497	0.225	0.636	1.265
Gender (2) *Reference*	0[a]				
Marital status (1)*	-1.544	0.524	8.681	0.003	0.213
Marital status (2) *Reference*	0[a]				
Working status (1)*	5.577	0.768	52.749	0.000	264.256
Working status (2) **	1.311	0.617	4.518	0.034	3.709
Working status (3) *Reference*	0[a]				
Assistance group (1)*	1.683	0.589	8.153	0.004	5.381
Assistance group (2)	-0.094	0.581	0.026	0.871	0.910
Assistance group (3) *Reference*	0[a]				
Residential area kind (1)	0.211	0.348	0.367	0.545	1.235
Residential area kind (2) *Reference*	0[a]				
Residential region (1)	0.257	0.527	0.238	0.625	1.293
Residential region (2)	-0.238	0.668	0.127	0.722	0.788
Residential region (3)	0.297	0.666	0.199	0.656	1.346
Residential region (4) *Reference*	0[a]				

Significance level: * P<0.01; ** P<0.05. Detailed description of variables was given Table 3

The reference category of marital status variable was determined category 2, which is single or divorced. Households with a married head of household are 0.21 times less likely to be in the high category compared to the other (Wald $\chi 2 = 8.681$, P<0.01). This finding is entirely in alignment with previous studies. Sigle-Rushton & McLanahan (2002), Hoddinott & Quisumbing (2003) and Biyase & Zwane (2018) stated that unmarried heads of households have lower income compared to those who are counterparts (the divorced, single and widowed). In parallel with these findings, a study conducted by Ekinci Hamamcı & Anık (2020) showed that most of the needy women who divorced in Türkiye reported that they ended their marriages due to financial difficulties, and the Family Structure Survey conducted by the Turkish Statistical Institute (2017) reported that 42.6% of the reasons for divorce were the inability to financially support the family.

The reference category of working status was considered as category 3 and include households with two or more working individuals. The households that did not have any working individual was considered category 1. For this group, the odds ratio of the household of being in high level for dependency on social aid increases 264.25 times compared to household within the

category 3 (Wald $\chi^2 = 52.749$, $P<0.01$). If only one individual living in the household works, the odds ratio value would decrease to 3.71 (Wald $\chi2 = 4.518$, $P<0.05$). This implies that one individual working in a household represent a potential to fight against the household's poverty. Similarly, Sekhampu (2013) found that the employment status is a key factor determining the probability of the household of being poor and Coulombe & Mckay (1996) and Grootaert (1997) discovered that economic factors such as employment have an important role in determining the poverty status of the household. Further, Çağlayan and Dayıoğlu (2011) emphasized that the number of employees in the household determine the poverty status of the household whereas Javed & Asif (2011) stated that the employment status of the individuals determines their income. Considering the relationship between employment status and household income, the study of Abdoul-Azize & Sayın (2022) found that an increase in the monthly income of the households would likely reduce its level of need for public social assistance.

The reference category of the group of social assistance was determined as the category 3. This category includes one-time payment, food assistance, pandemic related assistance, and home improvement payment. Households in category 1, include regularly paid assistance such as old-age poor assistance, home care assistance, were more likely to be in high level for dependency on social aid 5.38 times compared to those in category 3 (Wald $\chi2 = 8.153$, $P<0.01$). This implies that unproductive individuals such as elders and disabilities are the most disadvantaged social group.

4. Conclusions

The use of determining level of dependency on public social aid as a focal point has revealed highly interesting results regarding the profiles of households receiving social assistance. In this study, the method that reveals the level of dependency on public social aid among the households receiving social assistance was applied and the effective factors were determined by ordinal logistic regression model.

The descriptive statistics results showed that most households were in low decency level for social aid. Additionally, the results of ordinal logistic regression analysis revealed that the working status of the individuals living in the recipient households, the marital status of the head of the household, and the group of social assistance were significant predictors of the level of dependency on social aid ($P<0.05$).

The researchers in this study recommend that social aid should be provided to those who are most dependent on social aid, and that the amount of aid should increase as dependency increases. In other words, standard social aids should be avoided. Besides, it is of great importance and priority that at least one individual is participated to working life by the help of the public so that the households receiving aid do not fall into the trap of poverty. Also, the fact that individuals who receive regularly paid aids are in a disadvantageous position in working life results in being more dependent on social assistance. Therefore, it increases the share of assistances in total household income. The employment of these individuals in the public or private sectors should be subject to positive discrimination by the public. Besides, steps towards ensuring family integrity through direct or indirect policies will contribute positively to the combating poverty by ensuring family unity.

Future studies should be conducted with beneficiaries by identifying the level of dependency on social aid. A comparison of the profiles of these beneficiaries will contribute to the literature. Furthermore, similar studies must be conducted in provinces with different development levels and in provinces hosting foreign refugees.

Acknowledgments

This study is quoted from a part of the PhD Thesis of Osman Doğan Bulut entitled: Sosyal Yardım Politikalarının İşleyiş Sistemi ve Etkilerinin Değerlendirilmesi: Iğdır İli Kırsal Alan Örneği. This PhD thesis was supervised of Prof. Dr. Cengiz Sayın, Department of Agricultural Economics, Akdeniz University.

Compliance with ethical standards

This research was conducted with the approval of the Ethics Committee of Scientific Research and Publication Ethics Board Presidency, Igdir University, Türkiye.

Conflict of interest

The authors declare that there is no conflict of interest concerning the publication of this article.

Funding

The authors declare that the study received no funding.

References

Abdoul-Azize H T (2020). Investigating the social assistance policy practices and factors affecting the household benefit levels in rural areas: case of the province of Antalya. Natural and Applied Sciences Institute. Phd Thesis (Advisor: Prof. Dr. Cengiz Sayın). Antalya. https://tez.yok.gov.tr/UlusalTezMerkezi/tezDetay.jsp?id=o7zjETrgkgRBEwuWC8yTAQ&no=0SL5Hn_tq--bH5BSkA1fMQ (In Turkish)

Abdoul-Azize H T & Sayın C (2022). Determining the factors affecting the levels of need for public social assistance of households: insights from the district of Konyaaltı, Antalya - Türkiye. *Middle Black Sea Journal of Communication Studies* 7(2): 225-240. https://doi.org/10.56202/mbsjcs.1109905

Agresti A (2010.) Analysis of ordinal categorical data. Second edition. University of Florida. A John Wiley & Sons, Inc., Publication

Ajakaiye D O & Adeyeye V A (2001). Concepts, measurement and causes of poverty. CBN Economic and Financial Review 39(4): 8-44

Arı A (2003). A local friend in the fight against poverty: Friendly hand. Poverty Symposium Book 3:160- 167. Deniz feneri yayınları: İstanbul (In Turkish)

Barros R, Fox L & Mendonca R (1997). Female-headed households, poverty, and the welfare of children in urban Brazil. Economic Development and Cultural Change 45(2): 231–257

Biyase M & Zwane T (2018). An empirical analysis of the determinants of poverty and household welfare in South Africa. *The Journal of Developing Areas* 52(1): 115–130. https://doi.org/10.1353/jda.2018.0008

Boxill I & Quarless R (2005). The determinants of poverty among the youth of the Caribbean. Social and Economic Studies 54(1): 129–160

Çağlayan E & Dayıoğlu T (2011). Comparing the parametric and semiparametric logit models: household poverty in Turkey. *International Journal of Economics and Finance* 3(5): 197-207

Çetinkaya Ş (2012). Social assistance organization in Turkey: Situation analysis, problems and solution proposals. Dumlupınar University. Social Sciences Institute. Phd Thesis (Advisor: Assoc. Prof. Dr. Ramazan Kılıç). Kütahya. https://tez.yok.gov.tr/UlusalTezMerkezi/tezDetay.jsp?id=FZ4McqppUgM0z59anqdZow&no=jxthwoWkRlq0SSqZtP0fqg (n Turkish)

Chen C K & Hughes J J (2004). Using ordinal regression model to analyze student satisfaction questionnaires. IR Applications 1-13

Clayton D G (1974). Some odds ratio statistics for the analysis of ordered categorical data. Biometrika 61: 525-531

Coulombe H & McKay A (1996). Modeling determinants of poverty in mauritania. World development 24(6): 1015-1031

Creswell J W (2013). Research design: Qualitative, quantitative, and mixed methods approaches. New York: Sage

Dama N (2016). The effect of social aid on social welfare in turkey. Yıldırım Beyazıt University, Social Sciences Institute. Phd Thesis (Advisor: Prof. Dr. Erdal Tanas Karagöl). Ankara. http://acikerisim.ybu.edu.tr:8080/xmlui/handle/123456789/1594 (In Turkish)

Daşlı Y (2016). Public social welfare as a reason of the poverty: Example of central district of Sivas. Social Sciences Institute. Phd Thesis (Advisor: Assist. Prof. Dr. Erem Sarıkoca). Erzirum. https://tez.yok.gov.tr/UlusalTezMerkezi/tezDetay.jsp?id=yu6iCIDLAUVpnvIYYdj9Dw&no=Drx19oqhE9ZPT-PD5V2cug (In Turkish)

Deaton A (2006). Measuring poverty. Understanding poverty pp. 3-15 Oxford University Press

Demissie B, S Kasie & T Asmmav (2017). "Rural Households' Vulnerability to Poverty in Ethiopia." *Journal of Poverty* 21(6): 528-542

Dewilde C (2004). The multidimensional measurement of poverty in Belgium and Britain: A categorical approach. Social Indicators Research 68: 331–369

Ekinci Hamamcı E D & Anık K (2020). The role of social assistance in women's struggle against poverty: An application in Erzurum. Erzurum Technical University Social Sciences Institute Journal (10): 105-131 (In Turkish)

Garson D G (2012). Ordinal regression. Asheboro: Statistical associates publishing

General Directorate of Family and Social Services (2010). Perception of social assistance and culture of poverty. Ankara. (In Turkish)

Grootaert C (1997). The determinants of poverty in Cote d'Ivoire in the 1980s. *Journal of African Economies* 6(2): 169-196

Hair J F, Jr Anderson R E, Tatham R L & Black W C (1995). Multivariate data analysis. New York: Macmillan

Haushofer J & Fehr E (2014). On the Psychology of Poverty. Science 344(6186): 862-867

Hoddinott J & Quisumbing A (2003). Methods for microeconometric risk and vulnerability assessments. Social Protection Discussion Paper No:0324. Washington, DC: World Bank

Ikizoğlu M (2002). Relationship between poverty and social assistance: An empirical research in Ankara: Mamak district. Society and Social Service 13(1): 86-115 (In Turkish)

Ishwaran H & Gatsonis CA (2000). A general class of hierarchical ordinal regression models with applications to correlated ROC analysis. The Canadian

Jansen A, Moses M, Mujuta S & Yu D (2015). Measurements and determinants of multifaceted poverty in South Africa. Development Southern Africa 32(2):151–169

Javed Z H & Asif A 2011. Female households and poverty: A case study of Faisalabad district. *International Journal of Peace and Development Studies* 2(2): 37-44

Kerlinger F N & Lee H B (1999). Foundations of behavioral research. New York: Harcourt College Publishers

Kim K, Lee Y & Lee Y (2010). A multilevel analysis of factors related to poverty in Welfare States. Social Indicators Research 99: 391–404

Klaeboe R, Turunen Harvik L & Madshus C (2003). Vibration In dwellings from road and rail traffic part II: Exposure– effect relationships based on ordinal logit and logistic regression models. Applied Acoustics 64:89-109

Koletsi D & Pandis N (2018). Ordinal logistic regression, *American Journal of Orthodontics and Dentofacial Orthopedics* 153(1): 157-158

Liang J, Bi G & Zhan C (2020). Multinomial and ordinal logistic regression analyses with multi-categorical variables using R. Annals of translational medicine 8:16

Lister R (2004). Poverty. Cambridge, UK: Polity Press

Lomax R G & Hahs-Vaughn (2012). An introduction to statistical concepts. New York: Routledge

Long J S (1997). Regression models for categorical and dependent variables. London

Maitra P (2002). The effect of household characteristics on poverty and living standards in south Africa. *Journal of Economic Development*, 27(1): 75–83

McCullagh P (1980). Regression models for ordinal data. journal of the royal statistical society: Series B (Methodological), 42(2):109-127

Mdluli-Maziya P & Dunga S (2022). Determinants of poverty in south africa using the 2018 general household survey data. *Journal of poverty* 26(3): 197–213

Midgley J & Tang K (2010). The role of social security in poverty alleviation: an international review, social policy and poverty in East Asia: the role of social security. Routledge. London-New York

Ministry of Family and Social Services (2021). Activity report. https://www.aile.gov.tr/media/100242/2021-yili-faaliyet-raporu.pdf (In Turkish)
Morgan D L & Morgan R K (2008). Single-Case research methods for the behavioral and health sciences. Sage Publications
Nándori E (2011). Subjective poverty and its relation to objective poverty concepts in Hungary. Social indicators research 102: 537–556.
Osborne J W (2015). Best practices in logistic regression. Los Angeles: Sage Publications
Patton M Q (2005). Qualitative research. New York: John Wiley & Sons Inc. Publication
Petrucci C J (2009). A primer for social worker researchers on how to conduct a multinomial logistic regression. *Journal of social service research* 35(2): 193-205
Pituch K A & Stevens J A (2016). Applied multivariate statistics for the social sciences (6th ed). New York: Routledge
Ravallion M (2004). Pro-Poor growth: A primer. policy research working paper; No.3242. World Bank, Washington, D.C. World Bank
Ringle C M, Wende S & Becker J M (2015). SmartPLS 3. Bönningstedt: SmartPLS. Sage Publications
Sekhampu T J (2013). Determinants of poverty in a South African Township. *Journal of Social Science* 34(2): 145–153
Sen A (2000). Social exclusion: concept, application and scrutiny. Social Development Paper No: 01. Manila: Asian Development Bank
Sigle-Rushton W & McLanahan S (2002). For richer or poorer? Marriage as an anti-poverty strategy in the United States. Population 57(3): 509–526
Smith T J & McKenna C M (2013). A comparison of logistic regression pseudo R2. Multiple Linear Regression Viewpoints 39: 17-26
Tabor S R (2002). Assisting the poor with cash: Design and implementation of social transfer programs. World Bank Social Protection Discussion Paper 223: 79-97
Taşcı F (2019). Poverty and Social Work Lecture Notes. Faculty of Open and Distance Education, Istanbul University. İstanbul. (In Turkish)
Turkish Statistical Institute (2017). Family Structure Survey 2016. http://www.tuik.gov.tr/PreHaberBultenleri.do;jsessionid=4l2FZ1kVpNL8JhhkfQtRLqmnJtJ6xk6sGXGwbChqTR9N0LY1K1bp!1813643467?id=21869 (In Turkish)
Turkish Statistical Institute (2021). Income and Living Conditions Survey. https://data.tuik.gov.tr/Bulten/Index?p=Gelir-ve-Yasam-Kosullari-Arastirmasi-2021-45581 (In Turkish)
Vogt W P, Gardner D C & Haeffele L M (2012). When to use what research design. New York: Guilford Press.
Wagle U (2002). Rethinking poverty: Definition and measurement. *International Social Science Journal* 54(171): 155–165
Zengin E, Ayhan Ş & Salih Ö (2012). Social assistance practices in Turkey. Management and Economics 19(2): 133-142 (In Turkish)

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Farmers' desire to make changes in their agricultural branches in the first wave of COVID-19 pandemic restrictions: The example of Türkiye

Celal Cevher[a], Yener Ataseven[b*], Sule Coskun-Cevher[c]

[a] *Field Crops Central Research Institute, TÜRKİYE*

[b] *Ankara University, Faculty of Agriculture, The Department of Agricultural Economics, TÜRKİYE*

[c] *Gazi University, Faculty of Science, The Department of Biology, TÜRKİYE*

ARTICLE INFO

Research Article

Corresponding Author: Yener Ataseven, E-mail: yenerataseven@hotmail.com

Received: 07 November 2023 / Revised: 29 July 2024 / Accepted: 15 August 2024 / Online: 14 January 2025

Cite this article

Cevher C, Ataseven Y, Coskun-Cevher Ş (2025). Farmers' desire to make changes in their agricultural branches in the first wave of COVID-19 pandemic restrictions: The example of Türkiye. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):110-125. DOI: 10.15832/ankutbd.1387490

ABSTRACT

The coronavirus pandemic and subsequent protectionary lockdowns have negatively impacted farmers, especially those producing perishable agricultural products worldwide and Türkiye. For this reason, most researchers began to investigate the effects of restrictions on agrarian branches during the sudden shocks of the pandemic. This study was carried out to reveal the farmers' desire for change in agricultural production branches during the period when coronavirus pandemic restrictions were implemented in Türkiye. A survey study was conducted with broad participation and telephone interviews involving 2125 farmers in different production branches in 22 provinces. Nonlinear Canonical Correlation Analysis was used to analyze variables. According to our analysis results, compared with other agricultural branches, it has been determined that the farmers who do vegetable farming, fruit farming, and livestock farming have a higher desire to change their bare agrarian branches. Approximately half of the farmers interviewed reported difficulties obtaining seeds, fertilizers, and chemicals and providing sufficient labor. Due to these difficulties, 16.5% of the farmers stated that if the first shock effect of the pandemic restrictions continues and this shock effect continues, they will abandon the current main agricultural branches of animal husbandry and vegetable growing. Fruit growing and switch to other farming branches that require less input and labor for agricultural production. For these reasons, policymakers should invest more in market-oriented strategies such as input supply, storage of products, logistics, and processing of manufactured products to maintain the supply chain during pandemic periods. Because these strategies require high costs and the necessary knowledge of post-harvest operations, they are investments that individual farmers cannot afford. Our study revealed that the initial shock effect of the pandemic restrictions in Türkiye had a limited impact on the production of agricultural products. Despite this little impact, 87.3% of farmers stated that their income decreased slightly during this period. We foresee that the restrictions made due to the coronavirus pandemic will affect the planning and social policy in the Turkish economy in the coming years. However, this situation will not change the basic structure of Turkish agricultural production and distribution.

Keywords: Farming systems, Farmer behavior, COVID-19 lockdown impact, Türkiye

1. Introduction

At the beginning of the coronavirus pandemic (COVID-19), each nation sought to respond to the pandemic threat with its own cultural, political, and institutional norms (Rahman & Das 2021). While the restrictions imposed immediately after the COVID-19 pandemic affected every sector of the economy, the magnitude of these effects varied across industries. These measures significantly affected global food security, people's lives, and economic resources, particularly in low-income countries (Cevher et al. 2021; Husse et al. 2021; Obayelu et al. 2021; Varshney et al. 2021). Unlike other outbreaks, the negative impacts of the COVID-19 pandemic on agricultural production, food trade, and food security were felt more (Ceballos et al. 2020; Pu & Zhong 2020; Ali et al. 2021; Ding et al. 2021; Ker & Biden 2021; Kumaran et al. 2021). Periodic and long-term quarantines imposed during the COVID-19 pandemic have affected farmers more, mainly due to transportation restrictions and the lack of buyers in the market (Stephens et al. 2020; Ali et al. 2021; Mishra et al. 2021; Worku & Ülkü 2021). Due to the COVID-19 restrictions during this period, the effects on all branches of agricultural production in the countries were extraordinary (Orden 2021).

Agricultural policy experts worldwide have reported that the COVID-19 pandemic is putting tremendous pressure on agricultural production (FAO 2020; OECD 2020). On the other hand, many countries have made quick decisions to prevent COVID-19 restrictions, and thanks to these decisions, a significant food shortage has yet to emerge in the world (Laborde et al. 2020). In the early stages of the COVID-19 pandemic, both supply-side and demand-side shocks initially disrupted agrifood markets. However, significant adjustments by farmers, processors, distributors, and governments have kept them relatively short-

lived (Orden 2021). All the above-mentioned situations reveal that farmers need to rethink many factors, such as increasing the resilience of production branches in their farm systems, the labor required for farm practices and the marketing possibilities of harvested products (Meuwissen et al. 2019).

Due to its geographical structure, Türkiye has a production pattern that will meet all kinds of food needs; therefore, there is a need for labor supply and agricultural input in every season. Agricultural structures are more sensitive to sudden pandemics and adverse environmental conditions. Thus, the necessity of investigating the effects of COVID-19 on the Turkish agricultural economy has emerged. Although there are many studies on the impact of the COVID-19 pandemic on agriculture in Türkiye at the time of writing the article (Akın et al. 2020; Akbudak & Şen 2021; Gürbüz & Özkan 2021), there will be hardly any field studies at the farmer level. This shows that it is necessary to prioritize field studies on the impact of COVID-19 restrictions on agricultural production. In this respect, our study is one of the essential field studies that examines and demonstrates the effects of COVID-19 pandemic restrictions on agricultural production branches. We believe the results obtained from Türkiye will set an example for similar studies, especially in developing countries, and will guide future research in making general assumptions.

These studies will assist policymakers in formulating effective policies and provide data on future efforts to prevent similar communicable diseases. For this reason, researchers have reported that due attention should be given to academic studies, primarily covering the effects of the pandemic on global and local agricultural production branches, the social and economic uncertainties caused by the pandemic, and the effects of the outbreak on the continuity of all agricultural branches (Savary et al. 2020; Yoshida & Yagi 2021).

Türkiye is one of the few countries that are self-sufficient (especially in producing cereals, fruits, and vegetables), except for a few products in agriculture (soy, corn, sunflower, animal feed, and red meat). Nonetheless, in the face of unexpected and exceptional circumstances like the pandemic, the necessity of introducing new solutions in agricultural production and developing related strategies has emerged. Despite the negativities during the COVID-19 pandemic in Türkiye, agricultural production in the country has been at a level that meets the needs of the country's population. However, Türkiye remains susceptible to food price hikes resulting from the adverse impacts of the coronavirus. Türkiye is an essential agricultural country where the pandemic started late and is managed better than in other countries. Our study comprehensively addresses all aspects of agricultural activities affected by COVID-19, striving to provide a holistic understanding of its overall impact. This study aims to investigate the restrictions applied immediately after the COVID-19 pandemic and the trends in farmers' transition from existing agricultural production branches to other agricultural production branches.

2. Material and Methods

2.1 Characteristics of Turkish agriculture

Türkiye has 23.1 million hectares of utilized agricultural area. Agriculture in Türkiye contributes 4.8% to the Gross Domestic Product (GDP) (TurkStat 2021) and is the source of livelihood and employment for approximately 8.8% of the country's population. Over half of agricultural production consists of field crops, followed by vegetable and fruit production. Farm crops are grains, legumes, tuber plants, oilseeds, tobacco, plants used in textiles, medicinal plants, and forage plants. Türkiye holds the top global position in vegetable production and is among the top five countries producing many vegetable types. Due to the geographical potential of Türkiye, animal husbandry is an essential source of livelihood. Although animal husbandry is quite suitable for these lands in terms of ecological and socioeconomic conditions, it takes place as a secondary field of activity besides plant production. Despite the changes in the world, the agricultural sector in Türkiye has grown continuously, except for 2007 (TurkStat 2021). It holds the top position in Europe and ranks seventh worldwide in terms of the scale of its agricultural economy. However, structural problems in the agricultural sector continue in Türkiye.

2.2 Study population and sample

The population of this study consists of the provinces where the highest production is made in different agricultural branches. As seen in figure 1, these provinces represent different branches of agricultural production in Türkiye. According to the 2019 Farmer Registration System, the number of farmers in Türkiye is 2 172 000 (Anonymous 2022). For this research, 2 125 farmers were selected randomly from 22 different cities using a multi-stage cluster sampling method (Balcı 1997). In this study, the provinces with the highest production in each agricultural branch were selected, and these provinces and the number of surveys conducted are as follows, respectively: Adana (195), Afyonkarahisar (52), Ankara (167), Antalya (147), Aydın (83), Bursa (145). Edirne (95), Erzurum (99), Gaziantep (27), Mersin (62), İzmir (73), Kastamonu (72), Konya (209), Malatya (48), Manisa (143), Mardin (39), Ordu (61), Sakarya (132), Samsun (102), Trabzon (63), Şanlıurfa (55) and Van (20). Once the sample size was determined, we identified the first responder farmers by consulting key information sources with whom the authors have worked for several years and provided their telephone numbers. An interview guide was prepared with five farmers, and a pre-test was done with the prepared questionnaire. Survey data was provided by interviews with farmers via mobile phone during the May-June 2020 period, when the first shock effect of the pandemic restrictions was experienced. We aim to measure in which production branches the farmer's behavior is felt more during this period due to the pandemic restrictions and determine which

production branches the farmers want to change. Individual interviews with the farmers who wanted to participate in the survey lasted an average of 25 minutes.

Researchers have formal training in applied social sciences and many years of experience conducting mixed-methods research. Therefore, in the first data collection stage, we consciously identified our initial participants (and obtained their phone numbers) by consulting key informants with whom the researchers had worked for many years. Later, with the farmers interviewed, the purpose of the research was clearly explained to the participants, the institutions where the researchers worked were informed, it was explained that the study was necessary for the continuity of agricultural production, and most importantly, after establishing trust between the farmers and the researchers as a result of the experiences gained by the researchers in the project experience they had carried out in 20-25 years, the survey was conducted. Data collection has started. Therefore, participants trusted this interview style, and no participant refused to participate in the survey (Cevher & Altunkaynak 2020). Detailed notes were taken both during and after all the interviews. The area where the survey was conducted is shown in Figure 1.

Figure 1- Map of the study area

2.3 Data and survey

Once the subjects were identified, questionnaire forms were prepared following the purpose of the research. The final version of the questionnaire was reviewed by expert researchers on agribusiness and agricultural economics. Sociologists and psychologists also provided valuable input in the questionnaire preparation process.

The questionnaires were crafted with support from a multidisciplinary team, which included sociologists, psychologists, agricultural economists, subject matter experts, academics, and researchers. The questionnaire consists of three main headings: individual characteristics of farmers (demographic characteristics of the farmer- age, education, marital status, farmer's residence, non-farm income), the infrastructure of the farm (land width, number of animals), and the effects of the COVID-19 pandemic on agricultural production branches (cereal farming, livestock farming, cereal and livestock, vegetable farming, fruit farming, mixed farming) the semi-standardized questionnaire combined short open-ended questions, multiple-choice questions, and Likert-style responses.

The information obtained from the research area and the analyses based on this information cover the agricultural production season data for 2019-2020. The variable list and optimal scaling levels are shown in Table 1 below.

Table 1- Variable list and optimal scaling levels (n= 2 125)

Optimal Scaling Name and Level	*Variable Categories*
Individual features of the farmers	
Ages	1= ≤ 30, 2= 31-40, 3= 41-50, 4= ≥ 51
Educational Status (Ordinal)	1=Primary School, 2=Middle school, 3=High School, 4=University
Gender (Nominal)	1=Male, 2=Female
Non-Farm Income (Nominal)	1=Yes, 2=None
Farmer's Annual Income (Ordinal)	1=Low (15.000 $), 2= Middle (43.000 $), 3=High (≥ 43.000 $)
Place of Residence (Nominal)	1=Rural, 2=Urban
Farm Infrastructure	
Land width (Hectare) (Ordinal)	1= ≤ 6.0, 2= 6.1-15.0, 3=15.1-25.0, 4= ≥ 26.0
Number of Cattle (Ordinal)	1= ≤ 5, 2= 6-10, 3= 11-20, 4= ≥21
Number of Ovine (Ordinal)	1= ≤ 50, 2= 51-100, 3= 101-150, 4= ≥151
Agricultural Production Branch	
Agricultural Production Branch (Ordinal) (Which best describes your farming system)	1=Cereal Farming, 2=Livestock Farming, 3=Cereal and Livestock, 4 =Vegetable Farming, 5=Fruit Farming, 6=Mixed Farming
Change in Agricultural Production Branch (Ordinal)	1= I'm thinking, 2= Undecided, 3= I don't think

All of the variables mentioned above are considered as the variables that directly or indirectly affect the change of agricultural production branches by the farmers during the restrictions imposed during the COVID-19 pandemic. These data were collected by a survey form prepared by the researchers in line with the existing literature (Uğur & Buruklar 2020; Ullah et al. 2021).

Agricultural Production Branch: the primary production branch constitutes more than 50% of the annual income and is the farmer's leading activity indicator in the agricultural farm. Accordingly, agricultural production branches are cereal farming, listed as livestock farming, cereal and livestock, vegetable farming, fruit farming, and mixed farming.

Making Changes in Agricultural Production Branches: the dependent variable is considered as making/not making changes in the agricultural production branch, the desire of the farmer to change the main production branch, and products in other agricultural production branches.

Farmer's Annual Income: the annual income of the head of the household was asked about their average monthly income based on the minimum wage and its multiples. It is accepted that on a farm, the head of the household and the adult family members consist of four people. According to the data of the Central Bank of the Republic of Türkiye, the average of 1 U.S. dollar was accepted as 7.02 TL when the research was conducted.

Institutional Review Board Statement: the study was conducted according to the guidelines of the Declaration of Helsinki, and approved by Gazi University Ethics Committee (protocol code 2021-398 and date of approval E-77082166-604.01.02-70583).

2.4 Research hypothesis

Since the COVID-19 pandemic's scale and consequences will have different effects across agricultural branches, a qualitative research approach, including semi-structured interviews and qualitative content analysis, was used to determine its impact on agricultural branches common within the Turkish agrarian system. In a similar study, they reported that approximately four months after the start of COVID-19 pandemic restrictions, it would be appropriate to collect information from various agricultural systems on how the pandemic affected the functioning of agricultural systems worldwide (Stephens et al. 2022). Ragasa et al (2021), in their study conducted via telephone survey in June 2020 to measure the effects of the shock pandemic restrictions (restrictions between February and May 2020), reported that it impacted all agricultural branches and rural livelihoods. Similar methods have been previously applied to address the effects of COVID-19 on agriculture (Perrin & Martin, 2021; Snow et al. 2021; Mastronardi et al. 2020; Goswami et al. 2021). Therefore, taking advantage of the strengths of qualitative methods, we investigated the impact of COVID-19 as a new phenomenon on farmers' switching between branches of agriculture and also generated hypotheses for further quantitative research on the subject.

Null hypothesis (H0): the COVID-19 pandemic does not have a significant effect on changes among the main agricultural production branches in Türkiye (cereal farming, livestock farming, cereal and livestock, vegetable farming, fruit farming, and mixed farming).

Alternate hypothesis (H1): the COVID-19 pandemic has had a significant impact on changes among the primary agricultural production branches in Türkiye (cereal farming, livestock farming, cereal and livestock, vegetable farming, fruit farming, and mixed farming).

2.5 Data analysis technique

The study investigated whether there is a relationship between the socioeconomic and business characteristics of the farmers and whether they will change the primary agricultural branches during the COVID-19 pandemic restrictions. Nonlinear Canonical Correlation Analysis (NLCCA) was used to analyze these variables. This technique seeks the assumptions made for Canonical Correlation Analysis (CCA). In this context, prerequisites such as normal distribution of variables, lack of full correlation between variables, a large number of samples forming the data set, and absence of outliers in the data set are not present in NLCCA (Aydın et al. 2014).

NLCCA analyzes the relationships between two or more clusters of variables. The analysis does not make any assumptions about the linearity of relationships or the distribution of variables that may have different levels of measurement. This analysis is used in many fields as it includes categorical variables in the analysis as well as numerical variables and also includes the graphical representation of variables on two-dimensional maps (Filiz & Kolukısaoğlu 2012; Johnson & Wichern 2014; Yılmaz & Pulatsü 2021). The analysis was based on telephone interviews with 2,125 randomly selected aged 29 to 77.

The study encompassed various variables, including socioeconomic characteristics of farmers, characteristics of their agricultural businesses, changes in their agricultural branches, primary agricultural production branches, and the impact of COVID-19 pandemic restrictions. These variables were analyzed with NLCCA, interpreted, and explained with the help of graphics and tables.

Loss Function: in the nonlinear canonical correlation analysis with more than two clusters, the loss function is included in Equation 1 and the boundary conditions (Kolukısaoğlu 2013). The purpose of calculating the loss function is to try to find the minimum value of the function. Gifi, on the other hand, applied the methodology of loss functions differently. The canonical variables provided for each set are independent of each other. In other words, the vectors of canonical variables are orthogonal. Accordingly, when the NLCCA matrix is provided, this matrix will turn into the identity matrix (Rkk) (Bülbül & Giray 2012). Therefore, the loss function and constraint created for k sets are explained in the equations below (Özkan 2019).

$$\sigma_m(x, y) = K^{-1} \sum_j SSQ\left(x - \sum_{j,jk} G_j y_j\right) \quad (1)$$

Where: x, object scores; yj, vector of category digitizations; m, total number of variables; Gj, j. H indicator matrix for variable; SSQ (), sum of squares of matrix elements; K, number of clusters; jk, k. is the number of variables in the set

Constraints: some constraints are created to minimize the function obtained in the homogeneity analysis. The iterative technique that provides the optimum function with the help of these constraints is called the Alternating Least Squares (ALS) method.

$$Restricts = x'x = nl \ \ and \ \ u'x = 0 \quad (2)$$

Where: u', is a vector of size (1 x n) with 1 element; The K notation represents the number of clusters used in the Nonlinear Canonical Correlation method k. signifies the number of variables in the set.

Finding Eigen Values: eigenvalues give the amount of the relationship shown in the dimensions and are calculated according to Equation 3 (Meulman & Heiser 2005).

$$Eigenvalue = 1 - \frac{1}{kn} \sum_{p=1}^{p} \sum_{k=1}^{k} (X_p \ - \ Q_k a_{kp})2 \quad (3)$$

Where: p, number of dimensions; k, number of sets; n, number of objects (observations); Q, digitized data matrix; X, object scores matrix; AK, weight (number of sets) matrix

Fit: the fit value, equal to the maximum possible number of dimensions, gives the total explained variance value. The relationship is perfect if the fit value equals the number of dimensions. The sum of the eigenvalues of each dimension gives the fit value. The fit value is calculated with Equation 4 (Kolukısaoğlu 2013).

$$Fit = \sum_{p=1}^{p} Eigenvalue\ (p) = p - \frac{1}{kn} \sum_{p=1}^{p} \sum_{k=1}^{k} (X_p - Q_k a_{kp})2 \quad (4)$$

Canonical correlation coefficients are calculated using the formula in equation (5).

$$p_d = ((K * E_d) - 1)/(K - 1) \quad (5)$$

Where: K: variable set (cluster); d, number of dimensions; E_d, indicates the eigenvalue in dimension

3. Results

The findings are presented in four sections. These sections: (1) socioeconomic characteristics of respondents, (2) general problems of quarantine restrictions on agricultural production and changes in agricultural branches, and (3) relationships between individual characteristics and major production branches *(NLCCA)*.

3.1 Socioeconomic characteristics of respondents

In this section, socioeconomic characteristics and farm characteristics that are thought to impact farmers' desire to make changes in essential agricultural branches are discussed as variables. Although not shown in the table, all farmers participating in the survey were male, averaging 49.90 years. Approximately half of the farmers (47.1%) are above average age. It was determined that more than half of the farmers (57.1%) received primary and secondary school education. In agricultural farms, land width and number of animals are important sources of capital in both economic indicators and agricultural production. For this reason, the number of animals owned by farmers and the land width were determined. Although the land width owned by farmers varies between 0.3 and 250 hectares, the average land width is determined as 13.5 hectares. The majority of land areas consist of producers' own properties. The number of cattle is between 1 and 600 and the average number of cattle is determined as 18.48. The number of small ruminants is between 10 and 550, with 159.81 animals per farm. According to the agricultural farm structure in Türkiye, 64.9% of the farmers participating in the survey had a middle-income level. The variables and percentages used in the study are given in Table 2.

Table 2- Variables and percentages (n= 2 125)

Variables	*Category*	*Number*	*Percentage*	*Variables*	*Category*	*Number*	*Percentage*
Farmer's Age	≤ 30 Age	8	0.8	Land width	≤ 6.0	827	38.9
	31-40 Age	457	21.5		6.1-15.0	806	37.9
	41-50 Age	659	31.0		15.1-25.0	241	11.3
	≥ 51	1001	47.1		≥ 25	251	11.8
Education Levels	Primary School	639	30.1	Farmer's Annual Income	Low	299	14.1
	Secondary school	573	27.0		Middle	1379	64.9
	High School	673	31.7		High	447	21.0
	University	240	11.3				
Number of Ovine	No animals	1884	88.7	Number of Cattle	None	1022	48.1
	≤ 50	22	1.0		≤ 5	401	18.9
	51-100	54	2.5		6-10	177	8.3
	101-150	84	4.0		11-20	229	10.8
	≥ 151	81	3.8		≥ 21	296	13.9
Agricultural Production Branch	Cereal Farming	627	29.5	Change in Agricultural Production Branch	I'm thinking	351	16.5
	Livestock Farming	144	6.8		Undecided	780	36.7
	Cereal and Livestock	379	17.8		I don't think	994	46.8
	Vegetable Farming	295	13.9	Non-agricultural Income	Yes	1226	57.7
	Fruit Farming	592	27.9		None	899	42.3
	Mixed Farming	88	4.1				

3.2 General problems of quarantine restrictions on agricultural production and changes in agricultural branches

This section discusses the common problems faced by farmers during the first COVID-19 lockdown and the impacts of these problems. The questions in the questionnaire presented to the farmers were prepared based on previous research findings and literature review and were evaluated on a five-point Likert-type scale (1: not important, 2: least important, 3: partially important, 4: important, and 5: very important). Farmers were asked what the problems they encountered in agricultural production branches

were in the first phase of the pandemic restrictions. Seven explanatory variables were used to evaluate the variables affecting agricultural production. The value of the total scores of the problems faced by farmers in agricultural production (according to the Likert scale) is determined as a maximum of 35 and a minimum of 7. Those with average Likert scores below 2.5 were classified as less important and unimportant, and those with average scores below 2.5 were classified as essential and very important. The most critical problems farmers face during the restriction process are the need for more support policies, sudden increases in input prices, problems in the supply of agricultural inputs, difficulties in the sale of products, and labor shortages, respectively. These difficulties encountered during the pandemic are similar to the problems of Türkiye's agricultural structure before the pandemic (Yılmaz et al. 2006; Cevher et al. 2021). These problems encountered during the pandemic are consistent with the studies of many researchers (Yegbemey et al. 2021; Middendorf et al. 2021; Ragasa et al. 2021; Taylor et al. 2022). Researchers have found that farmers primarily experienced problems in accessing pesticides, fertilizers, and seeds during the pandemic restrictions, and these problems were followed by farmers' loss of income and difficulties in accessing local and urban markets.

On the other hand, during pandemic restrictions, problems of low importance, according to the Likert scale, such as access to agricultural land, problems with producers on the neighboring farm, and fear of pandemic due to restrictions, are listed as variables. The average Likert score of farmers regarding the problems encountered in agricultural production is $\bar{X}$= 17.49. Since this value ($\bar{X}$= 35/2) is lower than the Likert score, it shows that pandemic restrictions do not negatively impact the sustainability of agricultural production. This result may have been influenced by the fact that farmers received exemptions as long as they complied with the pandemic restrictions and the lack of farming activities (planting and harvesting) during this period. As a result, it can be said that farmers are less affected by the pandemic compared to individuals in other sectors, especially the factors mentioned above. The average Likert scores of the problems faced by farmers due to pandemic restrictions are shown in the table below (Figure 2).

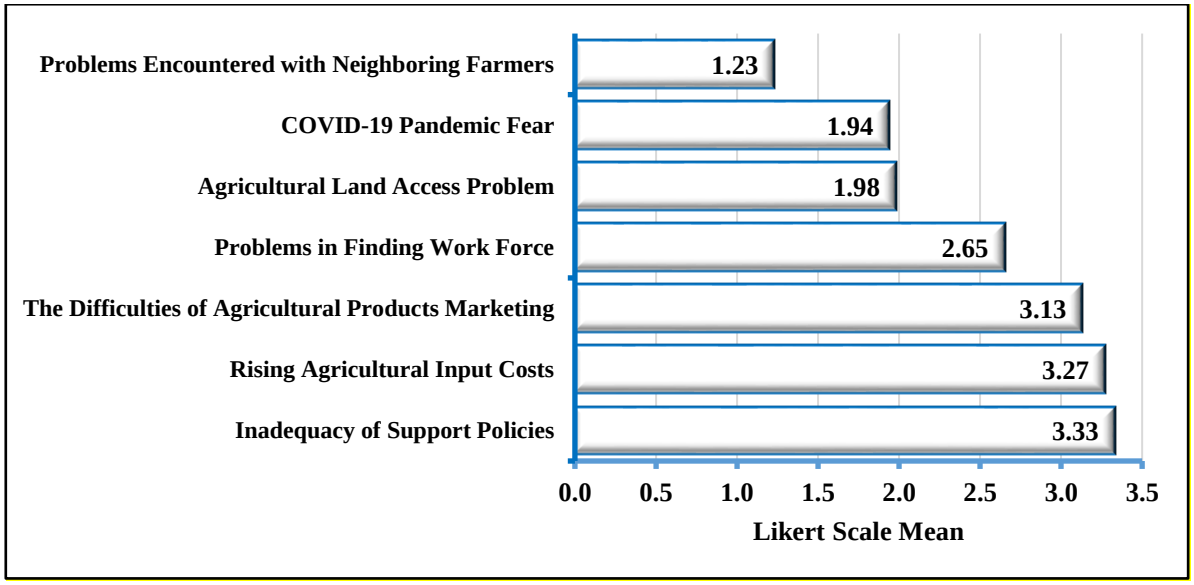


Figure 2- The scale of challenges faced by farmers in COVID-19 pandemic restrictions

During the pandemic restrictions applied across Türkiye (the shock period of the pandemic) and the period when survey data was collected (May-June 2020), farmers' wishes to make changes in essential agricultural branches are shown in Figure 3. Due to the shock restrictions applied throughout Türkiye, the rate of farmers who want to change their basic agricultural branch has been determined as 16.5%. Almost all (92.3%) of the farmers' requests for change in these branches of agriculture consist of farmers producing animal husbandry, vegetable and fruit production. In a similar study on this subject (Karakoç & Kısa 2023), it was reported that restrictions during the pandemic had an impact on farmers' decision-making and behavioral change. Data on farmers' desire to make changes in essential agricultural branches are shown in Figure 3.

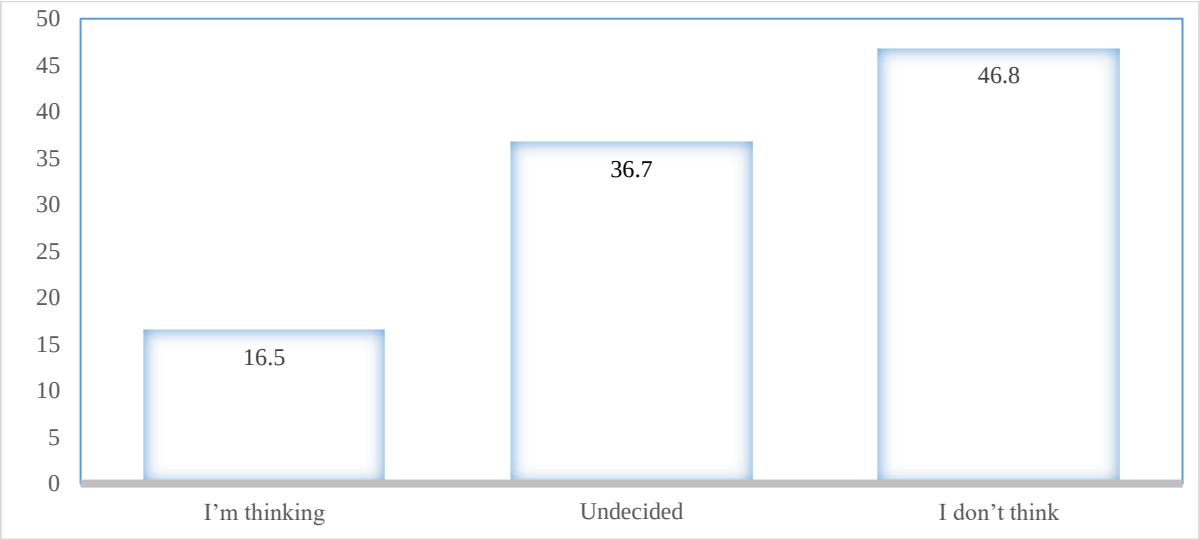


Figure 3- Willingness to Make Changes in Basic Agricultural Branches (%)

Pandemic restrictions have affected farmers' essential agricultural branches in different ways. For this reason, an attempt was made to determine the farmers' desire to change which agricultural branches and to what extent. In addition, the background information and essential characteristics of their perceptions of changing agricultural production branches were also determined. When Figure 4 is examined, it can be seen that the greatest desire to make changes (34.2%) is among farmers producing in the vegetable branch. It has been determined that this change request was caused by disruptions in input supply and marketing problems in the production process. This desire for change was followed by farmers who produce grain and animal products and grow only animal products.

Farmers in this production branch stated that they could not slaughter animals due to the closure of restaurants and other institutions (slaughterhouses), they could not market dairy products, and they wanted to change the production branch due to increased animal feed costs. On the other hand, grain was the branch of agricultural production in which farmers had the lowest desire to make changes. This was because the government announced grain purchase prices early and input supply in grain production was at its lowest level. According to these results, the effects of pandemic restrictions on agricultural branches are different. Ragasa et al. (2021) reported that the global economy and production branches were negatively affected during the pandemic's early periods and pandemic restrictions. Data on farmers' desire to make changes in essential agricultural branches are given in Figure 4.

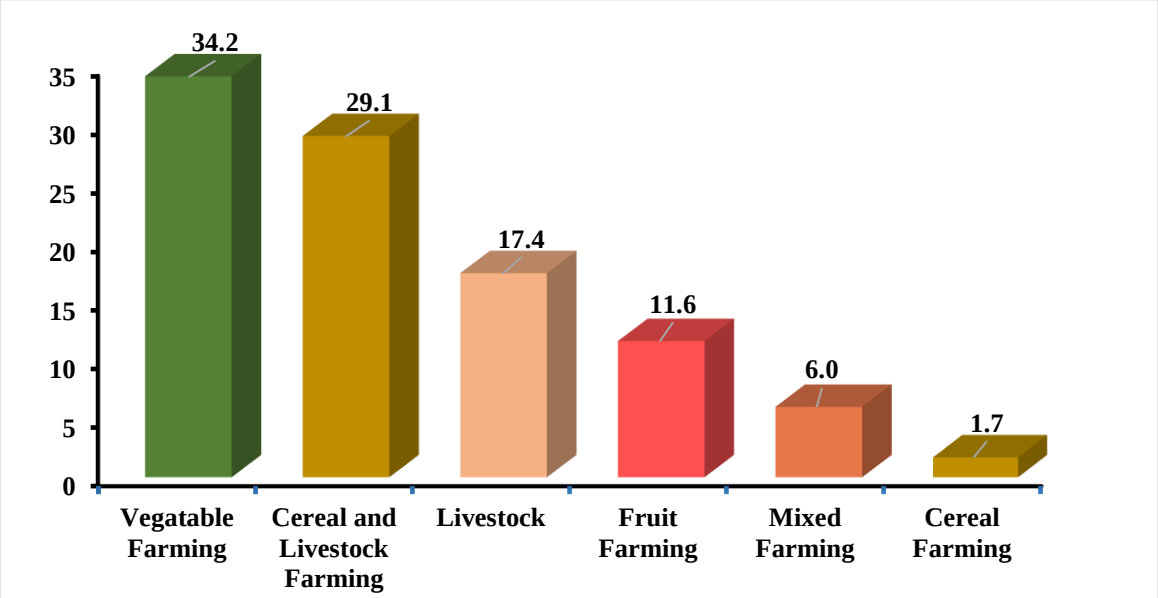


Figure 4- Willingness to make changes in Basic Agriculture Branches (%)

3.3. Relationships between individual characteristics and major production branches (NLCCA)

Table 3 shows the loss function and fit values that show how good the solution is. According to these values, in the nonlinear canonical correlation analysis, the uncalculated loss value for the first dimension was 0.247, which was determined as 0.312 for the second dimension. These values were found to be 0.248 and 0.313 in the second dimension, respectively. The averages of these values were found to be 0.247 in the first dimension and 0.313 in the second dimension. The calculation of the eigenvalues was obtained by subtracting the uncalculated loss values of the dimensions obtained from 1, and the eigenvalue for the first dimension was 0.753; it was calculated as 0.687 for the second dimension. The eigenvalue provides information about the relative efficiency of each discriminant function. This value is the most helpful measurement for OVERALS because it is equivalent to the intergroup correlation (www.ibm.com). In OVERALS, the fit is equal to the number of sets used most, and since there are two sets in this study, the highest fit probability is 2. The fit value, which constitutes the eigenvalue totals of both dimensions, was found to be 1.440. The maximum difference between the two fits was found to be 0.560 (2 - 1.440). Since the highest fit value will be as much as 2, which is the number of dimensions, a score of 1.440 can be accepted as an appropriate score. According to Table 3, since the total fit of the model is calculated as 1.440, it has a high value of 72.0% (1.440/2). However, while it can explain the real fit value of Set-1, which is 1.440, 52.3% (0.753/1.440), Set-2 can explain 47.7% (0.687/1.440) of the proper fit. In crafting the analysis summary, the order is structured by the eigenvalues of the dimensions. As shown in Table 3, the eigenvalue of the first dimension is higher than the second dimension. Canonical correlation coefficients are calculated using the formula outlined in the equation. Canonical correlation coefficient for the first dimension is 2 x 0.753 – 1 = 0.506; the correlation coefficient in the second dimension was calculated as 2 x 0.687 – 1 = 0.374. In short, according to the first dimension, there is a moderate relationship of 50.6% between the desire to make changes in primary production branches and age, education level, non-agricultural income, and annual business income.

Table 3- Summary of analysis

		Dimension		
		1	*2*	*Total*
Loss	Set 1	0.247	0.312	0.560
	Set 2	0.248	0.313	0.561
	Mean	0.247	0.313	0.560
Eigenvalue		0.753	0.687	
Fit				1.440

When examined in Table 4, the age and education variables in the first set and dimension make the highest contribution to the weight loads of the variables, while the farmer's annual income variable provides the highest contribution in the second set and dimension.

Table 4- Weight loads of variables

Set		*Dimension*	
		1	*2*
1	Change in Agricultural Production Branch	- 0.126	- 0.904
	Agricultural Production Branch	- 0.028	0.287
	Ages	0.841	- 0.206
2	Education Levels	- 0.874	0.204
	Non-farm Income	- 0.077	- 0.002
	Farmer’s Annual Income	- 0.012	- 0.853

Based on the idea that the above questions cannot determine the demand for change in agricultural production branches, the study is based on the fact that different indicators should also be used to define the desire for change in agricultural production branches. It was thought that not only the desire for change should be focused on, but also the groups that wanted the most change should be included in the analysis. For this purpose, the degree of relationship between sociodemographic characteristics, business infrastructure, COVID-19 restrictions, and the desire to make changes in agricultural production branches and agricultural production branches, as well as which categories are related to each other and whether they form homogeneous clusters have been determined.

When the component loads of the variables are examined (Figure 5), it is expected that the variables considered will be as far from the origin as possible. The greater the distance of the variables from the origin, the greater the importance of the variables considered. As shown in Table 3, the farmers under the age of 30, university graduates with low and high annual income, making changes in the relevant agricultural production branches, not making changes, and being undecided are essential variables. The component load values are the correlation coefficients between the transformed variables and the object scores, and the graph of the component load values is shown in Figure 4.

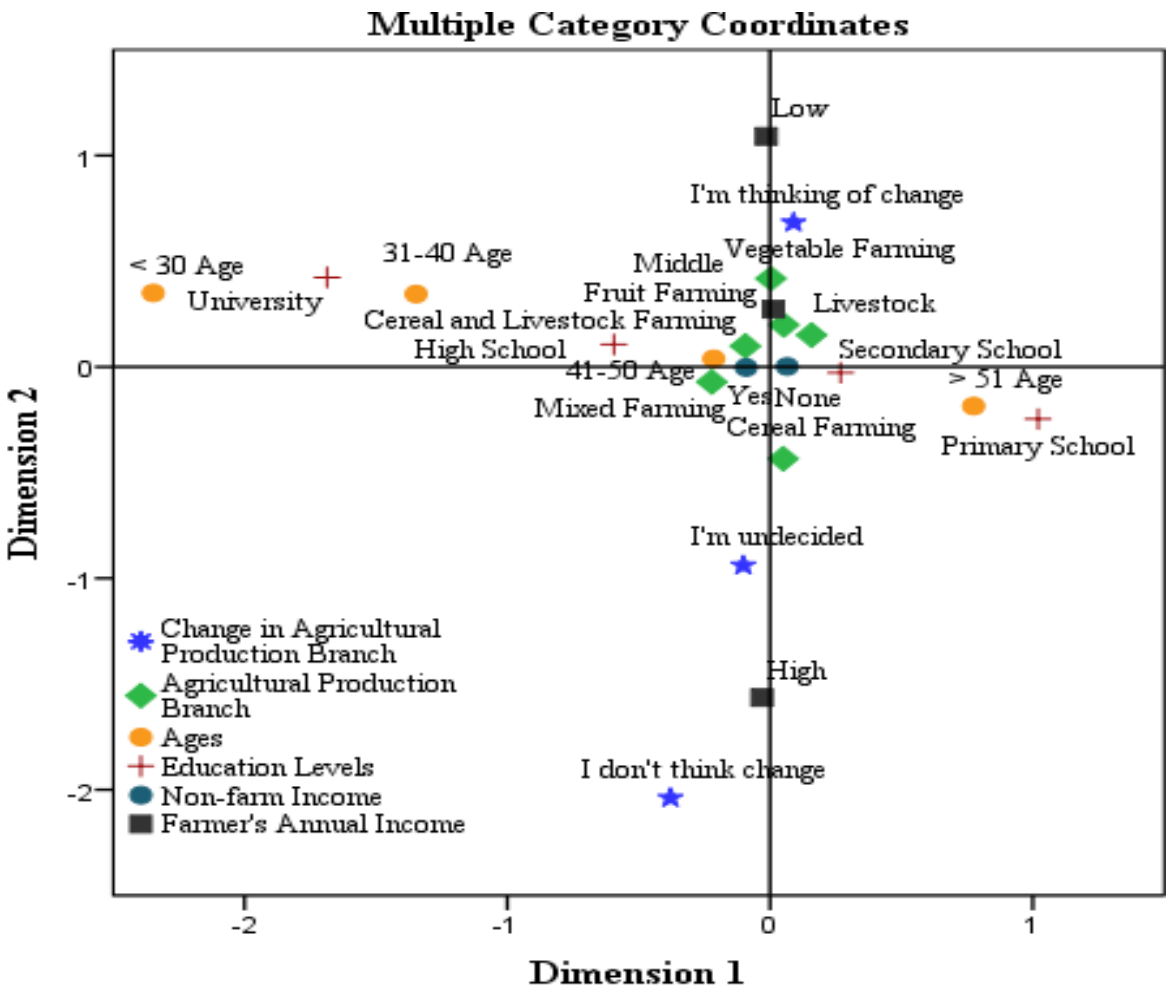


Figure 5- Graphical representation of component loads between category variables of individual variables and the desire to make changes in production branches

When the graph of the categories of the variables was examined, it was determined that the categories of the variables formed four homogeneous groups. In the first group, the relations between the categories of farmers under the age of 30 and in the group of 31-40, high school and university graduates, the desire to make changes in agricultural production branches due to COVID-19 restrictions, the state of being undecided, and the desire not to make any changes were found to be weak. In the second group, farmers who want to change the branches of vegetable growing, fruit growing, animal husbandry, low-income, and agricultural production constitute a homogeneous group.

According to these results, it can be seen that farmers who want to make changes in essential agricultural branches consist of farmers who produce animal products, vegetables and fruits. On the other hand, it has been determined that the desire of farmers who produce in the grain and mixed agriculture branches (farms with high grain production) to make changes in the essential agricultural branch is at a low level. Another remarkable result obtained from the study is that young and highly educated farmers have a low level of desire to make changes in their essential agricultural branches.

There is a strong relationship between the levels of variables with these characteristics. The willingness of those in this group to change their agricultural branches was higher than those in the other group. The strong relationship in this category was influenced by the inadequacy of the food supply chain during the pandemic, the closure of restaurants and shopping centers due to restrictions, the difficulty in obtaining production inputs, the inability to ship products daily, and the deterioration of products due to insufficient storage. It can be said that the third group, primary school and secondary school education level, farmers over 51 years of age, and cereal farming farmers have a weak desire to make changes in agricultural branches. It can be said that the farmers in the fourth group have high annual incomes and that the farmers in this group do not have the desire to make changes in their agricultural branches. It is seen that the desire to make changes in the agricultural branch of those with and without non-agricultural income, who are outside of these groups, is optional.

3.4 Relationships between farm infrastructure variables and the situation of making changes in basic production branches

In the fit analysis conducted to determine the relationships between farm infrastructure and basic agricultural branches (Table 5), the eigenvalues of the variables were found to be higher in the first dimension and the fit value constituting the sum of both

dimensions was determined as 1.965. This value is within acceptable fit values. Therefore, it can be seen that the total fit rate of the model is as high as 98.25%. While the true fit value explained 50.2% of the data in the first (set-1) dimension, this rate was determined as 49.8% in the second dimension. However, the canonical correlation coefficient for the first dimension was calculated as 0.972, and for the second dimension as 0.958. According to these results, there is a high positive correlation of 97.2% between the desire to make changes in agricultural production branches in the first dimension and the characteristics of the place where the farmer resides. In the second dimension, it was determined that there was a 95.8% high correlation between the desire to make changes in agricultural production branches and the characteristics of land width, number of cattle, and number of ovine. These values show a strong positive relationship between the variable sets considered in both dimensions.

Table 5- Summary of analysis

		Dimension		
		1	*2*	*Total*
Loss	Set 1	0.014	0.021	0.035
	Set 2	0.014	0.022	0.036
	Mean	0.014	0.021	0.035
Eigenvalue		0.986	0.979	
Fit				1.965

When Table 6 is examined, the desire for change in agricultural production branches in the first dimension and the weight value of the farmer's place of residence variable in the second dimension was found to be higher than other variables. Among the variables, the effect of the variable number of sheep and goats on the desire for change in agricultural production branches in the first dimension was found to be higher than other variables. In the second dimension, the effect of land width on the place of residence variable was found to be higher than other variables.

Table 6- Weight loads of variables

		Dimension	
		1	*2*
1	Change in Agricultural Production Branch	1.357	- 0.200
	Place of Residence	0.169	1.378
2	Land width	- 0.037	0.188
	Number of Cattle	- 0.157	0.031
	Number of Ovine	- 0.270	0.050

The further the variables seen in the component loadings plot are from the origin, the more critical that variable is for analysis (IBM 2021). When we inspect the lines represented in Figure 6, we can observe that they pertain to farmers residing in both rural and urban areas, farmers who express a desire to change their agricultural branch, those who do not wish to make changes, and farmers with a land width of 6.0 hectares or less. Component load values, correlation coefficients between transformed variables and object scores, and the graph of component load values are shown in Figure 6. When the figure was examined, it was determined that the variables formed two homogeneous groups. In the first group, we find farmers who desire to change their primary production branches, those who are undecided about making changes, farmers with a cattle number exceeding 21, and farmers with varying levels of small cattle numbers. There is a strong relationship between the levels of variables with these characteristics. It can be inferred that farmers with these characteristics are more inclined to either make changes in their primary production branches or remain undecided compared to individuals with different variable profiles. It has been determined that the farmers in this group are the farmers who produce for the market. According to these findings, it has been concluded that farmers with medium and large-scale farms are more willing to make changes in their essential agricultural branches.

On the other hand, farmers with less than five cattle and 6-10 cattle, with a land width of fewer than 6.0 hectares and between 6.1-15.0 hectares, tend not to make changes in these dimensions, essential production branches, and forming a separate homogeneous group. Factors such as the fact that the farmers in this category are small family businesses, the majority of the products produced are consumed on the farm, they do not need much in the food supply chain, and some of the production inputs are met within the farm have had an impact on the solid relationship in this category. It is observed that farmers with small-scale agricultural farms have a lower desire to make changes in essential agricultural branches than farmers with medium and large-scale farms. Our findings are similar to the studies (Middendorf et al. 2021; Menon & Schmidt-Vogt 2022). It can be clearly stated that the strong relationship in this category is influenced by the fact that the input use is the least in grain production, the absence of the harvest period, and the government's early announcement of grain prices.

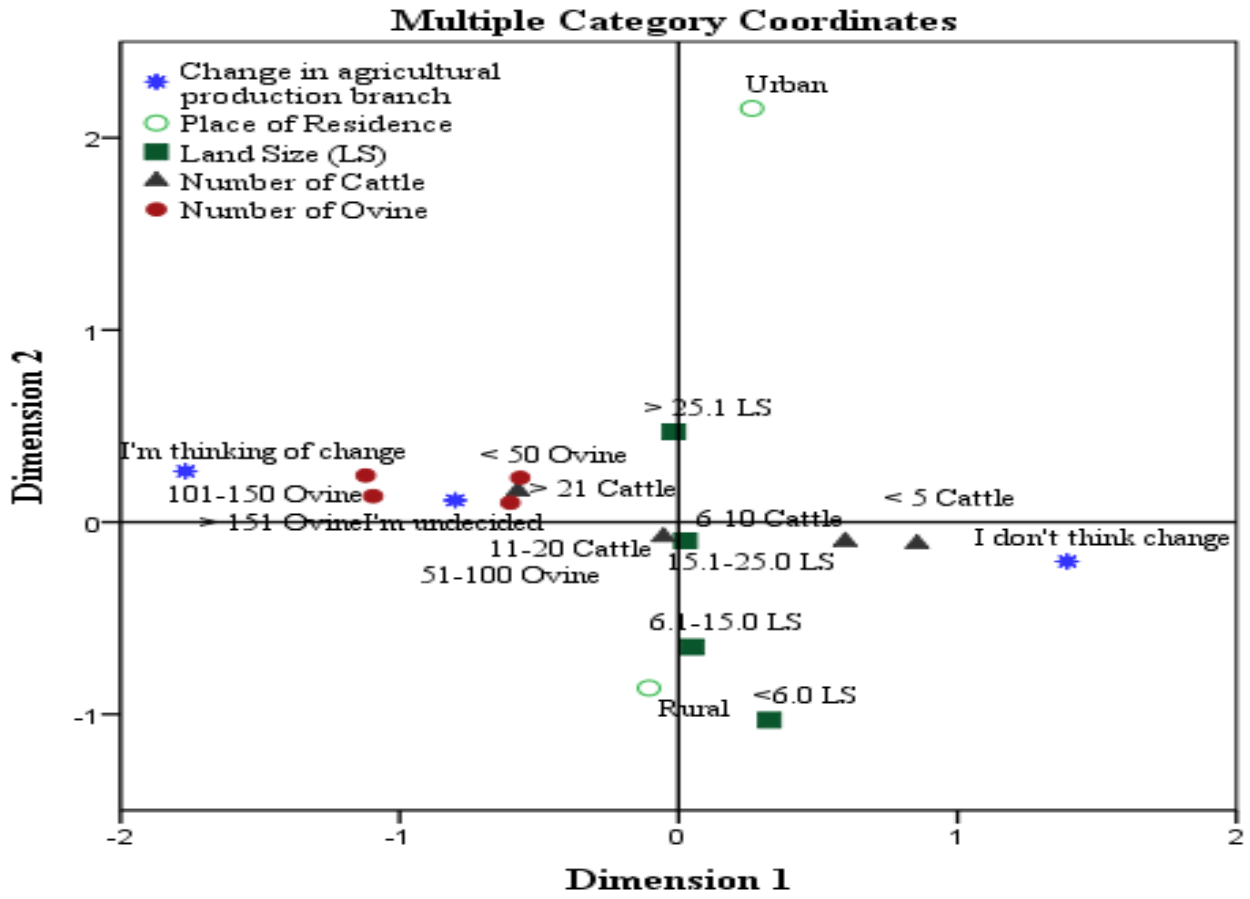


Figure 6- Graphical representation of the component loads between the category variables of the business infrastructure and the desire to make changes in production branches

4. Discussion

This study attempted to determine farmers' preferences for changes in their main agricultural production branches during the COVID-19 pandemic restrictions. In addition, the difficulties encountered in agricultural production during the pandemic, obstacles to agrarian production branches, factors affecting production decisions, problems encountered in the supply chain, the effectiveness of agricultural policies, and research gaps on this subject were systematically analyzed.

It was determined that the most common problems faced by farmers during the pandemic were (1 = not important; 5 = very important), the inadequacy of agricultural support policies implemented ($\bar{X}$= 3.33) and sudden increases in agrarian input costs ($\bar{X}$= 3.27). Farmers stated that they are considering seriously changing their farm practices (branches of agriculture) (16.5%) if COVID-19 and current challenges continue for another year. These change requests were made in animal husbandry and fruit and vegetable production branches, where agricultural inputs and labor are used intensively. During the first period of the pandemic, when the curfews began, difficulties encountered in the fruit and vegetable supply chain caused disruptions in consumers' access to food and farmers' agricultural production due to the closure of restaurants and the failure to establish neighborhood markets. In addition, the inadequacy of existing storage facilities for perishable products such as fruits, vegetables, and meat has caused difficulties in supplying and consuming these agricultural products. Middendorf et al. (2021) reported that there were difficulties in supplying input due to pandemic restrictions in animal production branches, vegetable growing, and fruit growing branches, and this caused a decrease in production.

Similarly, many studies have reported that farmers producing vegetables experienced difficulty accessing farm inputs and storing or selling produced fresh vegetables during the COVID-19 pandemic (Yegbemey et al. 2022; Husse et al. (2021). In the livestock branch, increased feed costs due to rising exchange rates, closure of restaurants and slaughterhouses, and disruptions in the product supply chain have caused production problems in livestock farms. In addition, the deficiencies in existing agricultural policies and lack of preparation for the pandemic have further increased the main threats and issues to agriculture during the pandemic. For this reason, it has been determined that the desire of low-income farmers who are engaged in vegetable farming, fruit growing, and animal husbandry to change their essential agricultural branches is higher than that of farmers in other agrarian branches.

In order to alleviate the problems encountered during the pandemic period and ensure sustainability in agricultural production, many agricultural productions could be continued successfully due to the government's relaxation of restriction measures (for farmers). Similar studies also reported that the impact of the pandemic on the increase in food prices was minimal (Hobbs 2020), and in another study, the supply of vegetables, fruits, and oils decreased by 10%, and the effect of this situation on prices was limited (Mahajan & Tomar 2020). Despite the possibility of prolonged pandemic periods, governments should develop agricultural policies that will protect farmers during pandemic periods, prevent significant changes such as transition between agricultural branches, and ensure the sustainability of food supply. Therefore, one of the main conclusions of our research is that policymakers need to produce more agricultural policies to promote low-cost and easy-to-use storage facilities for vegetables, fruits, and animal products.

Most farmers engaged in livestock and vegetable production in Türkiye consist of small family businesses (Aşkan & Dağdemir 2015) and farmers with low agricultural income. These farmers can make more emotional and intuitive decisions (Cevher et al. 2021). In our study, it was determined that the desire of farmers in animal husbandry, vegetable growing, and fruit growing branches to make changes in their bare agricultural branches was at a higher level than that of farmers in other agricultural branches. The most critical problems farmers in this group faced during the pandemic were, respectively, sudden price increases in animal feed inputs, lack of small-scale cold storage for the preservation of products, and difficulties encountered in the supply chain. In order to overcome these difficulties, it is necessary to increase product diversity, popularize mixed agricultural production, and expand the use of agricultural technology in small agricultural farms, which will contribute more to the sustainability of agricultural production. This approach will increase resilience to pandemics and create additional income, employment, and risk reduction opportunities for small farmers (Behera & France 2016; Nzekwe et al. 2018). Biswal et al. (2020) and Stephens et al. (2020) stated that small-scale farmers are preferred during pandemic periods due to labor movement outside the farm and decreased input supply. Small-scale farms account for the largest share of food production, especially in developing countries (Frelat et al. 2016), and are therefore critical for food security (Husse et al. 2021). On the other hand, Gu & Wang (2020) determined that due to the COVID-19 pandemic, the incomes of farmers with small businesses generally decreased, and they suffered more losses. Therefore, their desire to make changes in agricultural production branches increased. During the pandemic, farmers with small businesses faced more risks due to insufficient tools and equipment (Hatab et al. 2020; Cordeiro et al. 2021).

Grain production in the crop production branch is expected to decrease by 15% in 2021 compared to 2020 (TurkStat 2021). It was concluded that this decrease was not due to the COVID-19 pandemic restrictions but to the negative environmental impact and excessive increase in input costs (Cevher et al. 2021). The decrease in vegetable production is due to the rise in input costs. While the production number of products such as apple, grape, cherry, and pistachio decreased in 2021 compared to the previous year, an increase in the production number of products such as peach, olive, strawberry, tangerine, and orange is expected (TurkStat 2021). According to these data, it can be said that the amount of fruit production was not affected much by the COVID-19 outbreak. Martínez-Azúa et al. (2021) supported this observation in their study, highlighting that among current economic indicators, the agrifood sector is one of the least affected by the pandemic-induced crisis.

As can be seen from official statistical data, it shows that there was no shortage of production in agricultural production branches during the pandemic period. This indicates that the first period of the pandemic shock did not cause significant problems in the amount of agricultural production in Türkiye, but it impacted the basic structure of the distribution of farm products. Therefore, more agricultural policies need to be developed to address the problems farmers face in the livestock, vegetable, and fruit fields. Paarlberg (2021) emphasized the importance of re-evaluating the challenges faced by modern science and technology in agriculture and food production and distribution in the post-COVID-19 era. Our study shows that, given the institutional weaknesses revealed by the pandemic, efforts to increase the resilience of agricultural production during pandemic periods should be at the center of agricultural policy programs. Therefore, there is a need to develop country- and region-specific policies, strategies, and reforms after the COVID-19 outbreak to ensure safe and sustainable agricultural production and to create more resilient agriculture-food systems that can withstand sudden shocks. Our current study provides insights into measures that can be implemented to ensure sustainable agriculture in Türkiye. However, there will be a need to investigate some of the "known unknowns" regarding the ongoing short- and long-term impacts of COVID-19 and similar outbreaks and potential future opportunities.

Key "unknowns" include: will changes in critical branches of agriculture affect food security in Türkiye? How will the change in agricultural branches affect sustainable production and the required product diversity? We suggest that these questions should be included in future research agendas. Therefore, in addition to the measures that governments need to take, there are many areas that individuals, producers, professional organizations, food trade, industrial organizations, and civil stakeholders in urban and rural areas need to pay attention to and take the initiative. New agricultural policies and a high level of coordination are needed in order to ensure the sustainability of agrarian branches negatively affected by pandemic restrictions and to increase cooperation between geographical regions and economic sectors.

5. Conclusions

As a result, the restrictions imposed during the pandemic impacted the vegetables, fruits, and livestock branches, especially in marketing. However, the adverse effects of the pandemic on production, supply of inputs, and access to labor were not felt much. These findings show that farmers did not encounter severe problems in agricultural production during the pandemic period. This situation is a positive outcome of the measures taken by the relevant ministries of the government to prevent the pandemic. With these measures, there was no shock decline in agricultural production due to the early announcement of base purchase prices in some agricultural branches (for example, the grain branch) and the fact that this was the period when input demand and access to labor were lowest. For this reason, the rate of farmers' desire to switch from existing agricultural production branches to other agricultural branches was 16.5%. It was concluded that the desire to make changes among the agricultural branches, respectively, consists of farmers who produce vegetables at a rate of 34.2%, livestock farming at 17.4%, and fruits at 11.6%.

Within the framework of the findings, the government needs to revitalize by supporting effective management, necessary financial support, and disadvantaged vegetable, fruit, and animal production branches. It is also expected that all stakeholders will take all possible measures to combat distressing situations by creating an enabling environment for their livelihood. Public and private sector organizations, non-governmental organizations, and municipalities are also considered necessary in this context. This study will constitute an essential preliminary step in determining the levels of similar variables that will be addressed in studies on the impact of changes in agricultural production branches on the basic agricultural production structure of the country. Our findings provide rich and rigorous information on the effects of COVID-19 on key branches of agriculture in Türkiye. However, the proposed short duration of telephone surveys and other disadvantages compared to face-to-face interviews have somewhat limited our insights into the depth of the impact of the COVID-19 crisis. It is the first study conducted in Türkiye to minimize the adverse effects on bare agricultural branches during COVID-19 restrictions and similar pandemics and identify existing problems.

References

Akbudak N & Şen Ö (2021). GLOBALGAP in the COVID-19 pandemic process. Turkish Journal of Agricultural Research 8(2): 248-255 (in Turkish)

Akın Y, Celen B, Çelen M F & Karagöz A (2020). Agriculture and pandemic: how should Turkish agriculture change after COVID-19? EJONS International Journal on Mathematics. Engineering and Natural Sciences 36(4): 904-914. (in Turkish)

Ali S S, Ahmad M R, Shoaib, J U M, Sheik M A, Hoshain M I, Hall R L, Macintosh K A & Williams P N (2021). Pandemic or environmental socioeconomic stressors which have greater impact on food security in the barishal division of Bangladesh: initial perspectives from agricultural officers and farmers. Sustainability 13(10): 5457

Anonymous (2022). Republic of Türkiye Ministry of Agriculture and Forestry General Directorate of Plant Production. Retrieved in February 23, 2024 from https://www.tarimorman.gov.tr/Haber/5503/Ciftci-Kayit

Aşkan E & Dağdemir V (2015). Analysis of the factors affecting the production rates of dairy livestock facilities benefitting from governmental financial supports and incentives: the sample of Erzurum, Erzincan, Bayburt Provinces. Turkish Journal of Agricultural Economics 21(2): 69-76 (in Turkish)

Aydın S, Görmüş A Ş & Altıntop M Y (2014). The Relationship between the satisfaction level of students and their demographic features with non-linear canonical correlation analysis: an application in Vocational High School. Abant Izzet Baysal University Journal of Abant Social Sciences, 14(1): 35-58 (in Turkish).

Balcı A (1997). Sosyal Bilimlerde Araştırma Yöntemi Teknik ve İlkeleri. 2. Baskı. Ankara: 72 TDFO Yay. Ltd. Şt.

Behera U K & France J (2016). Integrated farming systems and the livelihood security of small and marginal farmers in India and other developing countries. Advances in Agronomy, Academic Press 138: 235-282

Biswal J, Vijayalakshmy K & Rahman H (2020). Impact of COVID-19 and associated lockdown on livestock and poultry sectors in India. Veterinary World 13(9): 1928-1933

Bülbül Ş & Giray S (2012). Examining of the Relationship structure between job and special life (nonwork) satisfaction with Nonlinear Canonical Correlation Analysis. Anadolu University Journal of Social Sciences 12(4): 101-116 (in Turkish).

Ceballos F Kannan S & Kramer B (2020). Impacts of a national lockdown on smallholder farmers' income and food security: empirical evidence from two states in India. World Development 136: 105069

Cevher C & Altunkaynak B (2020) Socioeconomic factors and sustainable forage crops production in Turkey Aegean Region: A Multivariate Modeling. Sustainability 12: 8061

Cevher C, Adanacıoğlu H & Doğan N (2021) Impacts of the first wave of the COVID-19 on the agricultural production in Turkey: assessing the roles of farmers during the pandemic. In: XII International Agriculture Symposium "AGROSYM 2021", 7-10 October 2021, Bosnia and Herzegovina.

Cordeiro M C, Santos L & Marujo L G (2021). COVID-19 and the fragility of Brazilian small farming resilience. Brazilian Journal of Operations & Production Management 18(2): 1-14

Ding Y, Wang C, He L, Tang Y, Li T & Yin Y (2021). Effect of COVID-19 on animal breeding development in China and its countermeasures. Animal Frontiers 11(1): 39-42

FAO (2020). Food systems and COVID-19 in Latin America and the Caribbean: labour market. https://repositorio.cepal.org/bitstream/handle/11362/46053/cb0973

Filiz Z & Kolukısaoğlu S (2012). Nonlinear canonical correlation analysis and an application. International Journal of Management Economics and Business 8(16): 59-74 (in Turkish).

Frelat R, Lopez-Ridaura S, Giller K E, Herrero M, Douxchamps S, Djurfeldt A A & Van-Wijk M T (2016). Drivers of household food availability in sub-Saharan Africa based on big data from small farms. Proceedings of the National Academy of Sciences 113(2): 458-463

Goswami R, Roy K, Dutta S Ray K, Sarkar S, Brahmachari K, Nanda M K, Mainuddin M, Banerjee H Timsina J & Majumdar K (2021). Multi-faceted impact and outcome of COVID-19 on smallholder agricultural systems: integrating qualitative research and fuzzy cognitive mapping to explore resilient strategies. Agricultural Systems 189: 103051

Gu H & Wang C (2020). Impacts of the COVID-19 pandemic on vegetable production and countermeasures from an agricultural insurance perspective. Journal of Integrative Agriculture 19(12): 2866-2876

Gürbüz I B & Özkan G (2021). Will agriculture beat the odds against Covid-19? The Covid-19 outbreak and its effect on agricultural supply in Turkey. New Media: Mediterranean Journal of Economics, Agriculture and Environment 20(2): 15-26

Hatab A A, Lagerkvist C J & Esmat A (2020). Risk perception and determinants in small- and medium-sized agrifood enterprises amidst the COVID-19 pandemic: evidence from Egypt. Agribusiness: An International Journal 37(1): 187-212

Hobbs J E (2020). Food supply chains during the COVID-19 pandemic. Canadian Journal of Agricultural Economics 68(2): 171-176

Husse M, Brander M, Kassie M, Ehlert U & Bernauer T (2021). Improved storage mitigates vulnerability to food-supply shocks in smallholder agriculture during the COVID-19 pandemic. Global Food Security 28: 100468.

IBM (2021). IBM Knowledge Center. Retrieved in February 23, 2024 from https://www.ibm.com/support/knowledgecenter/en/sslvmb_24.0.0/spss/disc-rim_bankloan_e-igen.html.

Johnson R A & Wichern D W (2014). Applied Multivariate Statistical Analysis; Pearson: London, UK.

Karakoç H & Kısa C (2023). Investigation of the effect of Covid-19 on anxiety, fear of death and obsessive compulsory disorder. Turkey Journal of Integrative Psychotherapy 6(11): 1-13 (in Turkish)

Ker A P & Biden S (2021). Risk management in Canada's agricultural sector in light of COVID-19: considerations one year later. Canadian Journal of Agricultural Economics 69(2): 299-305

Kolukısaoğlu S (2013). Nonlinear canonical correlation analysis' application on depression anxiety and stress scale. M.Sc. Thesis, Eskisehir Osmangazi University Graduate School of Natural and Applied Sciences, Department of Statistics, Türkiye. (in Turkish)

Kumaran M, Geetha R, Antony J, Vasagam K P K, Anand P R, Ravisankar T, Angel J R J, Muralidhar D D M, Patil P K & Vijayan K K (2021). Prospective impact of Coronavirus disease (COVID-19) related lockdown on shrimp aquaculture sector in India-a sectoral assessment. Aquaculture 531: 735922

Laborde D, Martin W, Swinnen J & Vos R (2020). COVID-19 risks to global food security. Science 369: 500-502

Mahajan K & Tomar S (2020). COVID-19 and supply chain disruption: evidence from food markets in India. The American Journal of Agricultural Economics 103(1): 35-52

Mastronardi L, Cavallo A, Romagnoli L (2020). Diversified farms facing the COVID-19 pandemic: first signals from Italian case studies. Sustainability 12 (5709)

Martínez-Azúa B C. López-Salazar P E & Sama-Berrocal C (2021). Impact of the COVID-19 pandemic on agrifood companies in the region of Extremadura (Spain). Agronomy 11(5): 971

Menon A & Schmidt-Vogt D (2022). Effects of the COVID-19 Pandemic on farmers and their responses: a study of three farming systems in Kerala, South India. Land 11: 144

Meuwissen M P M, Feindt P H, Spiegel A, Termeer C J A M, Mathijs E, Finger R, Balmann A, Wauters E, Urquhart J, Vigani M, Zawalińska K, Herrera H, Nicholas-Davies P, Hansson H, Paas W, Slijper T, Coopmans I, Vroege W, Ciechomska A, Accatino F, Kopainsky B, Poortvliet M P, Candel J J L. Maye D, Severini S, Senni S, Soriano B, Lagerkvist C J, Peneva M, Gavrilescu C, Reidsma P (2019). A framework to assess the resilience of farming systems. Agricultural Systems 176: 102656

Middendorf B J, Faye A, Middendorf G, Stewart Z P, Jha P K & Vara-Prasad P V (2021). Smallholder farmer perceptions about the impact of COVID-19 on agriculture and livelihoods in Senegal. Agricultural Systems 190: 103108

Mishra A, Bruno E & Zilberman D (2021). Compound natural and human disasters: managing drought and COVID-19 to sustain global agriculture and food sectors. Science of The Total Environment 754: 142210

Meulman J J & Heiser W J (2005). SPSS categories 14.0 SPSS Inc.

Nzekwe C A, Onwudinjo L E & Okonkwo P N (2018). The challenges and prospects of small-scale farmers in Enugu state, Nigeria. International Journal of Novel Research in Humanities, Social Science and Management 1(1): 19-24

Obayelu A E, Obayelu, O A, Bolarinwa, K B, Oyeyinka, R A (2021). Assessment of the Immediate and potential long-term effects of COVID-19 outbreak on socioeconomics, agriculture, security of food and dietary intake in Nigeria. Food Ethics 6: 5

OECD (2020). Evaluation of the Impact of the Coronavirus (COVID-19) on Fruit and Vegetables Trade. 1-14. Retrieved in February 23, 2024 from https://www.oecd.org/agriculture/fruit-vegetables/oecd-covid-19-impact-on-fruit-and-vegetables-trade.pdf

Orden D (2021). Agrifood markets and support in the United States after 1 year of COVID-19 pandemic. The Canadian Journal of Agricultural Economics 69(3): 243-249

Özkan M (2019). Evaluation of selected demographic and socio-cultural concepts by Nonlinear Correlation Analysis. Eskişehir Osmangazi University Journal of Economics and Administrative Sciences 14(2): 391-408 (in Turkish)

Paarlberg R (2021). Resetting the table: straight talk about the food we grow and eat. New York: Alfred A. Knopf.

Perrin A & Martin G (2021) Resilience of French organic dairy cattle farms and supply chains to the COVID-19 Pandemic. Agricultural Systems 190: 103082

Pu M & Zhong Y (2020). Rising concerns over agricultural production as COVID-19 spreads: lessons from China. Global Food Security 26: 100409

Rahman M S & Das G C (2021). Effect of COVID-19 on the livestock sector in Bangladesh and recommendations. Journal of Agriculture and Food Research 4: 100128

Ragasa C, Lambrecht L, Mahrt K, Aung Z. W & Wang W (2021). Immediate impacts of COVID-19 on female and male farmers in central Myanmar: phone-based household survey evidence. Agricultural Economics 52(3): 505-523

Savary S, Akter S, Almekinders C, Harris J, Korsten L, Rötter R, Waddington S & Watson D (2020). Mapping disruption and resilience mechanisms in food systems. Food Security 12: 695-717

Stephens E C, Martin G, Wijk M V, Timsina J, Snow V (2020). Editorial: impacts of COVID-19 on agricultural and food systems worldwide and on progress to the sustainable development goals. Agricultural Systems 83: 102873

Stephens E, Timsina J, Martin G, Wijk M V, Klerkx L, Reidsma P, Snow V (2022). The immediate impact of the first waves of the global COVID-19 pandemic on agricultural systems worldwide: reflections on the COVID-19 special issue for agricultural systems. Agricultural Systems 201: 103436

Snow V, Rodriguez D, Dynes R, Kaye-Blake W, Mallawaarachchi T, Zydenbos S, Cong L, Obadovic I Agnew R, Amery N (2021). Resilience Achieved via multiple compensating subsystems: the immediate impacts of COVID-19 control measures on the agrifood systems of Australia and New Zealand. Agricultural Systems 187: 103025

Taylor D E, Farias L M, Kahan L M, Talamo J, Surdoval A, McCoy E.D, Daupan, S M (2022). Understanding the challenges faced by Michigan's family farmers: race/ethnicity and the impacts of a pandemic. Agriculture and Human Values 39: 1077-1096.

TurkStat (2021). Production of cereals and other crops. Retrieved in February 23, 2023 from https://data.tuik.gov.tr.

Uğur A & Buruklar T (2020). Effects of COVID-19 pandemic on agrifood production and farmers. Food Science and Technology 42, 1-10.

Ullah A, Zeb A, Liu J, Mahmood N & Kächele H (2021). Transhumant pastoralist knowledge of infectious diseases and adoption of alternative land use strategies in the Hindu-Kush Himalayan (HKH) region of Pakistan. Land Use Policy 109: 105729.

Varshney D, Kumar A, Mishra A K, Rashid S & Joshi P K (2021). COVID-19, Government transfer payments, and investment decisions in farming business: evidence from Northern India. Applied Economic Perspectives and Policy 43(1): 248-269

Worku A & Ülkü M A (2021). Analyzing the impact of the COVID-19 pandemic on vegetable market supply in Northwestern Ethiopia. Journal of Agribusiness in Developing and Emerging Economies 12(3): 371-385

Yegbemey RN, Komlan Ahihou CM, Olorunnipa I, Benali M, Afari-Sefa V, Schreinemachers P (2021). COVID-19 effects and resilience of vegetable farmers in North-Western Nigeria. Agronomy 11(9): 1808

Yoshida S & Yagi H (2021). Long-term development of urban agriculture: resilience and sustainability of farmers facing the COVID-19 pandemic in Japan. Sustainability 13: 4316

Yılmaz E & Pulatsü S (2021). Evaluation of fishermen opinions about inland fisheries with nonlinear canonical correlation analysis. Ege Journal of Fisheries and Aquatic Sciences 38(1): 11-19 (in Turkish)

Journal of Agricultural Sciences (Tarim Bilimleri Dergisi) 2025, 31 (1) : 126 – 136 DOI: 10.15832/ankutbd.1458732

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

Evaluation of Oxidation Stability of Organic and Conventional Hazelnut Oils

Sümeyye Şahin[a*], Caner Ümit Topçu[a]

[a]*Department of Food Engineering, Faculty of Agriculture, Ordu University, Ordu, TÜRKIYE*

ARTICLE INFO
Research Article
Corresponding Author: Sümeyye Şahin, E-mail: sumeyyesahin@odu.edu.tr
Received: 25 March 2024 / Revised: 08 August 2024 / Accepted: 16 August 2024 / Online: 14 January 2025

Cite this article
Şahin S, Topçu C Ü (2025). Evaluation of Oxidation Stability of Organic and Conventional Hazelnut Oils. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):126-136. DOI: 10.15832/ankutbd.1458732

ABSTRACT
In this study, the oxidation stabilities of organic hazelnut oils (OHO) were compared to conventional hazelnut oils (CHO). For this purpose, oxidation parameters such as peroxide value (PV), free fatty acid (FFA), anisidine value (AV), volatile lipid oxidation compounds (VLOC), fatty acid composition, and total antioxidant capacity (TAC) of the hazelnut oil samples were investigated under conditions of accelerated storage. At the beginning of storage, OHO had lower PV, higher AV, TAC and linoleic acid content than CHO and similar FFA, VLOC, palmitic, stearic, and oleic acids contents to CHO. As expected, PV, FFA and AV increased in OHO and CHO during storage, while TAC decreased. No significant difference between OHO and CHO was observed in terms of FFA, TAC, palmitic and stearic acid contents at the accelerated storage condition. When compared to CHO, OHO showed lower oleic acid and higher linoleic acid at the end of storage. During storage, the highest AV and PV were determined in OHO. Results reveal that OHO has lower oxidation stability than CHO. According to these results, it can be recommended to consume the OHO without being exposed to long-term oxidation.

Keywords: Oxidation stability, Organic hazelnut oil, Accelerated storage

1. Introduction

The ongoing rise in the global population has led to an increase in the need for food. To meet this increasing need for food, several farming practices have been implemented to increase productivity in agricultural activities. Some of these practices are the use of chemical fertilizers, herbicides, and pesticides. Conventional agriculture is a farming system including the use of chemical fertilizers, herbicides, pesticides, and GMO (genetically modified organisms) (Unluturk et al. 2021). Since the 1950s, conventional farming has produced a significant increase in productivity, but at the same time has also resulted in higher environmental pressures (Mondelaers et al. 2009). Organic farming has surfaced as a viable alternative to counteract the adverse impacts of conventional agriculture on both human health and the environment (Unluturk et al. 2021). Organic agriculture practices to uphold natural equilibrium and strives to yield high-quality food free from residues that may pose risks to human and animal health. To achieve this goal, it refrains from utilizing chemical fertilizers, pesticides, and GMO (Mondelaers et al. 2009). Therefore, products obtained from organic farming are considered safer and healthier by consumers and the interest in these foods is increasing day by day despite their higher prices. The increasing demand for organic foods has led to an increase in organic production. According to the latest available data on organic farming worldwide, the area used for organic agricultural production continues to increase and currently, 1.6 percent of the world's farmland is organic (Willer et al. 2023). The rapid development in organic products has prompted scientists to study the nutritional quality of organic products. In this regard, some researchers investigated differences in nutrient content and bioactive compounds of food items produced through organic and conventional methods. Many researchers compared the nutrient content and bioactive compounds content in organic vs. conventional food products in the following species: virgin olive oils (Jimenez et al. 2014); apples (Weibel et al. 2000); apple juice (Heinmaa et al. 2017); strawberry (Anttonen et al. 2006), red oranges (Tarozzi et al. 2006): peach and pear (Carbonaro et al. 2002); tomatoes (Juroszek et al. 2009); sunflower oil (Perretti et al. 2004). Recently, Karaosmanoğlu (2022) determined the chemical differences between organic and conventional hazelnuts. While numerous studies have explored the chemical composition of organic products, a significant gap exists in the literature concerning their storage stability.

Hazelnut (*Corylus avellana* L.) stands as the second most extensively cultivated tree nut globally, following almonds (Alasalvar & Shahidi 2009). Considering world hazelnut production, a total of 1 077 117 tons of hazelnuts were produced in 2021 and 684 000 tons of this was grown in Türkiye. With this production amount, Türkiye meets 64% of the world hazelnut production as the leading hazelnut-producing country (Food and Agriculture Organization 2023). The main component of unshelled hazelnut kernels is fat varying between 54 and 66% (Li & Parry 2011; Karaosmanoğlu 2022; Şahin et al. 2022).

Hazelnut oil contains very high amounts of mono- and polyunsaturated fatty acids (Alasalvar et al. 2006; Karaosmanoğlu, 2022; Şahin et al. 2022); which are known to be effective in preventing heart diseases (Alasalvar & Shahidi 2009). Among the fatty acid constituents, oleic acid is the primary constituent, ranging from 78 to 85% (Karaosmanoğlu 2022; Şahin et al. 2022; Şahin 2023; Tüfekci & Karataş 2018). Hazelnut oil is also used in the kitchen for its slightly nutty taste, in pharmacy and medicine for its cholesterol-lowering and vasoconstrictive effects, in the food industry for baking and as a massage oil in cosmetics (Krist 2020). Moreover, hazelnut oil is recognized for its ability to resist oxidation (Şahin 2023).

Primary goal of this study was to compare stability of organic hazelnut oil during accelerated oxidation with conventional hazelnut oil in terms of the free fatty acid, peroxide value, anisidine value, fatty acid composition, total antioxidant capacity, and volatile lipid oxidation compounds.

2. Material and Methods

2.1. Materials

The organic refined hazelnut oil (OHO) and conventional refined hazelnut oil (CHO) were supplied from Altaş Oil Company (Ordu, Turkey). OHO was certified by Letis Organic Quality Certifier (certificate no 1646/2023-01). Sodium thiosulfate, DPPH, potassium methoxide, trolox, potassium iodide, chloroform, ethyl alcohol, sodium carbonate, and diethyl ether were acquired from Sigma Aldrich. Butanol, p-anisidine, isooctane, and sulfuric acid were supplied by Merck.

2.2. Methods

2.2.1. Accelerated storage conditions

The organic and conventional hazelnut oils were stored in accelerated storage conditions, as outlined in literature (Petersen et al. 2012a). Briefly, the oils were filled into 100 mL brown glass bottles and kept in a drying cabinet at 80 °C for 14 days with the glass bottles caps open. The oil samples were taken from the drying cabinet for analysis at two-day intervals (0, 2, 4, 6, 8, 10, 12 and 14 days) and maintained in a frozen state until the analysis was conducted.

2.2.2. Peroxide value (PV) analysis

PV was determined using the DGF method (Deutsche Gesellschaft für Fettwissenschaft 2002). For this purpose, 2-2.5 grams of the oil samples were dissolved in an equal volume of isooctane: acetic acid. The distilled water and saturated solution of potassium iodide were introduced into the mixture. Subsequently, the resulting combination underwent titration against sodium thiosulfate using a starch indicator.

2.2.3. Free fatty acid analysis (FFA)

The quantification of FFA in the samples was conducted according to the AOAC method (American Oil Chemists' Society 1997). Briefly, the oil samples were mixed with an equal volume of diethyl ether: ethyl alcohol and titrated with sodium hydroxide (0.1 N).

2.2.4. p-anisidine value analysis

The measurement of the p-anisidine value followed the standard procedure specified in AOCS Cd 18-90 (American Oil Chemists' Society 2017). For this purpose, 2 g of oil samples were dissolved in isooctane (25 mL). 5 mL of the dissolved samples were mixed well with 1 mL of p-anisidine solution (2 M in acetic acid). Absorbance of the mixture was measured at 350 nm by UV-visible spectrophotometer (Perkin Elmer 710, USA).

2.2.5. Fatty acid analysis

Fatty acid composition of oil samples was conducted by gas chromatograph (GC) method according to the literature (Şahin & Özata, 2022) using a GC 2010 (Shimadzu, Tokyo, Japan) equipped with a flame ionization detector (FID). For this purpose, fatty acids were converted to fatty acids methyl esters (FAMEs). FAMEs were separated with a GC capillary column (TR-CN100 column, 60 m × 0.25 mm I.D., 0.20 µm film thickness: Technorama, Spain). The oven of GC (Shimadzu, Tokyo, Japan) was programmed from 140 °C (kept for 5 min) to 250 °C at 4 °C /min and kept at the final temperature for 15 min. The injector temperature was 250 °C and the injection volume was 1µL. The split ratio was 1:100. The carrier gas was nitrogen with a flow rate of 30 mL/min. The standard FAME (Restek, USA) was used for identification of FAMEs in the oil samples.

2.2.6. Identification of volatile substances

To determine the volatile substances of oil samples, a HS-SPME-GC-MS (headspace solid phase microextraction-gas chromatography-mass spectroscopy) method as outlined by Petersen et al. (2012a) with minor adjustments was applied. A Rxi-

5MS capillary column (30 m × 0.25 mm × 0.25 µm; Restek) was used separation of volatile compounds. 3 g of oil samples were weighed into a 10 mL vial. A carboxen/polydimethylsiloxane (CAR/PDMS) SPME fiber (Supelco, Bellefonte, PA, USA) was inserted into the vial at 40 °C for 60 min. The SPME fiber was then desorbed at 270 °C for 5 min. The carrier gas was helium with a flow rate of 0.6 mL/min. The oven temperature was initially set at 40 °C and maintained for 6 minutes. Subsequently, it was raised to 100 °C at a rate of 5 °C per minute. Following this, a more rapid increase to 300 °C was implemented at a rate of 30 °C per minute. The final temperature of 300 °C was sustained for 2 minutes. Concurrently, the ion source temperature was held constant at 230 °C, while the transfer line temperature was maintained at 250 °C. The split ratio was set at 1:5. The ionization voltage was 70 eV. The mass spectra were obtained using electron ionization (m/z 35-350). Identification of volatile compounds in the oil samples involved a comparison between the mass spectra of unidentified peaks and those present in the NIST mass spectral library (NIST MS Search Software). This comparative analysis was employed to determine the presence and characteristics of specific compounds based on their mass spectral patterns.

2.2.6. Total antioxidant capacity (TAC) analysis

The assessment of the TAC in oil samples was conducted using DPPH radical scavenging method as outlined by Şahin & Özata (2022). The oil samples were dissolved in n-butanol. 40 µL of the dissolved sample were mixed with 1500 µL of DPPH solution (0.6 mM in n-butanol). After incubation at room temperature for 30-minute, absorbance of the mixed solution was determined at 515 nm using an UV-visible spectrophotometer (Perkin Elmer 710, USA). TAC was given as millimoles of (±)-6-hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid (trolox) equivalents per liter (mmol TE/L).

2.2.7. Analysis of data

The data underwent analysis through repeated measures two-way ANOVA, employing Tukey's honest significant difference multiple comparison test with a significance level set at $P<0.05$. The statistical analysis was performed using SPSS Statistics software, version 26.0. (IBM, USA).

3. Results and Discussion

The free fatty acid (FFA) values (%) of organic hazelnut oil (OHO) and conventional hazelnut oil (CHO) are reported in Table 1. FFA values of OHO and CHO were 1.04 – 1.66% and 1.01 – 1.64%, respectively. As anticipated, FFA values showed an increase during accelerated storage, corresponding with longer storage periods. When FFA values of organic hazelnut oil and conventional hazelnut oils were compared, statistical analysis revealed a significant difference between the oils specifically on the 6th and 8th days of the storage period, while the oils had the same FFA values on the other days of storage.

Table 1- Increase in the free fatty acid (FFA) values of organic and conventional hazelnut oils under accelerated storage

	Duration of storage (in days)							
	0	*2*	*4*	*6*	*8*	*10*	*12*	*14*
OHO	1.04 ± 0.03^{aA}	1.22 ± 0.01^{bA}	1.32 ± 0.04^{cA}	1.41 ± 0.00^{cdA}	1.44 ± 0.04^{dA}	1.58 ± 0.00^{eA}	1.63 ± 0.00^{eA}	1.66 ± 0.03^{eA}
CHO	1.01 ± 0.01^{aA}	1.24 ± 0.00^{bA}	1.32 ± 0.03^{cA}	1.47 ± 0.00^{dB}	1.52 ± 0.00^{deB}	1.57 ± 0.01^{efA}	1.63 ± 0.00^{fgA}	1.64 ± 0.01^{gA}

The reported values denote means with corresponding standard deviations. Statistically significant differences ($P\leq .05$) are indicated when values in the same columns are associated with distinct capital letters or values in the same rows are marked by different lower-case letters. Abbreviations: OHO, organic hazelnut oil; CHO, conventional hazelnut oil.

Peroxide values of the oil samples are given in Figure 1. The peroxide value (PV) of OHO varied from 7.97 meq/kg to 66.66 meq/kg, while CHO had PV ranging from 12.55 meq/kg to 48.69 meq/kg. It was observed that the peroxide formation in OHO significantly increased up to 12th days of oxidation, followed by a tremendous significantly decrease on the 14th day. PV of CHO significantly increased until the 4th days of oxidation, significantly decreased on the 6th day, then unchanged between the 6th and 8th days, then significantly increased again until the 10th days, then remained unchanged between the 10th and 12th days, and significantly decreased after the 12th days. On the other hand, Yalcin (2011) reported a regular rise in PV of hazelnut oil stored at lower temperature (40 °C) up to 120 days. This unexpected decrease in PV after prolonged oxidation at high temperature observed in our study may be related to the volatilization of some lipid peroxidation products, formed in the earlier stages of oxidation (Wanasundara et al. 1994). Similar decreases in PV due to long time oxidation at high temperature were reported by Iqbal & Bhanger (2007). They stored the sunflower oil at 65 °C and observed that PV increased up to 20th days and decreased

on the 24th day. Crapiste et al. (1999) also suggested that PV in sunflower oil heated at 67 °C increased during storage, reached its maximum value on the 41st days of storage and thereafter decreased due to decomposition of lipid hydroperoxides. When they stored the oil at low temperature (37 °C and 47 °C), they observed that PV went on increasing with the increase in storage time. OHO had significantly lower PV than CHO up to 6th days of storage, while PV in OHO were significantly higher than CHO on the 8th, 10th, and 12th days of storage. The highest PV (66.67 ± 0.002 meq/kg) was observed in OHO on the 12th day of oxidation. Although OHO had the highest PV, it showed significantly lower PV (9.493 ± 0.001 meq/kg) than CHO (21.28 ± 0.019 meq/kg) at the end of storage (14th day).

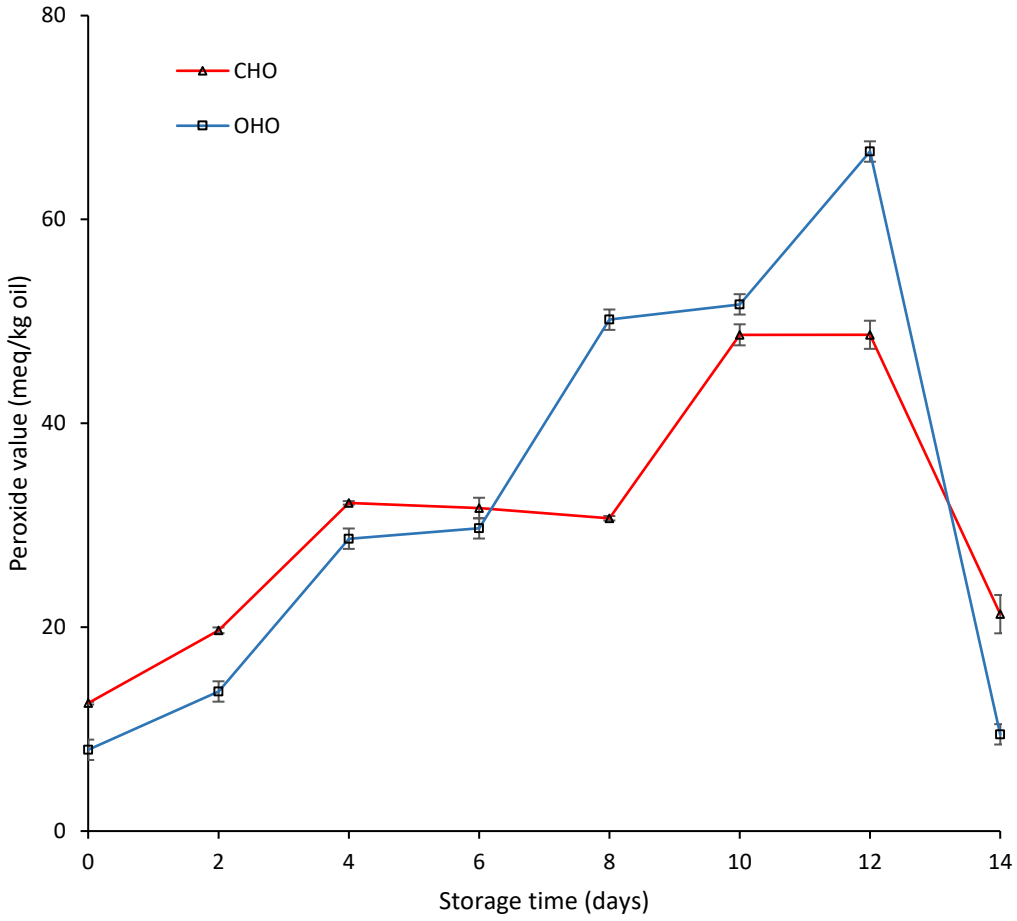


Figure 1- Increase in the peroxide values (meq/kg) of organic and conventional hazelnut oils under accelerated storage. Values are means ± (100 x standard deviation). Abbreviations: OHO, organic hazelnut oil; CHO, conventional hazelnut oil

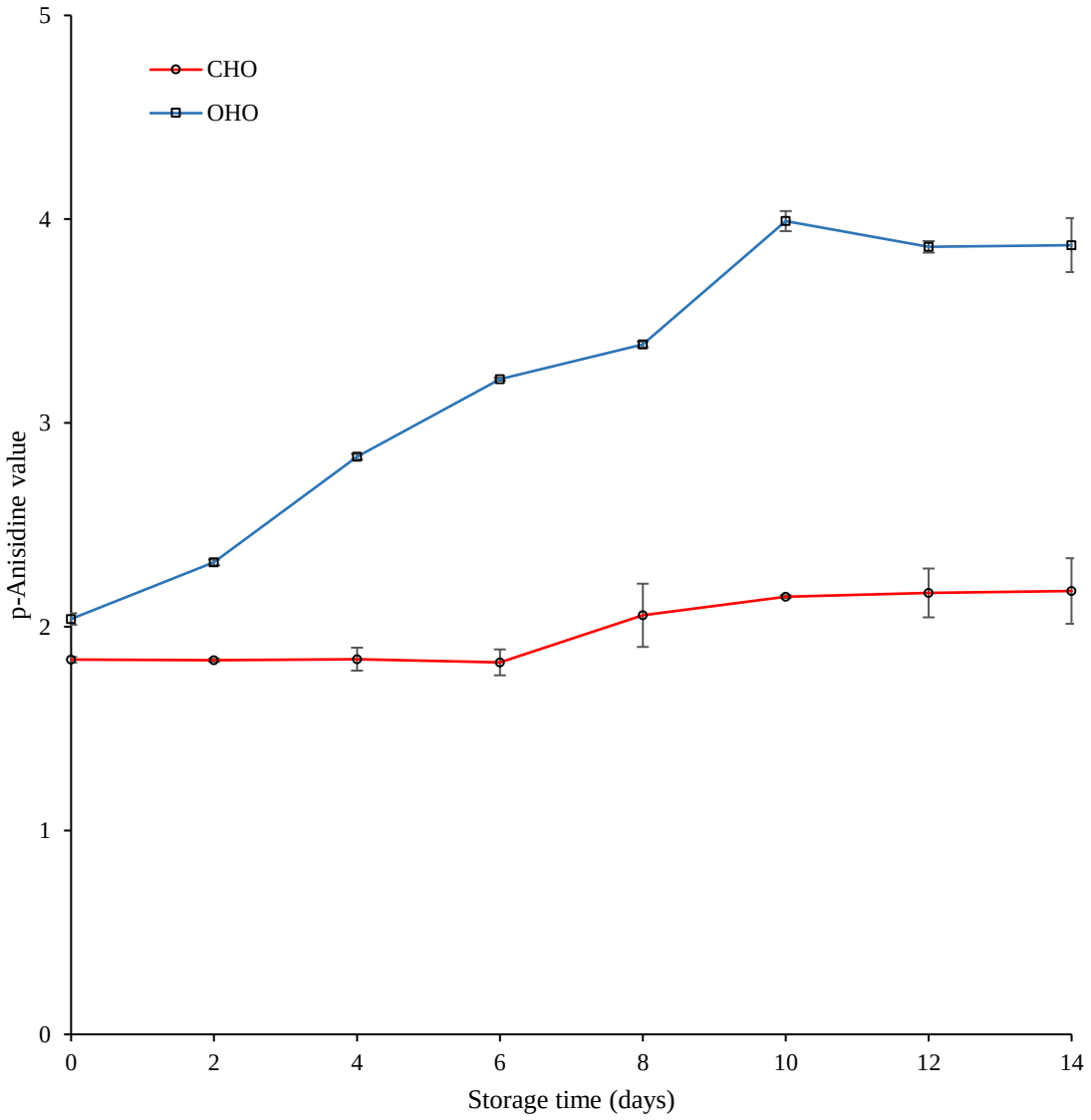


Figure 2- Relative increase in p-anisidine value (AV) of organic and conventional hazelnut oils under accelerated storage. Values are means ± (10 x standard deviation). Abbreviations: OHO, organic hazelnut oil; CHO, conventional hazelnut oil

The main secondary products of lipid peroxidation are the aldehydes such as hexanal, and propanal (Ayala et al. 2014). P-anisidine value (AV) is used as an indicator for secondary oxidation. AVs of OHO and CHO during storage are shown in Figure 2. OHO had significantly higher AVs compared to CHO during accelerated storage. In OHO, AVs significantly increased up to 10 days of storage, then significantly decreased on the 12th days of storage, and remained stable thereafter. AVs in CHO did not change until the 6th day of storage, then started to increase until the 10th days of oxidation, then unchanged until the end of storage. As observed in OHO, AVs in CHO also did not increase after 10 days of storage. Similar change in AVs during accelerated storage was reported by Petersen et al. (2012a) for sunflower oil and by Petersen et al. (2012b) for high-oleic rapeseed oil. The reduction in anisidine formation at long time oxidation may be related to the observation of Petersen et al. (2012a), who reported that some aldehydes such as hexanal, a secondary product of lipid oxidation, decomposed and decreased in conventional sunflower oil after 8 days of accelerated storage. These findings for AVs also align with the peroxide results mentioned earlier, where the peroxide values of oils increased during the initial stages of oxidation but decreased in the later stages, likely due to the breakdown of lipid oxidation products.

Antioxidative capacities of OHO and CHO during accelerated storage were given in Table 2. In terms of total antioxidant capacity, a significant difference was detected between OHO and CHO only at the beginning of storage, and no significant difference was observed throughout storage. Initially, OHO showed a higher antioxidative capacity (0.25 mmol /L TE) compared to CHO (0.20 mmol /L TE). Recently, Karaosmanoğlu (2022) reported that the hazelnuts grown under organic production system exhibited a higher antioxidative capacity than the hazelnuts grown with conventional production system and suggested that a decrease in the amounts of antioxidants of hazelnuts may be due to pesticides used in conventional production.

Table 2- Changes in total antioxidant capacity (mmol/L TEAC) of organic and conventional hazelnut oils during accelerated storage

	Duration of storage (in days)							
	0	*2*	*4*	*6*	*8*	*10*	*12*	*14*
OHO	0.25 ± 0.01^{aA}	0.15 ± 0.01^{bA}	0.10 ± 0.00^{cA}	0.07 ± 0.01^{cA}	0.05 ± 0.01^{cA}	0.05 ± 0.01^{cA}	0.06 ± 0.00^{cA}	0.05 ± 0.01^{cA}
CHO	0.20 ± 0.01^{aB}	0.13 ± 0.00^{bA}	0.12 ± 0.01^{bA}	0.06 ± 0.00^{cdA}	0.04 ± 0.00^{dA}	0.06 ± 0.01^{cdA}	0.08 ± 0.01^{cA}	0.06 ± 0.00^{cdA}

The reported values denote means with corresponding standard deviations. Statistically significant differences ($P \leq .05$) are indicated when values in the same columns are associated with distinct capital letters or values in the same rows are marked by different lower-case letters. Abbreviations: OHO, organic hazelnut oil; CHO, conventional hazelnut oil.

As can be seen in Table 2, total antioxidative capacity in OHO decreased until the 4th day of storage, then remained stable. Total antioxidative capacity of CHO decreased until the 2nd day of storage, remained constant between the 2nd and 4th days of storage, then decreased again until the 6th day and thereafter remained constant. These results mean that although the total antioxidant capacity for both OHO and CHO declined in the early stages of oxidation, it remained unchanged in the later stages of oxidation. This may be due to the fact that almost all antioxidant effective substances are used to prevent oxidation in the early stages of oxidation. The results observed here are also in agreement with the results of PV and AV. Indeed, PVs and AVs of OHO and CHO increased at the first time of exposure to oxidation, then AVs did not change, and PVs decreased.

The reported values denote means with corresponding standard deviations. Statistically significant differences ($P \leq .05$) are indicated when values in the same rows are associated with distinct lower letters Abbreviations: OHO, organic hazelnut oil; CHO, conventional hazelnut oil; TSFA, total saturated fatty acids; TMUFA, total monounsaturated fatty acids; TPUFA, total poly unsaturated fatty acids; TUFA, total unsaturated fatty acids.

As shown in Table 3, the predominant fatty acids in both hazelnut oil samples are palmitic (C 16:0), stearic (C 18:0), oleic (C 18:1 ω9), and linoleic acid (C 18:2 ω 6). Similarly, Karaosmanoglu & Ustun (2021) and Karaosmanoğlu (2022) determined these four fatty acids (palmitic, stearic, oleic and linoleic acids) as the dominant fatty acids in organic and conventional hazelnuts. In our study, oleic acid constituted the largest quantity among all fatty acids (79.74% for OHO and 81.39% for CHO), followed by linoleic acid (10.51% for OHO and 8.89% for CHO), palmitic acid (6.15% for OHO and 6.04% for CHO), and stearic acid (2.85% for both OHO and CHO).

It was observed that total monounsaturated fatty acids (TMUFA) contents in OHO and CHO were similar (80.11% and 81.78%, respectively), while OHO had higher total polyunsaturated fatty acids (TPUFA) compared to CHO (10.61% for OHO and 8.99% for CHO; Table 3). This can be explained by the fact that OHO contains similar oleic acid and more linoleic acid compared to CHO. Total saturated fatty acid content (TSFA) in OHO was the same as CHO (9.3% for OHO and 9.23% for CHO; Table 3). This result related to their similar palmitic and stearic acid contents. Similar results for TMUFA, TPUFA, and TSFA contents of organic and conventional hazelnut oils have been shown before by Karaosmanoglu & Ustun (2021) who analyzed chemical composition of commercially important hazelnut varieties such as Çakıldak, Foşa, Mincane, Palaz, Sivri and Tombul. In contrast, Karaosmanoğlu (2022) reported slightly higher TMUFA content and slightly lower TPUFA and TSFA contents for organic and conventional Tombul hazelnuts than the TMUFA, TPUFA and TSFA levels found in this work. The TUFA/TSFA ratio in OHO and CHO was determined as 9.75% and 9.83%, respectively (Table 3). These values were found to be less than the values determined by Karaosmanoğlu (2022) for organic and conventional Tombul hazelnuts (11.54 % and 12.85 %, respectively) and by Ghirardello et al. (2013) for Tonda Gentile delle Langhe hazelnuts grown in Italy (12.03 %). The reason for these differences observed in the fatty acid contents may be due to the different origins and varieties of hazelnuts examined in these studies. Indeed, it has been reported in previous studies that the fatty acid composition in hazelnuts varies depending on the origin, variety and growing conditions of the hazelnuts (Li & Parry 2011; Şahin et al. 2022; Karaosmanoğlu 2022).

Table 3- Fatty acid composition (%) of organic and conventional hazelnut oils

Fatty acids	*OHO*	*CHO*
$C_{14:0}$	0.03 ± 0.01a	0.04 ± 0.01a
$C_{15:0}$	0.01 ± 0.00a	0.02 ± 0.01a
$C_{16:0}$	6.15 ± 0.24a	6.04 ± 0.27a
$C_{16:1}\ \Delta^{9}$	0.19 ± 0.02a	0.19 ± 0.03a
$C_{17:0}$	0.04 ± 0.00a	0.05 ± 0.01a
$C_{17:1}\ \Delta^{10}$	0.06 ± 0.01a	0.06 ± 0.00a
$C_{18:0}$	2.85 ± 0.02a	2.85 ± 0.00a
$C_{18:1}\ \Delta^{9}$	79.74 ± 0.67a	81.39 ± 0.25a
$C_{18:2}\ \Delta^{9,12}$	10.51 ± 0.46a	8.89 ± 0.06b
$C_{18:3}\ \Delta^{9,12,15}$	0.10 ± 0.00a	0.10 ± 0.00a
$C_{20:0}$	0.16 ± 0.00a	0.16 ± 0.00a
$C_{20:1}\ \Delta^{11}$	0.12 ± 0.00a	0.14 ± 0.00a
$C_{22:0}$	0.02 ± 0.01a	0.03 ± 0.00a
$C_{24:0}$	0.04 ± 0.00a	0.04 ± 0.00a
TSFA	9.30	9.23
TMUFA	80.11	81.78
TPUFA	10.61	8.99
TUFA	90.72	90.77
TUFA/TSFA	9.75	9.83
Oleic/linoleic	7.59	9.15

Table 4 shows the changes of the predominant fatty acids including palmitic, stearic, oleic, and linoleic acids in the organic and conventional hazelnut samples during accelerated storage. No significant changes were observed in the amounts of predominant fatty acids in both hazelnut samples during storage. It was found that the production system (organic and conventional) did not significantly affect the palmitic and stearic acid contents. Similarly, Karaosmanoglu & Ustun (2021) did not detect statistically significant differences in palmitic and stearic acid contents between organic and conventional hazelnuts depending on the production system. In addition, Perretti et al. (2004) did not observe a significant difference between sunflower seed oils produced from conventionally cultivated crops and organically cultivated ones with regards to palmitic and stearic acid contents. In contrast, Karaosmanoğlu (2022) reported that the palmitic and stearic acid contents in hazelnuts were affected by the production system.

Table 4- Changes in the fatty acid contents (%) of organic and conventional hazelnut oils during accelerated storage

	Duration of storage (in days)							
	0	*2*	*4*	*6*	*8*	*10*	*12*	*14*
Palmitic acid								
OHO	6.15 ± 0.24^{aA}	6.19 ± 0.26^{aA}	6.08 ± 0.26^{aA}	6.03 ± 0.05^{aA}	5.83 ± 0.23^{aA}	5.97 ± 0.25^{aA}	5.72 ± 0.02^{aA}	5.98 ± 0.02^{aA}
CHO	6.04 ± 0.27^{aA}	5.79 ± 0.04^{aA}	5.72 ± 0.07^{aA}	5.67 ± 0.31^{aA}	5.68 ± 0.01^{aA}	5.81 ± 0.19^{aA}	5.60 ± 0.17^{aA}	6.12 ± 0.17^{aA}
Stearic acid								
OHO	2.85 ± 0.02^{aA}	2.83 ± 0.04^{aA}	2.81 ± 0.01^{aA}	2.82 ± 0.01^{aA}	2.81 ± 0.01^{aA}	2.84 ± 0.03^{aA}	2.87 ± 0.01^{aA}	2.93 ± 0.17^{aA}
CHO	2.85 ± 0.00^{aA}	2.89 ± 0.06^{aA}	2.87 ± 0.02^{aA}	2.79 ± 0.01^{aA}	2.82 ± 0.07^{aA}	2.88 ± 0.05^{aA}	2.89 ± 0.04^{aA}	2.93 ± 0.12^{aA}
Oleic acid								
OHO	79.74 ± 0.67^{aA}	78.34 ± 0.02^{aA}	78.69 ± 0.78^{aA}	78.32 ± 0.36^{aA}	78.64 ± 0.22^{aA}	78.97 ± 0.32^{aA}	79.43 ± 0.40^{aA}	78.37 ± 0.23^{aA}
CHO	81.39 ± 0.25^{aA}	80.89 ± 0.14^{aB}	80.63 ± 0.17^{aA}	80.62 ± 0.11^{aB}	80.88 ± 0.47^{aB}	80.60 ± 0.22^{aB}	80.79 ± 0.28^{aA}	81.10 ± 0.48^{aB}
Linoleic acid								
OHO	10.51 ± 0.46^{aB}	10.53 ± 0.14^{aB}	10.06 ± 0.32^{aB}	10.31 ± 0.56^{aB}	9.87 ± 0.09^{aB}	9.99 ± 0.01^{aB}	9.54 ± 0.31^{aB}	9.99 ± 0.40^{aB}
CHO	8.89 ± 0.06^{aA}	8.74 ± 0.02^{aA}	8.56 ± 0.32^{aA}	8.67 ± 0.01^{aA}	8.39 ± 0.44^{aA}	8.51 ± 0.43^{aA}	8.19 ± 0.27^{aA}	8.54 ± 0.36^{aA}

The reported values denote means with corresponding standard deviations. Statistically significant differences ($P \leq .05$) are indicated when values in the same columns are associated with distinct capital letters or values in the same rows are marked by different lower-case letters. Abbreviations: OHO, organic hazelnut oil; CHO, conventional hazelnut oil

Hazelnut oils are primarily composed of oleic acid (Alaşalvar et al. 2006; Karaosmanoglu & Ustun 2021; Karaosmanoğlu 2022; Şahin et al. 2022; Şahin 2023). During accelerated storage, the amount of the oleic acid in OHO varied from 78.32 to 79.74 %, while the oleic acid content in CHO ranged between 80.60 and 81.39 % (Table 4). These values are lower than the oleic acid contents (84.64% for organic hazelnut and 84.59% for conventional hazelnut) found by Karaosmanoğlu (2022) but are similar to the values reported by Karaosmanoglu & Ustun (2021). As shown in Table 4, OHO exhibited lower oleic acid than CHO at all periods of storage, but this difference observed between organic and conventional oils was not significant on the 0^{th}, 4^{th}, and 12^{th} days of storage ($P>0.05$). Similarly, it has been shown before that there was no significant difference between organic and conventional hazelnuts in terms of oleic acid content depending on the production system (Karaosmanoglu & Ustun 2021; Karaosmanoğlu 2022). Moreover, it was determined that the production system had no significant impact on the amount of oleic acid in sunflower seed oils (Perretti et al. 2004).

As can be seen from Table 4, the linoleic acid contents of OHO and CHO were observed in the range of 9.54–10.51% and 8.19-8.89%, respectively. Similar linoleic acid content (9.3-10.3% for organic hazelnuts and 8.1-10.3% for conventional hazelnuts from Çakıldak, Sivri, Foşa, Tombul, and Palaz varieties) were determined by Karaosmanoglu & Ustun (2021). However, it has been reported that organic and conventional Tombul hazelnuts had lower levels of linoleic acid than the results of this study (Karaosmanoğlu 2022). Compared to CHO, OHO exhibited higher linoleic acid at all storage periods (Table 4). Contrary to our results, Karaosmanoğlu (2022) reported that the organic Tombul hazelnuts contain lower linoleic acid (6.9%) than conventional Tombul hazelnuts (7.8%). On the other hand, it was previously reported that the amount of linoleic acid did not change depending on the production technique in hazelnuts (Karaosmanoglu & Ustun 2021) and sunflowers (Perretti et al. 2004).

Table 5- Primary volatile compounds were extracted from both non-stored and stored organic and conventional hazelnut oils, with their corresponding peak areas identified

		Area of peak x 10^3			
Retention index	*Volatile compounds*	*OHO day 0*	*OHO day 14*	*CHO day 0*	*CHO day 14*
634	Pentane	1371	5495	350	5149
648	Acetic acid	361	*	495	*
651	n-Hexane	2716	310	816	572
688	2,2-dimethyl- Hexane	893	*	132	*
696	Heptane	587	2296	277	973
797	Hexanal	2971	6150	512	6120
957	(Z)-2-Heptenal	*	11766	*	10219
980	1-Octen-3-ol	*	1316	*	1180
992	2-pentyl-furan	*	2722	*	2707
1003	Octanal	*	111	*	957
1068	(E)-2-Octenal	*	6345	*	6352
1070	6,7-Dodecanedione	*	749	*	895
1106	Nonanal	*	1378	*	1450
1166	(E)-2-Nonenal	*	1091	*	1157
1268	(E)-2-Decenal	*	114	*	97
1324	2,4-Decadienal	*	1399	*	1157
1371	2-Undecenal	*	969	*	989

Abbreviations: OHO, organic hazelnut oil; CHO, conventional hazelnut oil; *: not detectable

Table 5 shows primary volatile compounds identified in OHO and CHO samples, both non-stored and stored. The volatile compounds isolated from OHO samples were similar to the volatile compounds isolated from CHO samples both at the beginning and after storage. Six volatile compounds were identified in non-stored OHO and CHO, while 15 volatile compounds were detected in OHO and CHO at the end of storage (14.day). Among the 15 compounds found in stored OHO and CHO samples, 9 were identified as aldehydes, 3 as alkanes, 1 as ketones, 1 as alcohols and 1 as furans, which previously identified as lipid oxidation products (Petersen et al. (2012a). These results are in agreement with AVs, which indicate secondary oxidation products such as aldehydes. Indeed, AVs also increased after accelerated storage (Figure 2). (Z)-2-Heptenal, 1-octen-3-ol, 2-pentyl-furan, octanal, (E)-2-octenal, 6,7-dodecanedione, nonanal, (E)-2-nonenal, (E)-2-decenal, 2,4-decadienal and 2-undecenal were found only in autoxidized (stored) OHO and CHO samples, whereas hexanal was in both fresh (non-stored) and oxidized (stored) OHO and CHO samples. As expected, the content of hexanal in oxidized OHO and CHO samples was higher than that in non-stored OHO and CHO. Similarly, Petersen et al. (2012a) reported that non-stored sunflower and high-oleic sunflower oil samples had hexanal and that the amount of hexanal in sunflower and high-oleic sunflower oil samples increased with oxidation (storage at 80 °C for 9-day). Some of these compounds including pentanal, hexanal, octanal, 1-octen-3-ol, nonanal and (E)-2-heptenal were reported before as the important oxidative degradation products of vegetable oils (Xu et al. 2017, Mildner-Szkudlarz et al. 2003 and Petersen et al. 2012a). Mildner-Szkudlarz et al. (2003) evaluated rapeseed, sunflower, peanut, soybean, sunflower, and olive oil that stored at 60 °C for a duration of 5 days. They found that these stored oils contain more volatile compounds than non-stored oils due to oxidation. They also reported that the amount and type of volatile components detected vary depending on the type of stored oil. Some of these reported compounds were identified in this study as well. Petersen et al. (2012a) stored sunflower and high-oleic sunflower oils at 80 °C for 9 days in a drying cabinet. They identified in total 74 volatile components encompassing the volatile substances detected in our study. They also reported that these compounds are predominantly generated as secondary oxidation products through the autoxidation process of oleic and linoleic acids. Furthermore, Petersen et al. (2012b) measured the volatile compounds of rapeseed and high-oleic rapeseed oils which were stored at 40 °C for 26 days. They identified 55 volatile lipid oxidation compounds and described E, Z-2,4-decadienal, propanal, and E,E-2,4-heptadienal as the prevailing volatile compounds in both rapeseed oils.

4. Conclusions

This study evaluated effect of production system (conventional and organic) in hazelnut oil samples on quality parameters including peroxide and anisidine values, free fatty acid, fatty acid composition, antioxidant capacity, and volatile substances during of accelerated storage. Organic hazelnut oils showed similar antioxidant capacity (DPPH assay) to conventional hazelnut oil during storage despite its high initial antioxidant capacity. During accelerated storage, OHO exhibited more linoleic acid and less oleic acid than CHO. Although OHO had lower PV at the end of storage, it showed higher PVs and AVs during storage. In

conclusion, these results indicate that OHO has less oxidation stability compared to CHO. Therefore, OHO should be consumed as fresh as possible without prolonged exposure to oxidation.

Acknowledgements

The research conducted for this study received financial support under the B-2136 research projects from the Scientific Research Project Coordination Unit at Ordu University in Turkey.

References

Alasalvar C, Amaral J S & Shahidi F (2006). Functional lipid characteristics of Turkish Tombul hazelnut (*Corylus avellana* L.). *Journal of Agricultural and Food Chemistry* 54(26): 10177-10183. https://doi.org/10.1021/jf061702w

Alasalvar C & Shahidi F (2009). Tree Nuts: Composition Phytochemicals and Health Effects. Taylor & Francis Group Boca Raton London New York: CRC Press.

Anttonen M J, Hoppula K I, Nestby R, Verheul M J & Karjalainen R O (2006). Influence of fertilization, mulch color, early forcing, fruit order, planting date, shading, growing environment, and genotype on the contents of selected phenolics in strawberry (Fragaria× ananassa Duch.) fruits. *Journal of Agricultural and Food Chemistry* 54(7): 2614-2620

American Oil Chemists' Society (1997). Free fatty acids (Vol. Ca 5a–40). Association of Official Analytical Chemists.

American Oil Chemists' Society (2017). P-Anisidine value (Vol. Cd 18-90). Official Methods and Recommended Practices of the American Oil Chemists' Society

Ayala A, Muñoz M F & Argüelles S (2014). Lipid peroxidation: production, metabolism, and signaling mechanisms of malondialdehyde and 4-hydroxy-2-nonenal. *Oxidative Medicine and Cellular Longevity*, 2014

Carbonaro M, Mattera M, Nicoli S, Bergamo P & Cappelloni M (2002). Modulation of antioxidant compounds in organic vs conventional fruit (peach, *Prunus persica* L., and pear, *Pyrus communis* L.). *Journal of Agricultural and Food Chemistry* 50(19):5458-5462. doi: 10.1021/jf0202584. PMID: 12207491.

Crapiste G H, Brevedan M I & Carelli A A (1999). Oxidation of sunflower oil during storage. *Journal of the American Oil Chemists' Society* 76(12): 1437

Deutsche Gesellschaft für Fettwissenschaft (2002). Deutsche Einheitsmethoden zur Untersuchung von Fetten, Fettprodukten, Tensiden und verwandten Stoffen, Bestimmung der Peroxidzahl, Teil 1: Methode nach Wheeler, Ersetzt C-VI 6a (98)

Food and Agriculture Organization (2023). FAOSTAT [online]. Website https://www.fao.org/faostat/en/#data/QCL [accessed 23 March 2023]

Ghirardello D, Contessa CE, Valentini N, Zeppa G, Rolle L, Gerbi V & Botta R (2013). Effect of storage conditions on chemical and physical characteristics of hazelnut (*Corylus avellana* L.). *Postharvest Biology and Technology* 81: 37-43

Heinmaa L, Moor U, Põldma P, Raudsepp P, Kidmose U & Scalzo R L (2017). Content of health-beneficial compounds and sensory properties of organic apple juice as affected by processing technology. *LWT-Food Science and Technology* 85: 372-379

Iqbal S & Bhanger M I (2007). Stabilization of sunflower oil by garlic extract during accelerated storage. *Food Chemistry* 100(1): 246-254

Jimenez B, Sánchez-Ortiz A, Lorenzo M L & Rivas A (2014). Effect of organic cultivation of Picual and Hojiblanca olive varieties on the quality of virgin olive oil at four ripening stages. *European Journal of Lipid Science and Technology* 116(12): 1634-1646

Juroszek P, Lumpkin H M, Yang R Y, Ledesma D R & Ma C H (2009). Fruit quality and bioactive compounds with antioxidant activity of tomatoes grown on-farm: comparison of organic and conventional management systems. *Journal of Agricultural and Food Chemistry* 57(4): 1188-1194

Karaosmanoglu H & Ustun N S (2021). Fatty acids, tocopherol and phenolic contents of organic and conventional grown hazelnuts. *Journal of Agricultural Science and Technology* 23(1): 167–177. http:// jast.modares.ac.ir/article-23-25566-en.html

Karaosmanoğlu H (2022). Lipid characteristics, bioactive properties, and mineral content in hazelnut grown under different cultivation systems. *Journal of Food Processing and Preservation* 46(7): e16717

Krist S (2020). Hazelnut Oil. In: Vegetable Fats and Oils. Springer, Cham. https://doi.org/10.1007/978-3-030-30314-3_53

Li H & Parry J W (2011). Phytochemical compositions, antioxidant properties, and colon cancer antiproliferation effects of Turkish and Oregon hazelnut. *Food and Nutrition Sciences* 02(10): 1142-1149

Mildner-Szkudlarz S, Jeleń H H, Zawirska-Wojtasiak R & Wąsowicz E (2003). Application of headspace—solid phase microextraction and multivariate analysis for plant oils differentiation. *Food Chemistry* 83(4): 515-522

Mondelaers K, Aertsens J & Van Huylenbroeck G (2009). A meta-analysis of the differences in environmental impacts between organic and conventional farming. *British Food Journal* 111(10): 1098-1119

Perretti G, Finotti E, Adamuccio S, Della Sera R & Montanari L (2004). Composition of organic and conventionally produced sunflower seed oil. *Journal of the American Oil Chemists' Society*, 81: 1119-1123

Petersen K D, Kleeberg K K, Jahreis G & Fritsche J (2012a). Assessment of the oxidative stability of conventional and high-oleic sunflower oil by means of solid-phase microextraction-gas chromatography. *International Journal of Food Sciences and Nutrition*, 63(2): 160-169

Petersen K D, Kleeberg K K, Jahreis G, Busch-Stockfisch M & Fritsche J (2012b). Comparison of analytical and sensory lipid oxidation parameters in conventional and high-oleic rapeseed oil. *European Journal of Lipid Science and Technology* 114(10): 1193-1203

Şahin S (2023). Oxidation Stability of Hazelnut Oil Supplemented with β-Carotene During Light Exposure Using Xenon Test Instrument. *Akademik Ziraat Dergisi* 12(Özel Sayı) pp. 233-240

Şahin S, Tonkaz T & Yarılgaç T (2022). Chemical composition, antioxidant capacity and total phenolic content of hazelnuts grown in different countries. *Tekirdağ Ziraat Fakültesi Dergisi*, 19(2): 262-270

Şahin S & Özata AB (2022). Substitution of cocoa powder with hazelnut skin powder in cocoa hazelnut spreads. *Journal of Food Processing and Preservation*, 46(12), e17276.

Tarozzi A, Hrelia S, Angeloni C, Morroni F, Biagi P, Guardigli M, Cantelli-Forti G & Hrelia P (2006). Antioxidant effectiveness of organically and non-organically grown red oranges in cell culture systems. *European Journal of Nutrition*, 45(3):152-158. doi: 10.1007/s00394-005-0575-6. Epub 2005 Aug 12. PMID: 16096701.

Tüfekci F & Karataş Ş (2018). Determination of geographical origin Turkish hazelnuts according to fatty acid composition. *Food Science & Nutrition* 6:557–562

Unluturk M S, Kucukyasar S & Pazir F (2021). Classification of organic and conventional olives using convolutional neural networks. *Neural Computing and Applications*, 33(23): 16733-16744

Xu L, Yu X, Li M, Chen J & Wang X (2017). Monitoring oxidative stability and changes in key volatile compounds in edible oils during ambient storage through HS-SPME/GC–MS. *International Journal of Food Properties*, 20 (sup3), S2926-S2938.

Wanasundara U N & Shahidi F (1994). Canola extract as an alternative natural antioxidant for canola oil. *Journal of the American Oil Chemists' Society* 71: 817-822

Weibel F P, Bickel R Leuthold S & Alfoldi T (2000). Are organically grown apples tastier and healthier? A comparative field study Rusing conventional and alternative methods to measure fruit quality. *Acta Horticulturae* 7417427

Willer H, Schlatter B & Trávníček J (2023). The World of Organic Agriculture Statistics and Emerging Trends 2023, Research Institute of Organic Agriculture FiBL and IFOAM https://www.fibl.org/fileadmin/documents/shop/1254-organic-world-2023.pdf

Yalcin H (2011). Antioxidative effects of some phenolic compounds and carotenoids on refined hazelnut oil. *Journal für Verbraucherschutz und Lebensmittelsicherheit* 6: 353-358

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Deep Learning based Individual Cattle Face Recognition using Data Augmentation and Transfer Learning

Havva Eylem Polat[a], Dilara Gerdan Koc[b], Ömer Ertugrul[c], Caner Koc[b]*, Kamil Ekinci[d]

[a]*Ankara University, Faculty of Agriculture, Agricultural Structures and Irrigation Department, 06110, Diskapi, Ankara, TURKEY*

[b]*Ankara University, Faculty of Agriculture, Department of Agricultural Machinery and Technologies Engineering, Ankara, TURKEY*

[c]*Kırşehir Ahi Evran University, Faculty of Agriculture, Department of Biosystems Engineering, Kırşehir, TURKEY*

[d]*Isparta University of Applied Sciences, Faculty of Agriculture, Department of Agricultural Machinery and Technologies Engineering, Isparta, TURKEY*

ARTICLE INFO
Research Article
Corresponding Author: Caner Koc, E-mail: ckoc@ankara.edu.tr
Received: 03 July 2024 / Revised: 13 August 2024 / Accepted: 23 August 2024 / Online: 14 January 2025

Cite this article
Polat H E, Koc D G, Ertugrul Ö, Koc C, Ekinci K (2025). Deep Learning based Individual Cattle Face Recognition using Data Augmentation and Transfer Learning. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):137-150. DOI: 10.15832/ankutbd.1509798

ABSTRACT
Accurate identification of cattle is essential for monitoring ownership, controlling production supply, preventing disease, and ensuring animal welfare. Despite the widespread use of ear tag-based techniques in livestock farm management, large-scale farms encounter challenges in identifying individual cattle. The process of identifying individual animals can be hindered by ear tags that fall off, and the ability to identify them over a long period of time becomes impossible when tags are missing. A dataset was generated by capturing images of cattle in their native environment to tackle this issue. The dataset was divided into three segments: training, validation, and testing. The dataset consisted of 15 000 records, each pertaining to a distinct bovine specimen from a total of 30 different cattle. To identify specific cattle faces in this study, deep learning algorithms such as InceptionResNetV2, MobileNetV2, DenseNet201, Xception, and NasNetLarge were utilized. The DenseNet201 algorithm attained a peak test accuracy of 99.53% and a validation accuracy of 99.83%. Additionally, this study introduces a novel approach that integrates advanced image processing techniques with deep learning, providing a robust framework that can potentially be applied to other domains of animal identification, thus enhancing overall farm management and biosecurity.

Keywords: Cattle identification, Deep learning, Face detection, Smart farming

1. Introduction

Developments in technology have led to significant progress in the application of fully automated monitoring and control systems in the field of animal husbandry. In recent times, breeders have demonstrated a preference for intelligent livestock systems that constantly monitor the manner in which animals engage in reproduction, nutrition, health, comfort, and their surroundings (Džermeikaitė et al. 2023). These systems employ a variety of modelling techniques to accomplish this. Further, these systems are able to anticipate significant events such as birth and disease and then take the appropriate precautions in response to those events. Robotic, automated, and artificial intelligence-based tools are utilized in the process of breeding dairy cattle. The ability to monitor the specific requirements of each animal, make appropriate adjustments to the diet of the animals, prevent illnesses, and ultimately improve the overall health of the herd is a greater capacity that breeders possess. For the purpose of improving milk yields and animal welfare, as well as reducing methane emissions from animal waste by thirty percent, digital technological systems are being utilized in livestock enterprises, according to research conducted by scientists (Polat 2022). At the same time that automated systems reduce the amount of work that is performed by humans, they also reduce the amount of time that is spent in a shelter. Consequently, this makes it possible to manage larger herds, which ultimately leads to the development of livestock businesses that are both healthier and more profitable. In order to effectively transmit information regarding product yield and quality, intelligent livestock management makes use of electronic radio frequency identification systems, in addition to herd management software and internet connections. For the purposes of ensuring production, regulating disease, administering vaccinations, monitoring animal well-being, and managing ownership, accurate identification of cattle is essential (Allen et al. 2008). Throughout the course of history, ear tags and tattoos that were used for the purpose of identifying cattle have been prone to experiencing fading, loss, and damage. In comparison to more traditional approaches, RFID systems offer a significant number of advantages and improvements in operational efficiency. However, they also present significant risks to both security and privacy, which makes them susceptible to a variety of vulnerabilities (Awad 2016). According to Ruiz-Garcia & Lunadei (2011) and Kumar et al. (2016), the application of techniques for identifying cattle based on ear tags is a common practice in the management of livestock farms. Managing the transmission of acute diseases and understanding the progression

of diseases are both possible outcomes that can be achieved with the assistance of these techniques (Wang et al. 2010). Methods that are based on tags make use of one-of-a-kind identifiers, which may take the form of permanent markings, temporary markings, or electronic devices. According to Awad (2016), ear notching is a technique that involves the removal of a section of an animal's ear or ears, which results in a differentiated shape. Combining the positions of the ear notch on different cattle allows for the identification of specific cattle. Ear notching, on the other hand, can have a negative impact on the well-being of animals, whereas alternative methods of identification are more beneficial to the welfare of animals. According to Noonan et al. (1994), ear notching is a method that is not only limited in its ability to identify specific cattle on a farm, but it is also not feasible for accurately identifying individual cattle while working on large-scale farms. An additional disadvantage of ear tags is that they have the potential to become detached, which makes it impossible to differentiate between different animals (Wang et al. 2010; Awad 2016). There is a specific implementation of object detection known as "face detection" which accurately identifies and localizes target faces in images. At the moment, there is a significant amount of research activity in the field of computer vision that is centered on the detection of objects. According to Xu et al. (2021), this field of research makes it possible to perform more complex undertakings, such as intelligent image recognition and automated person identification. Kusakunniran and Chaivirooonjaroen (2019), Cai & Li (2013), and Xiao et al. (2022) have all reported that in recent times, machine learning and deep learning algorithms have been utilized as alternatives to the conventional methods that have been utilized in the identification of cattle. The convolutional neural network (CNN) is a method of deep learning that is extremely well-liked, as stated by Kaixuan & Dongjian (2015). CNNs have become increasingly common in recent years (Tsai et al. 2018). This surge in popularity can be attributed to the increased capacities of graphics cards. And this research emphasizes the importance of adopting face detection technologies in cattle identification, which can significantly enhance the efficiency of monitoring systems in large-scale farms, thereby reducing labor costs and minimizing human error. To summarise, the following is a list of the primary accomplishments that the current study has achieved:

1. We developed a monitoring system that enables the objective identification of particular animals in order to construct a database specifically for the purpose of recognizing the faces of cattle.

2. The findings of this research demonstrated that transfer learning results in an improvement in the ability to extract features from photographs of cattle faces. In order to improve the accuracy of transfer learning-based cattle face recognition, hyperparameter optimization was performed. Additionally, data augmentation techniques such as random flip, random rotation, random zoom, and others were utilized in order to prevent overfitting. Following the investigation, it was found that the DenseNet201 model had the highest level of performance.

3. The findings indicate that sophisticated computer vision models have been developed and are able to be utilized in the livestock industry. The system that has been proposed, which provides a variety of deployment options and the possibility of future feature enhancements, is designed to eliminate the need for a large number of wearable sensors and physical tags in the future.

2. Material and Methods

2.1. Dataset collection and preparation

The Alaca Livestock and Agriculture Enterprise, which can be found in the village of Seyran in Karacabey, Bursa (Turkey), served as the primary location for collecting data for this study. Numerous breeds, including Holstein Friesian, Montofon, and Simmental, are among those that can be discovered on the farm. Taking pictures of cattle faces both inside and outside was accomplished with the help of an RGB camera that had a resolution of 1920 x 1080 pixels (HD 1080). Creating a face image database for the purpose of cattle recognition was the reason for this action being taken. This study also emphasizes the importance of collecting diverse image data under various environmental conditions to enhance the robustness of the recognition model, thus ensuring accurate identification across different settings. First, the videos of the cattle faces were captured for the purpose of constructing the dataset. After that, the videos were converted to JPEG format by utilizing a free converter from the internet. It was discovered that the dataset initially contained a considerable number of images that were extremely similar to one another; consequently, these images were eliminated manually. In the beginning, there were 43 627 pictures of 30 cattle in the dataset; however, after the selection process was carried out manually, the number of pictures was reduced to 10,326. Following the implementation of image enhancement techniques, the final dataset was comprised of 15 000 photographs of the faces of cattle, captured from thirty different breeds of cattle (Figure 1). The dataset was divided into three sections, namely training, validation, and testing, with a ratio of 64:16:20. By employing a meticulous selection process and subsequent enhancement techniques, the study ensures that the dataset is of high quality, significantly reducing the risk of overfitting and improving model generalization. This arrangement was carried out in the following manner. Since this was the case, 9 600 images were utilized for training purposes, 2 400 images were utilized for validation, and 3 000 images were utilized for testing purposes.

Figure 1- Sample images from the training data set for cattle recognition

2.2. Image augmentation

Image augmentation employs various techniques such as image mixing, generative adversarial networks, random flip, random rotation, random zoom, random height, and random width. Incorporating advanced augmentation techniques such as GANs not only increases the diversity of the training set but also enhances the model's ability to generalize across unseen data, making it more resilient in real-world applications. Although image augmentation is commonly performed with supervision, it has diverse applications (Xu et al. 2022). Rice et al. (2020) and Schmidt et al. (2018) have identified a range of strategies for image augmentation, which include simple techniques like horizontal flipping and random cropping, as well as more advanced methods that leverage unlabeled data for semi-supervised learning. An initial task is created by employing image augmentation techniques, such as predicting the rotation angles and relative positions of image patches (Komodakis & Gidaris 2018; Doersch et al. 2015). Furthermore, augmented images that share similarity with the original can serve as positive examples for contrastive learning (Grill et al. 2020; Caron et al. 2021). Currently, the most widely used approach to improve data involves applying affine image transformations and color corrections, such as rotation, reflection, scaling (zooming in/out), and shearing. There are currently two categories of image augmentation techniques: deep neural network-based black-box methods and conventional white-box methods (Mikołajczyk & Grochowski 2018). The dataset in this study underwent random transformations such as flipping, rotation, zooming, and adjustments to its height and width (Figure 2). The study proposes the integration of semi-supervised learning methods alongside traditional augmentation to further refine the training process, enabling the model to leverage both labeled and unlabeled data efficiently.

Figure 2- Image augmentation process

The dataset also contained images of the left, full frontal, and right faces of cattle, each captured from a distinct viewpoint. Chen et al. (2022) stated that including the dataset scenarios mentioned above ensured a wide range of images, which made the dataset more complex and challenging. As a result, the proposed model became more resilient and capable of generalizing. The shuffle tool is used to input the initial data, including both the original and enhanced data, into the system in a completely random form.

2.3. Deep learning algorithms

A larger neural network with numerous layers, nodes, and activation functions is the basis of the DL based method. DL techniques have recently gained a widespread popularity and are frequently used in image classification tasks. In this study, InceptionResNetV2, MobileNetV2, DenseNet201, Xception, and NasNetLarge DL algorithms were used for the identification of individual cattle faces. The study includes a comparative analysis of multiple state-of-the-art deep learning models, providing insights into their respective strengths and weaknesses in the context of cattle face recognition. Pooling (GlobalAverage-Pooling2D), dropout (Dropout (0.2)) and dense layer were added classifier part of pre-trained TensorFlow models. In many different deep learning frameworks, InceptionResNetV2 has been extensively employed. With 164 layers, it is now widely accepted that larger networks allow for more accurate image comprehension (Bhatia et al. 2019). Using multiple convolution kernels of different sizes can improve the network's adaptability and extract more abundant features on a variety of scales, which is something that the Inception network structure takes into consideration. Furthermore, by applying the model, the Inception network structure can significantly cut down on the model's parameters, which allows it to represent model features accurately with fewer convolution kernels. As a result, the model becomes less complex (Wang et al. 2021).

2.3.1. MobileNetV2

MobileNetV2 is one of the most popular and portable CNNs. An inverted residual and a linear bottleneck are used in this network. It is designed to be used with images and has the ability to generate features and classify them. A total of 154 layers makes up MobileNetV2. Compared to other popular CNN models, MobileNetV2 uses 3.4 million parameters, which is fewer. MobileNetV2's lightweight architecture makes it particularly suitable for deployment in resource-constrained environments, such as farms where computing power may be limited. This enhances the practical applicability of the model in real-world settings. To solve the classification problem, MobileNetV2, a deep neural network, is utilized. (Sandler et al. 2018; Shahi et al. 2022).

2.3.2. DenseNet201

Utilizing a condensed network, the DenseNet201 is able to deliver highly parametric models that are simple to train and that allow for the reuse of features across multiple layers. Because of this feature reuse, performance is enhanced, and the input variety in the subsequent layer is increased. DenseNet201's ability to effectively reuse features across layers not only enhances model accuracy but also significantly reduces the computational cost, making it ideal for large-scale deployments. Within the DenseNet architecture, a straightforward connectivity pattern is utilized in order to establish direct and feed-forward connections

between all of the layers. As a consequence of this, each layer transmits its own feature maps to all of the other layers, and these layers also receive additional inputs from all of the layers that came before them (Huang et al. 2017).

2.3.3. Xception

This is an improved version of Inception-v3, which is known as Xception. Instead of the traditional convolution operation, Inception-v3 makes use of a technique known as depth-wise separable convolution. Traditional convolution is divided into two stages: spatial convolution, which handle each input channel independently, and pointwise convolution, which convolves each point using a 1×1 kernel. Depth-wise separable convolution is a method that divides traditional convolution into two stages. In order to construct its architecture, the Xception model is comprised of fourteen blocks. The depth-wise separable convolution employed by Xception significantly reduces the model's complexity while maintaining high accuracy, thus providing a balanced approach between computational efficiency and performance. The total number of depth-wise separable convolution layers among these 14 blocks is 33. Additionally, there are three common convolution layers that are distributed throughout these blocks. With the exception of the first and last blocks, each and every block is surrounded by linear residual links (Szegedy et al. 2016).

2.3.4. NasNetlarge

NASNet model employ reinforcement learning-based search strategies. It creates search space by factoring the network into cells and then further breaking it up into multiple blocks. CNN models with varying kernel sizes support a variety of commonly used operations for each block, including convolutions, max pooling, average pooling, dilated convolutions, and depth-wise separable convolutions. By utilizing a reinforcement learning-based approach, NasNetLarge can automatically optimize its architecture for cattle face recognition, potentially discovering novel structures that outperform manually designed models (Zoph et al. 2018; Punn & Agarwal 2021).

2.4. Transfer learning

Increased data sizes have resulted in improved performance of deep learning models. When confronted with a scarcity of data, conventional approaches are frequently superior to deep learning methods. The integration of transfer learning in this study allows for the leveraging of pre-trained models on large-scale datasets, thereby enhancing the model's performance in cattle face recognition even when limited labelled data is available. Traditional learning theory posits that the generalization behaviour of a learning system is contingent upon the nth training case. From this standpoint, deep learning networks exhibit the anticipated behaviour: an increase in training data results in a decrease in test errors (Poggio et al. 2018).

In order to overcome the problem of limited data, one can utilize synthetic data generation or "learning transfer" techniques to augment the dataset by transferring features. Transfer learning is a technique used to address the issue of overfitting by leveraging knowledge acquired from solving one problem and applying it to a similar problem. Adopting this approach is essential for reducing overfitting. The study further refines transfer learning by incorporating hyperparameter optimization, which tailors the model to the specific nuances of cattle face images, thereby enhancing overall accuracy. Transfer learning involves training a model on a large dataset initially, and then using the acquired weights as the starting point for training new models (Khosla & Saini, 2020).

Deep learning models have the ability to achieve zero training error, meaning they can effectively memorize the training set without compromising their ability to generalize (Srivastava et al. 2014). Experimental evidence has shown that augmenting training data by increasing its volume and diversity can effectively mitigate overfitting in modern deep learning tasks that deal with high-dimensional data. Data augmentation is a commonly employed technique that has been demonstrated through empirical evidence to alleviate overfitting. The combination of data augmentation and transfer learning in this study provides a robust framework for improving model generalization, ensuring that the model performs well across different cattle breeds and environmental conditions (DeVries & Taylor 2017; Zhang et al. 2017; Schmidt et al. 2018).

Transfer learning facilitates expedited and more effective resolution of novel problems that bear resemblance to previously addressed ones. Transfer learning differs from traditional machine learning methods by leveraging knowledge from related domains to enhance predictive modelling with diverse data patterns in the current domain. In recent times, computational intelligence has been utilized to improve the effectiveness of transfer learning methods and regulate the process of transferring knowledge in real-world systems. The study highlights the potential of computational intelligence in optimizing transfer learning processes, making the approach more adaptable to real-world scenarios in livestock management (Lu et al. 2015).

Transfer learning involves addressing two fundamental questions: "what specific knowledge should be transferred?" and "what is the most effective method for transferring this knowledge?" Diverse transfer learning algorithms facilitate the transfer of distinct types of knowledge from a source to a target domain, resulting in varied enhancements in the target domain. Identifying the best option to maximize performance improvement requires thorough research or significant expertise. According to

educational psychology, it is widely acknowledged that humans develop the ability to decide what to transfer through meta-cognitive reflection on inductive transfer learning practices (Ying et al. 2018).

2.5. Model performance metrics

The most common metrics used to measure the performance of multi-label classifiers currently include the F-score for each class, accuracy, precision, and recall. Each indicator's equation is organized as follows:

As mentioned in the second sentence, the true negative accurately predicts the negative class. The real deal: The positive class utilized in the model is accurately predicted by the true positive. A false positive happens when the model generates an inaccurate prediction regarding the positive class. Negative class prediction (FN) that was computed incorrectly. The incorrectly identified positive class is denoted by the acronym FP, the incorrectly identified negative class by the acronym FN, and the correctly identified positive class by the acronym TP. Precision, recall, and F1 score are the terms that are used to describe the actual values that the system predicts, and they are the components that make up the overall passing assessment.

The term "precision" refers to the proportion of all relevant results that are true positives (TP) or correct predictions. This proportion takes into account both true positives and false positives (FP). In tasks that involve the classification of multiple classes, the average of the classes is denoted by the letter P. In order to achieve precision, the formula is as follows.

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

Recall: False negatives and the percentage of TP from the total amount of TP (FN). Recall is averaged across all classes in problems involving multi-class classification. The recall formula is as follows.

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

The F1 score, which fully reflects the overall index of the model, is the harmonic average of precision and recall:

$$F1 = 2 * \frac{Precision*Recall}{Precision+Recall} \quad (3)$$

The most commonly utilized activation function is SoftMax, which solely considers the accuracy of classification and disregards the inter-class distance. In this study, the researchers have selected the face detection model and the latest loss functions for face identification, as described by Xu et al. (2022), to be used specifically for identifying cattle faces. The selection of SoftMax as the activation function is due to its efficiency in handling multi-class classification problems, which is essential for differentiating among numerous cattle individuals. The hyperparameters that have been chosen are listed in Table 1. These values yielded the most advantageous training results following thorough experimentation. Categorical cross-entropy is a commonly employed loss function in tasks like classification, where the output variable is a categorical variable with multiple classes (Gerdan Koc et al. 2023). The learning rate, which began at 0.01 and decreased correspondingly every 5 epochs to 0.00001, followed a precise applying schedule. The dynamic adjustment of the learning rate helps in fine-tuning the model, ensuring it converges more effectively without overshooting, leading to higher accuracy in cattle face recognition tasks. By lowering the learning rate according to a predefined schedule, learning rate schedules aim to modify the learning rate during training. The models were compared based on their performance on the testing set. The number of epochs without improvement is the patience value, after which training will be stopped. According to the training set and GPU performance, setting the batch size to 32 resulted in better model performance.

Table 1- Hyperparameters of algorithms

Parameters	***Values***
Batch size	32
Epochs	25
Momentum	0.9
Learning rate	0.00001
Metric	Categorical cross entropy
Patience	3
Optimization method	Adam
Activation function	SoftMax

3. Results

A Google Colab notebook, coded in Python, served as the platform for the whole investigation. The system specs for the computer included Windows 10, an Intel® CoreTM i7-10750H CPU, 16 GB of RAM, and a graphics processing unit from NVIDIA called an RTX 2060. Individual cattle face recognition with deep convolutional neural networks trained on augmented datasets was the primary goal of the research. The use of Google Colab with GPU acceleration significantly reduced the training time, enabling rapid experimentation and iteration on various model architectures. Through the utilization of open-source libraries, experiments were carried out with a variety of CNN models, such as InceptionResNetV2, MobileNetV2, DenseNet201, Xception, and NasNetLarge. Each of the three sections of the cattle face dataset consisted of nine thousand six hundred images: training (9,600 images), testing (2,400 images), and validation (three thousand images). The results of training each CNN model with fixed learning rates, epochs, and batch sizes were compared after the training was completed. The accuracy of the training and validation have been summarized in Table 3. By utilizing the Adam optimizer, DenseNet201 was able to achieve the highest validation accuracy possible, which was 99.83%. A validation accuracy of 98.87% was achieved by NasNetLarge, while MobileNetV2 achieved 98.54%, Xception achieved 98.88%, and InceptionResNetV2 achieved 92.54%. The DenseNet201 model's superior performance in validation accuracy highlights its robustness in generalizing from the training data, making it a prime candidate for real-world deployment in cattle face recognition systems (Table 2).

Table 2- Variation of validation accuracy and loss depending on the DL algorithms selected

Algorithms	*Validation Accuracy (%)*	*Validation Loss*
DenseNet201	99.83	0.0836
NasNetLarge	98.87	0.0971
MobilNetV2	98.54	0.1354
Xception	98.88	0.1678
InceptionResNetV2	92.54	0.3665

A value of 0.0836 was obtained for the DenseNet201 architecture, which was the CNN model that experienced the least amount of loss in validation. The low validation loss for DenseNet201 further supports its efficiency and accuracy, indicating minimal overfitting and strong performance across various data splits. Furthermore, according to the outcomes of the training, DenseNet201 was found to be the most efficient architecture with the least amount of loss. InceptionResNetV2 was found to have the highest loss, with a value of 0.3665, according to the findings. The epochs of accuracy for the models that were utilized in the experiments are displayed in Figure 3.

The individual algorithm-based results are presented in Table 3, which includes the average accuracy, macro average, weighted average, precision, recall, and F1 score values that were obtained by each of the models on the test dataset. The total amount of time spent training was determined by starting from the epoch in which the loss values of the models started to increase. This was done because early stopping was utilized during the training process. Implementing early stopping during training ensured that the models did not overfit, preserving their ability to generalize effectively to unseen data. Table 3 presents a comparison of the identification accuracy of the following algorithms: DenseNet201, MobileNetV2, Xception, NasNetLarge, and InceptionResNetV2. The results of the experiments presented in Table 3 make it abundantly clear that the DenseNet201 algorithm performed significantly better than other methods when applied to the cattle face dataset.

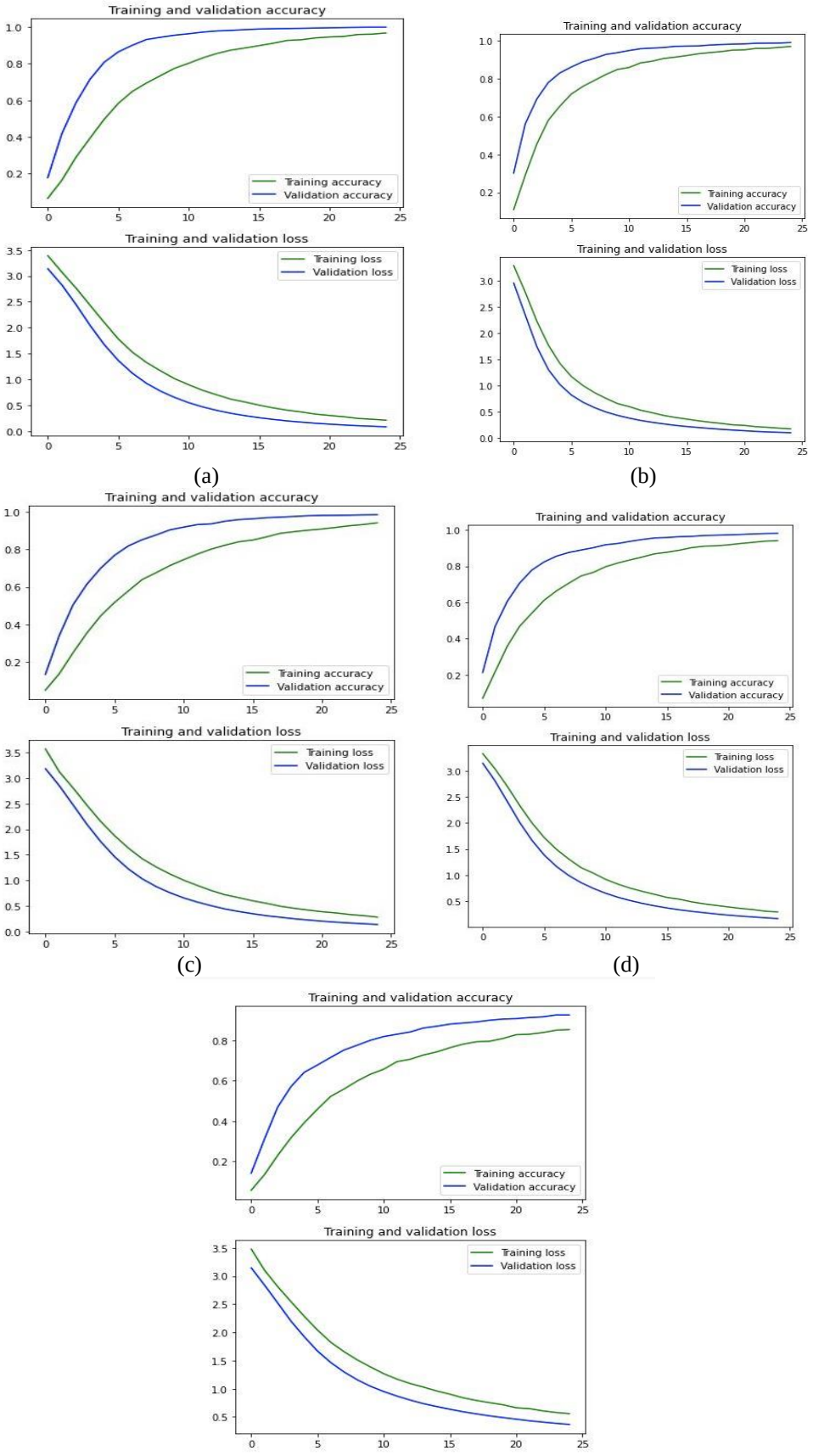


Figure 3- Training and validation accuracy and loss plots of DenseNet201 (a), NasNetLarge (b), MobilNetV2 (c), Xception (d), InceptionResNetV2 (e)

Table 3- Performance metrics and accuracy criteria of the algorithms for test results

Algorithms	*Accuracy (%)*	*Macro avg. Precision (%)*	*Weighted avg. Precision (%)*	*Macro avg. Recall (%)*	*Weighted avg. Recall (%)*	*Macro avg. F1 score (%)*	*Weighted avg. F1 score (%)*	*Total test time (min)*
DenseNet201	99.53	99.52	99.54	99.51	99.53	99.50	99.53	142.5
NasNetLarge	98.96	98.95	98.98	98.96	98.97	98.95	98.97	196.25
MobilNetV2	98.80	98.76	98.81	98.74	98. 80	98.74	98.79	44.6
Xception	98.03	98.01	98.05	97.99	98.03	97.98	98.02	85.42
Inception ResnetV2	93.60	93.42	93.62	93.46	93.60	93.36	93.53	92.1

DenseNet201 not only achieved the highest accuracy but also demonstrated consistent performance across all evaluation metrics, making it the most reliable model for cattle face recognition. According to Table 3, the architectures of the DenseNet201 and NasNetLarge models achieved the highest accuracy (99.53% and 98.96%) and precision (99.52% and 98.95%), respectively, when compared to the other deep learning models that were included in the dataset. With a weighted average precision of 99.53%, the DenseNet201 algorithm achieved the highest level of precision, while the InceptionResNetV2 algorithm achieved the lowest level of precision, which was 93.60%. Values of recall ranged from 99.51% for DenseNet201 to 93.46% for InceptionResNetV2, between the two networks. Similarly, the weighted average recall for DenseNet201 was 99.53%, while the recall for InceptionResNetV2 was only 93.60%. The superior recall of DenseNet201 indicates its ability to accurately identify cattle across a variety of conditions and environments, minimizing the likelihood of missed identifications.

After 196.25 minutes of testing, the NasNetLarge network was found to have the longest test time, while MobileNetV2 was found to have the shortest test time, which was 44.6 minutes. Following the completion of the analysis and reaching a conclusion regarding the training and validation data, the test data were analysed. The results of the test are displayed in Figure 4, which shows five outcomes that were chosen at random.

Figure 4- Results of five random test samples

The use of Convolutional Neural Networks (CNNs) is employed in order to guarantee that the network's "attention" is concentrated on the actual, distinguishing characteristics of the animal, as opposed to other regions of the image that may also contain information that is pertinent (Selvaraju et al. 2016; Hansen et al. 2018). The application of CNNs ensures that the model focuses on key distinguishing features of cattle faces, which is crucial for accurate identification, particularly in complex environments. The Grad-CAM system offers graphical explanations for the decisions made by CNN. Grad-CAM, in contrast to other methods, typically backpropagates the gradient to the final convolutional layer rather than the entire image. This results in the production of a coarse localization map that highlights significant regions of the image. Using Grad-CAM enhances model interpretability, allowing researchers and practitioners to understand which features are being used for decision-making, thus ensuring transparency in the cattle identification process. Figure 5 depicts a method that can be utilized to generate a coarse localization map for a particular class that the network has been trained on. This method is referred to as Gradient-weighted Class Activation Mapping (Grad-CAM). This image illustrates how the algorithm places significant emphasis on the face of the cattle.

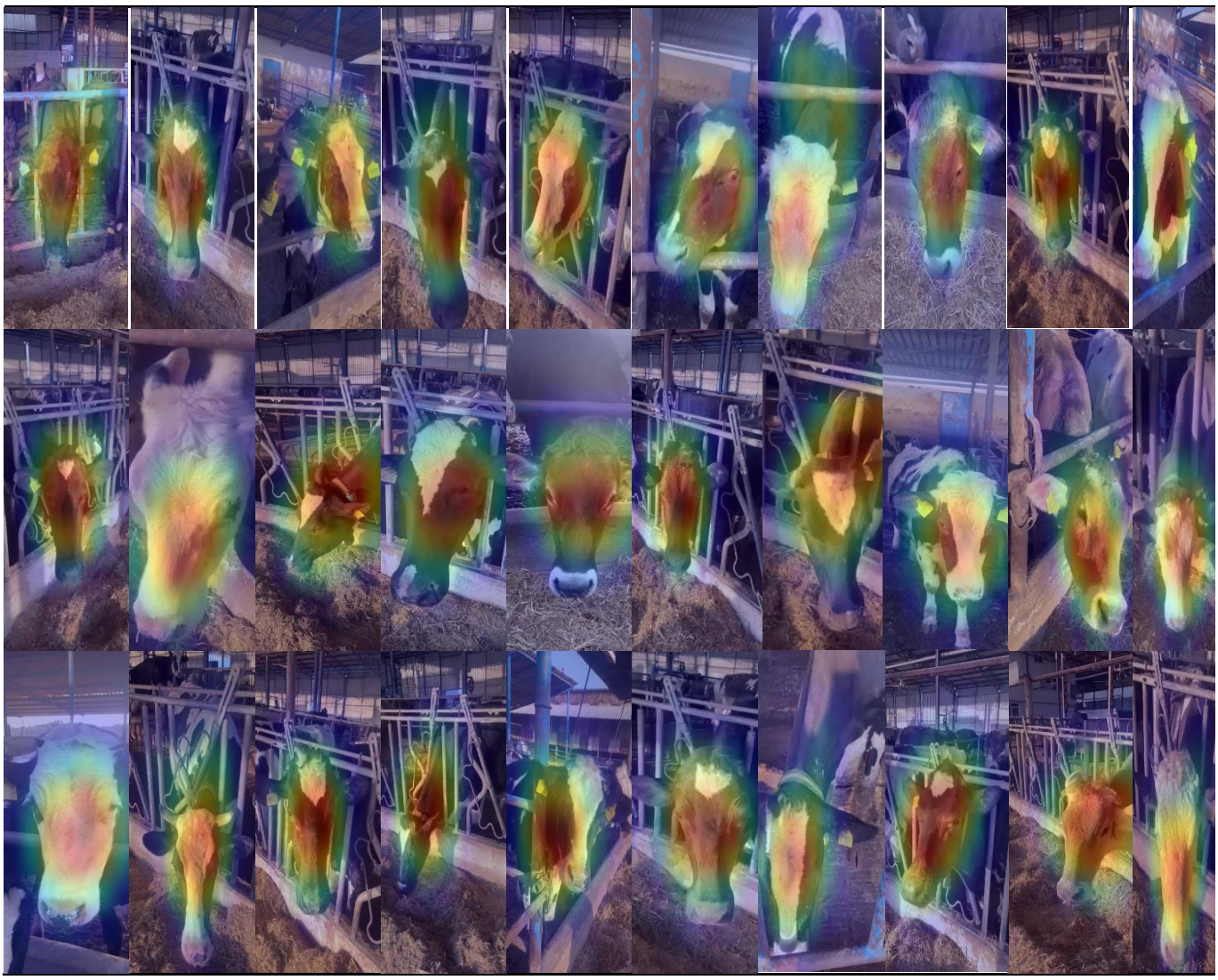


Figure 5- Gradient-weighted Class Activation Maps of cattle's

4. Discussion

A CNN-based model was developed by Qiao et al. (2021) as an alternative to RF-based ear tags for cattle identification. In their study, the researchers used photos from the top and back of the cattle to perform a deep learning analysis. The CNN model Inception-V3 was employed, and they achieved success rates of up to 88% and 99% for categorizing the rear and top photos of 41 calves, respectively. In the current study, five different deep learning methods were used along with 15 000 photos of 30 different cattle faces: InceptionResNetV2, MobileNetV2, DenseNet201, Xception, and NasNetLarge. The DenseNet201 algorithm demonstrated the highest accuracy with a score of 99.53%. Furthermore, the DenseNet201 and other algorithms for identifying livestock by their faces were more accurate at classifying them compared to images taken from the top and rear of the cattle. This finding suggests that facial features provide more reliable and consistent data points for identification than other body parts, which can vary significantly in appearance due to factors like posture, movement, and environmental conditions. Shen et al. (2020) conducted a deep learning study to classify dairy cattle using the YOLO model and the AlexNet model. In the research, 105 cattle were attempted to be classified by side views. The study findings showed that the proposed model had an accuracy of 96.65%. These results outperformed InceptionResNetV2, one of the deep learning algorithms used in our study on cattle facial recognition, in terms of accuracy (93.60%). The accuracy of the algorithms used in the current study varied from 93.60% to 99.53%. The comparative analysis with existing studies highlights the effectiveness of DenseNet201 in achieving higher accuracy, underscoring its potential as a superior model for cattle identification in practical applications.

Li et al. (2022) implemented deep learning algorithms to identify Simmental cows from images of 103 individuals. The models recommended by the researchers, including AlexNet, VGG16, MobileNetV1, SqueezeNet, and CNN, were used in the study. Further-more, AlexNet and the model recommended by the researchers both had the greatest ac-curacy of 98.86% and 98.37%, respectively. The accuracy of the CNN models used in our study on the classification of cattle was 99.53% for DenseNet201. The results showed that the accuracy obtained in our study was higher than the accuracy (98.7%) obtained by Li et al. (2022). The adopted methodologies and the variations in the face patterns of the cattle species used as test subjects might

have caused this variance. The methodology and dataset used in this study allowed for more refined and accurate cattle face recognition, which could be attributed to the enhanced image quality and diversity incorporated through data augmentation techniques.

Detecting the muzzle point patterns of cattle and classifying them using deep learning algorithms is another alternative for the classification of dairy cattle (Kumar et al. 2018). In the study, the researchers used 5 000 images collected from 500 samples that were selected using a filtering process. The accuracy levels found in the study using Deep Belief Network (DBN), Stacked Denoising Auto-encoder (SDAE), and CNN deep learning algorithms were 95.99%, 88.46%, and 75.98%, respectively. It was concluded that the DBN-based method can be used for such investigations due to its high accuracy. The classification accuracy achieved in the current study was substantially higher than that of Kumar et al. (2018), which compared the spots on the nose that were utilized to classify cattle. By focusing on facial features rather than muzzle point patterns, this study was able to achieve higher accuracy, suggesting that facial recognition may be a more reliable method for cattle identification. This can be attributed to the fact that the study was built around the cattle face im-ages that we used in our study. The points on the noses of cattle were used as a reference in the study by Kumar et al. (2018). Because they were easier to be distinguished from the photos with a point structure on the nose, images with a larger surface area and cattle faces were chosen.

Andrew et al. (2017) used drones to take upper body pictures of cattle from the air and create a data set in order to identify animals using deep learning techniques. In the study, they attempted to distinguish cattle faces based on the variations in the shape and pattern of their backs. They built their own data sets for the study using the video they captured with the drone. They discovered that using photos of 89 animals after milking resulted in an accuracy of 86.1%, whereas using images of 23 animals while they were grazing resulted in an accuracy of 98.1%. The researchers used faster R-CNN, which is a part of the Caffe deep learning package as a deep learning technique. The accuracy level detected in their study was lower than that of our study. This comparison emphasizes the importance of using high-resolution, close-up images for facial recognition, as it leads to significantly higher accuracy than aerial or distant photography. It might be argued that these variations are caused by continually moving livestock and the shooting angles used for aerial photography. The upper body shots of the cattle were taken from a distance of 5 m. Images that were taken from a distance of around 1 m were used in the samples that made up our data collection. In our study, it can be said that the separation of cattle face images is more decisive than the separation from the back photographs, which is another explanation for these disparities (Andrew et al. 2017).

Similarly, cattle face recognition was studied by Jaddou et al. (2020), who used Sup-port Vector Machine to identify the faces of 702 images of cattle and achieved a 99.00% success rate. The use of CNNs, as opposed to SVMs, allows for the handling of more complex image data and offers scalability to larger datasets, which likely contributed to the higher success rates observed in this study. In their study, Kumar et al. (2018) used 5,000 images of crossbred, Holstein Friesian, hybrid Ongole, and Balinese cattle. For cattle face determination, various methods including batch-CCIPCA, ICA, IND-CCIPCA, ISVM, LDA, LDA-LiBSVM, PCA, and PCA-LiBSVM were used. According to the findings, 95.87% of their study was successful. Lin et al. (2019) achieved 96.76% accuracy using the Fast R-CNN method with 900 images of cattle faces. Guo et al. (2022) used 3,152 Holstein Friesian cattle data points to analyze the YOLO V3-tiny. The accuracy success rate was 90.0%. This study's higher accuracy underscores the effectiveness of DenseNet201 and other advanced CNN models in providing superior performance in cattle face recognition tasks, especially when compared to older or simpler methods. The identification of Simmental and Holstein Friesian cattle faces was studied by Weng et al. in 2022. Deep image processing algorithms like VGG 16, AlexNet, GoogleNet, ResNet34, and two-branch convolutional neural networks (TB-CNN) were used for this purpose. Simmental cattle had an accuracy of 99.85%, while Holstein Friesian cattle had an accuracy of 99.81%. For detecting cattle faces, Yao et al. (2019) used the VGGNet, Inception V2, ResNet50, and Res-Net101 models. Their study had a 98.3% accuracy rate.

5. Conclusions

The fields of object detection and classification have seen significant advancements thanks to the application of deep learning principles. Increasing the gap between classes while simultaneously narrowing the gap between classes is the objective of classification. This field of machine learning for image processing has been dominated by deep learning models. The advancements that have been made in deep learning and image processing have provided an opportunity to broaden the scope of research and applications of plant disease detection and classification through the use of images. It is necessary to have models that are both quick and accurate in order to put into effect effective measures as quickly as possible. The exploration of deep learning models in animal identification, as demonstrated in this study, further expands the horizon for smart farming applications, ensuring more precise and efficient livestock management.

It is dependent on the development of deep learning for object detection and picture processing in order to replace wearable devices such as ear tags with the livestock identification system. This system also reduces the amount of harm of animals. The shift from traditional identification methods, such as ear tags, to facial recognition using deep learning, significantly reduces animal stress and improves welfare, making this approach more humane and sustainable. This study focused on the detection of faces on cattle, which is an essential component of the technology that is expected to emerge in the near future. The goal of this research was to develop a livestock machine vision system that is capable of monitoring individuals. The pretrained DenseNet algorithm, which had a 99.83% accuracy rate, was evaluated using a variety of unstructured scenes. All of these scenes were

used to evaluate the algorithm. The use of DenseNet201, with its exceptional accuracy, suggests that this model could serve as a cornerstone in future livestock monitoring systems, enabling precise individual identification even in diverse and challenging environments.

The purpose of this research was to identify the faces of cattle for the purpose of developing intelligent farming systems. A significant increase in the success rates of deep learning can be attributed to the utilization of transfer learning and data augmentation techniques. By integrating transfer learning and data augmentation, this study enhances the robustness and generalizability of the model, making it adaptable to various farming conditions and cattle breeds. The findings indicate that DenseNet performs at the highest level possible, as evidenced by its F1 score of 99.52%, its precision score of 99.50%, and its average total processing time of 142.5 minutes. Transfer learning is an essential component of deep learning because prior CNN models can be improved and retrained to perform new tasks even when there is a lack of labelled data for training. This makes transfer learning an essential part of deep learning. The implementation of transfer learning not only accelerates the training process but also mitigates the issue of overfitting, which is critical in scenarios with limited data availability. One possible factor that may influence the degree to which different deep neural networks generalize across a variety of datasets is the architecture of those networks.

Following the findings of this study, it is possible that the incorporation of robotic and early diagnosis systems on cattle farms will become feasible in the future. It is possible that the development of an infrastructure for artificial intelligence and deep learning will make it possible to quickly address a number of vital parameters pertaining to cattle, such as diseases, live weight, feed consumption, and estrus. The integration of AI and deep learning infrastructure in livestock management promises not only improved health monitoring but also optimized resource usage, contributing to more sustainable farming practices. It is therefore possible that this could make it possible to raise healthy cattle in a more efficient manner.

References

Allen A, Golden B, Taylor M, Patterson D, Henriksen D & Skuce R (2008). Evaluation of retinal imaging technology for the biometric identification of bovine animals in Northern Ireland. Livestock science 116(1-3): 42-52. DOI: https://doi.org/10.1016/j.livsci.2007.08.018

Andrew W, Greatwood C & Burghardt T (2017). Visual localisation and individual identification of holstein friesian cattle via deep learning. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2850-2859. Available from https://openaccess.thecvf.com/content_ICCV_2017_workshops/w41/html/Andrew_Visual_Localisation_and_ICCV_2017_paper.html

Awad A I (2016). From classical methods to animal biometrics: A review on cattle identification and tracking. Computers and Electronics in Agriculture, 123(2016): 423-435, https://doi.org/10.1016/j.compag.2016.03.014

Bhatia Y, Bajpayee A, Raghuvanshi D & Mittal H (2019). Image captioning using Google's inception-resnet-v2 and recurrent neural network. In 2019 Twelfth International Conference on Contemporary Computing (IC3), IEEE Publish-ing, pp. 1-6. DOI: https://doi.org/10.1109/IC3.2019.8844921

Cai C & Li J (2013). Cattle face recognition using local binary pattern descriptor. In 2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, IEEE Publishing pp. 1-4. DOI: https://doi.org/10.1109/APSIPA.2013.6694369

Caron M, Touvron H, Misra I, Jégou H, Mairal J, Bojanowski P & Joulin A (2021). Emerging properties in self-supervised vision transformers. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 9650-9660. DOI: https://doi.org/10.48550/arXiv.2104.14294

Chen X, Yang T, Mai K, Liu C, Xiong J, Kuan Y & Gao Y (2022). Holstein Cattle Face Re-Identification Unifying Glob-al and Part Feature Deep Network with Attention Mechanism. Animals, 12(8), DOI: https://doi.org/10.3390/ani12081047

DeVries T & Taylor G W (2017). Improved regularization of convolutional neural networks with cutout. arXiv preprint arXiv:1708.04552. DOI: https://doi.org/10.48550/arXiv.1708.04552

Doersch C, Gupta A & Efros A A (2015). Unsupervised visual representation learning by context prediction. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2015, pp. 1422-1430. Available from https://www.cv-foundation.org/openaccess/content_iccv_2015/html/Doersch_Unsupervised_Visual_Representation_ICCV_2015_paper.html

Džermeikaitė K, Bačėninaitė D & Antanaitis R (2023). Innovations in Cattle Farming: Application of Innovative Technologies and Sensors in the Diagnosis of Diseases. Animals, 13(5): 780

Fosgate G T, Adesiyun A A & Hird D W (2006). Ear-tag retention and identification methods for extensively managed water buffalo (Bubalus bubalis) in Trinidad. Preventive veterinary medicine, 73(4): 287-296. DOI: https://doi.org/10.1016/j.prevetmed.2005.09.006

Gerdan Koc D, Koc C &Vatandas M (2023). Diagnosis of tomato plant diseases using pre-trained architectures and a proposed convolutional neural network model. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi)* 29(2): 627-638. doi.org/10.15832/ankutbd.957265

Grill J B, Strub F, Altché F, Tallec C, Richemond P, Buchatskaya E & Valko M (2020). Bootstrap your own latent-a new approach to self-supervised learning. Advances in neural information processing systems, 33, pp: 21271-21284. ISBN: 9781713829546. Available from https://proceedings.neurips.cc/paper/2020/hash/f3ada80d5c4ee70142b17b8192b2958e-Abstract.html

Guo S S, Lee K H, Chang L, Tseng C D, Sie S J, Lin G Z & Lee T F (2022). Development of an Automated Body Temperature Detection Platform for Face Recognition in Cattle with YOLO V3-Tiny Deep Learning and Infrared Thermal Imaging. Applied Sciences, 12(8), DOI: https://doi.org/10.3390/app12084036

Hansen M F, Smith M L, Smith L N, Salter M G, Baxter E M, Farish M & Grieve B (2018). Towards on-farm pig face recognition using convolutional neural networks. Computers in Industry 98: 145-152

Huang G, Liu Z, van der Maaten L &Weinberger K Q (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition. arXiv. pp: 4700-4708. DOI: https://doi.org/10.48550/arXiv.1608.06993

Jiang B, Wu Q, Yin X, Wu D, Song H & He D (2019). FLYOLOv3 deep learning for key parts of dairy cow body detection. Computers and Electronics in Agriculture 166 (2019), DOI: https://doi.org/10.1016/j.compag.2019.104982

Kaixuan Z & Dongjian H (2015). Recognition of individual dairy cattle based on convolutional neural networks. Transactions of the Chinese Society of Agricultural Engineering, 31(5): 181-187. Available from https://www.cabdirect.org/cabdirect/abstract/20153218172

Kang X, Zhang X D & Liu G (2020). Accurate detection of lameness in dairy cattle with computer vision: A new and individualized detection strategy based on the analysis of the supporting phase. Journal of dairy science, 103 (11): 10628-10638. DOI: https://doi.org/10.3168/jds.2020-18288

Khosla C & Saini B S (2020). Enhancing performance of deep learning models with different data augmentation techniques: A survey. In 2020 International Conference on Intelligent Engineering and Management (ICIEM), IEEE, pp. 79-85, DOI: https://doi.org/10.1109/ICIEM48762.2020.9160048

Kumar S, Pandey A, Satwik K S R, Kumar S, Singh S K, Singh A K &Mohan A (2018). Deep learning framework for recognition of cattle using muzzle point image pattern. Measurement, 116 (2018), pp: 1-17, DOI: https://doi.org/10.1016/j.measurement.2017.10.064

Kumar S, Singh S K, Dutta T & Gupta H P (2016). A fast cattle recognition system using smart devices. MM '16: Proceedings of the 24th ACM international conference on Multimedia, October 2016, pp: 742–743, DOI: https://doi.org/10.1145/2964284.2973829

Kumar S, Singh S K, Singh R & Singh A K (2017). Recognition of Cattle Using Face Images. In: Animal Biometrics. Springer, Singapore. DOI: https://doi.org/10.1007/978-981-10-7956-6_3

Kusakunniran W & Chaiviroonjaroen T (2018). Automatic cattle identification based on multi-channel lbp on muzzle images. In 2018 International Conference on Sustainable Information Engineering and Technology (SIET), IEEE Publishing, pp:1-5. DOI: https://doi.org/10.1109/SIET.2018.8693161

Li G, Erickson G E & Xiong Y (2022). Individual Beef Cattle Identification Using Muzzle Images and Deep Learning Techniques. Animals, 12(11), DOI: https://doi.org/10.3390/ani12111453

Li Z, Lei X & Liu S (2022). A lightweight deep learning model for cattle face recognition. Computers and Electronics in Agriculture, 195 (2022), DOI: https://doi.org/10.1016/j.compag.2022.106848

Lu J, Behbood V, Hao P, Zuo H, Xue S & Zhang G (2015). Transfer learning using computational intelligence: A survey. Knowledge-Based Systems, 80 (2015), pp: 14-23, https://doi.org/10.1016/j.knosys.2015.01.010

Mikołajczyk A & Grochowski M (2018). Data augmentation for improving deep learning in image classification problem. In 2018 international interdisciplinary PhD workshop (IIPhDW), IEEE publishing, pp. 117-122, Poland. DOI: https://doi.org/10.1109/IIPHDW.2018.8388338

Noonan G J, Rand J S, Priest J, Ainscattle J & Blackshaw J K (1994). Behavioural observations of piglets undergoing tail docking, teeth clipping and ear notching. Applied Animal Behaviour Science, 39(3-4), 203-213. DOI: https://doi.org/10.1016/0168-1591(94)90156-2

Poggio T, Kawaguchi K, Liao Q, Miranda B, Rosasco L, Boix X & Mhaskar H (2018). Theory of deep learning III: ex-plaining the non-overfitting puzzle. arXiv preprint arXiv:1801.00173. DOI: https://doi.org/10.48550/arXiv.1801.00173

Polat H E (2022) New Technologies in Good Agricultural Practices – Smart Farming (In Turkish). In: Yaldız G, Çamlıca M (Eds.), Innovative Approaches in Medicinal and Aromatic Plants Production. Iksad Publications, Ankara/Turkey pp: 27- 54. ISBN:978-625-8246-33-9

Psota E T, Luc E K, Pighetti G M, Schneider L G, Fryxell R T, Keele J W & Kuehn L A (2021). Development and validation of a neural network for the automated detection of horn flies on cattle. Computers and Electronics in Agriculture, 180 (2021), DOI: https://doi.org/10.1016/j.compag.2020.105927

Punn N S & Agarwal S (2021). Automated diagnosis of COVID-19 with limited posteroanterior chest X-ray images using fine-tuned deep neural networks. Applied Intelligence, 51(5), 2689-2702

Qiao Y, Clark C, Lomax S, Kong H, Su D & Sukkarieh S (2021). Automated Individual Cattle Identification Using Video Data: A Unified Deep Learning Architecture Approach. Front. Anim. Sci., 2 (2021), https://doi.org/10.3389/fanim.2021.759147.

Rice L, Wong E & Kolter Z (2020). Overfitting in adversarially robust deep learning. Proceedings of the 37 th International Conference on Machine Learning, Vienna, Austria, PMLR 119, 2020. pp. 8093-8104. http://proceedings.mlr.press/v119/rice20a

Ruiz-Garcia L & Lunadei L (2011). The role of RFID in agriculture: Applications, limitations and challenges. Computers and Electronics in Agriculture, 79(1), 42-50. DOI: https://doi.org/10.1016/j.compag.2011.08.010

Sandler M, Howard A, Zhu M, Zhmoginov A & Chen L C (2018). MobilenetV2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4510–4520. DOI: https://doi.org/10.48550/arXiv.1801.04381

Schmidt L, Santurkar S, Tsipras D, Talwar K & Madry A (2018). Adversarially robust generalization requires more data. In Advances in Neural Information Processing Systems, pp. 5014–5026. https://proceedings.neurips.cc/paper/2018/hash/f708f064faaf32a43e4d3c784e6af9ea-Abstract.html

Selvaraju R R, Das A, Vedantam R, Cogswell M, Parikh D & Batra D (2016). Grad-CAM: Why did you say that? arXiv preprint arXiv:1611.07450

Shahi T B, Sitaula C, Neupane A & Guo W (2022). Fruit classification using attention-based MobileNetV2 for industrial applications. Plos one, 17(2), DOI: https://doi.org/10.1371/journal.pone.0264586

Shen W, Hu H, Dai B Wei X, Sun J, Jiang L & Sun Y (2020). Individual identification of dairy cows based on convolutional neural networks. Multimed Tools Appl (79) 14711–14724, https://doi.org/10.1007/s11042-019-7344-7

Srivastava N, Hinton G, Krizhevsky A, Sutskever I & Salakhutdinov R (2014). Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research, 15(1):1929–1958, DOI: https://dl.acm.org/doi/abs/10.5555/2627435.2670313

Szegedy C, Vanhoucke V, Ioffe S, Shlens J & Wojna Z (2016). Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2818-2826. https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Szegedy_Rethinking_the_Inception_CVPR_2016_paper.html

Tsai H, Ambrogio S, Narayanan P, Shelby R M & Burr G W (2018). Recent progress in analog memory-based accelerators for deep learning. Journal of Physics D: Applied Physics, 51(28), DOI: https://doi.org/10.1088/1361-6463/aac8a5

Wang H, Qin J, Hou Q & Gong S (2020). Cattle face recognition method based on parameter transfer and deep learning. In Journal of Physics: Conference Series, 1453 (2020), IOP Publishing, DOI: https://doi.org/10.1088/1742-6596/1453/1/012054

Wang J, He X, Faming S, Lu G, Cong H & Jiang Q (2021). A Real-Time Bridge Crack Detection Method Based on an Improved Inception-Resnet-v2 Structure. IEEE Access, (9) pp: 93209-93223. DOI: https://doi.org/10.1109/ACCESS.2021.3093210

Wang Z, Fu Z, Chen W & Hu J (2010). A RFID-based traceability system for cattle breeding in China. International Conference on Computer Application and System Modeling (ICCASM 2010), IEEE, Taiyuan. DOI: https://doi.org/10.1109/ICCASM.2010.5620675

Weng Z, Meng F, Liu S, Zhang Y, Zheng Z & Gong C (2022). Cattle face recognition based on a Two-Branch convolutional neural network. Computers and Electronics in Agriculture, 196 (2022), DOI: https://doi.org/10.1016/j.compag.2022.106871

Xiao J, Liu G, Wang K & Si Y (2022). Cow identification in free-stall barns based on an improved Mask R-CNN and an SVM. Computers and Electronics in Agriculture, 194 (2022), DOI: https://doi.org/10.1016/j.compag.2022.106738

Xu B, Wang W, Guo L, Chen G, Li Y, Cao Z & Wu S (2022). CattleFaceNet: A cattle face identification approach based on Retina Face and ArcFace loss. Computers and Electronics in Agriculture, 193 (2022), DOI: https://doi.org/10.1016/j.compag.2021.106675

Xu B, Wang W, Guo L, Chen G, Wang Y, Zhang W & Li Y (2021). Evaluation of deep learning for automatic multi-view face detection in cattle. Agriculture, 11(11), https://doi.org/10.3390/agriculture11111062

Xu M, Yoon S, Fuentes A & Park D S (2023). A comprehensive survey of image augmentation techniques for deep learning. Pattern Recognition, 137 (2023) DOI: https://doi.org/10.1016/j.patcog.2023.109347

Yao L, Liu H, Hu Z, Kuang Y, Liu C & Gao Y (2019). Cow face detection and recognition based on automatic feature extraction algorithm. ACM TURC '19: Proceedings of the ACM Turing Celebration Conference—China, May 2019, Article No.: 95, Pp 1–5, https://doi.org/10.1145/3321408.3322628

Ying W, Zhang Y, Huang J & Yang Q (2018). Transfer learning via learning to transfer. In International Conference on Machine Learning, pp. 5085-5094. Available from https://proceedings.mlr.press/v80/wei18a.html.

Zhang C, Bengio S, Hardt M, Recht B & Vinyals O (2016). Understanding deep learning requires rethinking generalization. arXiv preprint arXiv:1611.03530. DOI: https://doi.org/10.48550/arXiv.1611.03530

Zhang J, Sun G, Zheng K & Mazhar S (2020). Pupil detection based on oblique projection using a binocular camera. IEEE Access, vol: 8, pp: 105754-105765, DOI: https://doi.org/10.1109/ACCESS.2020.3000063

Zoph B, Vasudevan V, Shlens J & Le Q V (2018). Learning transferable architectures for scalable image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 8697-8710. DOI: https://doi.org/10.48550/arXiv.1707.07012

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Assessment of the Effects of Organic Fertilizer Applications on the Biochemical Quality of Basil

Tuğba Özbucak[a]*, Meltem Ocak[a], Melek Çol Ayvaz[b], Ömer Ertürk[a]

[a]*Ordu University, Department of Molecular Biology and Genetics, Ordu, TÜRKIYE*

[b]*Ordu University, Department of Chemistry, Ordu, TÜRKIYE*

ARTICLE INFO

Research Article

Corresponding Author: Tuğba Özbucak, E-mail: tsiozbucak@hotmail.com

Received: 16 May 2024 / Revised: 07 August 2024 / Accepted: 28 August 2024 / Online: 14 January 2025

Cite this article

Özbucak T, Ocak M, Çol Ayvaz M, Ertürk Ö (2025). Assessment of the Effects of Organic Fertilizer Applications on the Biochemical Quality of Basil. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):151-160. DOI: 10.15832/ankutbd.1485365

ABSTRACT

It is important to characterize the biochemical potential of medicinal and aromatic plants, which have significant therapeutic and commercial value. However, fertilizers obtained from natural resources are used in sustainable agricultural practices. In this study, basil (*Ocimum basilicum* L.) was grown from seed with different doses of barnyard manure, vermicompost and chicken manure were added. This work aimed was to investigate and compared the effect of these different fertilizer doses applications on the biochemical potential of basil plants.

Total phenolic content, antioxidant capacity and antimicrobial activity values were statistically significant according to the type and doses of fertilizer studied. Linalool was found to be the highest amount of volatile compound in leaf samples of manure treated plants. The highest antioxidant capacity values were determined in the samples where farm manure and chicken manure were applied at low (2.5%) and medium (20%) doses and worm manure was applied at high doses. The highest antibacterial effect was detected in the essential oil extract of leaf samples with 10% vermicompost against *Bacillus cereus*. Among the Gram (-) bacteria, the highest antibacterial effect against *E. coli* was determined in the essential oil extract of leaf samples with 20% farm manure. The essential oil extract of leaf samples with 20% farm manure also showed significant and high degree of antifungal effect against *S. cerevisae*.

Keywords: Basil, Manure, Antioxidant capacity, Antimicrobial activity

1. Introduction

The use of excessive amounts of chemical inputs in agriculture to meet the increasing need for food due to the rapid increase in population has become an important ecological problem. Especially in developing countries, the need to obtain more products from a unit area due to malnutrition, famine and economic reasons has led to the use of more chemical inputs. This situation negatively affects the deterioration of the ecological balance and the health of nature and the living creatures living there. For this reason, in recent years, there has been an increase in sustainable ecologically based practices that are in harmony with nature, where high- and high-quality products are obtained from unit area. Especially the interest in medicinal and aromatic plants has increased the need for quality development of these plants. The quality and quantity of active substances in plants may vary according to the genetic characteristics of the plant, climatic conditions and agronomic processes applied. Among these agronomic processes, organic farming, which is at the forefront today, is practiced in almost all countries.

O. basilicum is a plenty popular and thoroughly grown herb global. Organic and inorganic fertilization investigations have been carried out regarding cultivation, breeding, large, herbage efficiency and large essential oil content. *O. basilicum* L., also known as sweet basil, is one of the most economically important aromatic plants belonging to the Lamiaceae family (Baczek at. al. 2019; Sharma et al. 2021) that is native to India and Pakistan but is cultivated as a short-lived annual crop worldwide (Tenore et. al. 2017; Nadeem et al. 2022). The fresh and dried herbs are used as a spice in food and its essential oil in perfume industries for commercial products (Murillo-Amador et al.2013; Santos et al. 2016). There are about 900 medicinal and aromatic plants cultivated for commercial purposes worldwide (Acıbuca & Budak 2018). The sweet basil is in high demand in the international market (Egata 2021). Basil is mainly produced in accordance with conventional agriculture. However, its cultivation in organic systems seems to be better adjusted to consumer demands connected with the lack of pesticide residues in foods and their safety. The amount of basil produced organically in Turkey in 2015 was 6 tons (Arslan et al. 2015). The use of human and environmentally friendly production systems in fertilization studies is very important in correcting the natural balance that has been disturbed by wrong methods (Tsvetkov et al. 2018; Feledyn-Szewczyk et al. 2020). It was stated that cultivation of medicinal and aromatic plants with natural, environmentally friendly practices can improve product quality (Rao et al. 2022). It

has been reported that organic poultry fertilization at an appropriate rate to the soil increases antibacterial content, antioxidant activity, total phenolics, flavonoids and essential oil components in basil (Yaldız et al. 2019 a).

Organic fertilizers from barn, goat, sheep, poultry manures are usually preferred in environmentally friendly fertilizer studies (Yolcu 2011; Loss et al. 2020; Abou-Sreea et al. 2021). Biological and organic fertilizer applications are accepted as alternative methods to chemical fertilizers in sustainable agriculture (Dehghani-Samani et al. 2021; Özbucak & Alan 2024). Organic fertilizers are not only beneficial for plant growth, but also a nutrient-rich environment for later plants (Bernal et al. 2009). The poultry manure application has a positive effect on fresh weight, dry weight and nutrient content development of sweet basil (Yaldız et al. 2019 b). It was reported that different types and doses of vermicompost applications positively affected the product and quality characteristics of lettuce plants (Özbucak & Alan 2024).

Since medicinal and aromatic plants, like other plants, are very sensitive to environmental conditions, great differences are observed between the essential oil ratios and the contents of the main components of the same species of plants growing in different ecologies (Mohamed & Alotaibi 2023). This causes changes in the biological and chemical components of the plant such as quality and quantity of essential oil, antimicrobial and antioxidant properties (Bistgani et al. 2018; Yaldız et al. 2019). It is known that *O. basilicum* L. species contains different chemotypes according to the ratio of essential oil component (Raina & Gupta 2018; Shahrajabian et al. 2020).

This is the first report comparing the effects of vermicompost, farm manure, and chicken manure applications on bioactive components of basil plants grown in greenhouse. In addition, it was tried to determine the appropriate form and dose of organic fertilizer, which has been widely used in recent years, for the growth of basil plants. Therefore, it will undoubtedly be important to determine the potential for sustainable agricultural production by comparing the effects of different organic fertilizer applications on the same plant.

2. Material and Methods

2.1. Material and experimental design

The study was carried out as a pot trial with 3 replications in a plastic greenhouse in 2021. Basil seeds were sown in viols containing peat: perlite at a ratio of (3:1). The first germination was started in 3-4 days. After about 30 days, the seedlings were planted in 3 kg pots filled with soil in a plastic greenhouse. 36 pots were used in the study and two seedlings were planted in each pot. Growing media to which barnyard manure, vermicompost and chicken manure were added at different doses were used in the study. 10%, 20% and 30% doses of vermicompost and barnyard manure, 2.5%, 5% and 10% doses of chicken manure were applied. Soil without fertilizer application was used as control group. At the beginning of the experiment, 300-600-900 g of barnyard manure and vermicompost and 75-150-225 gr chicken manure were added to pots. The pots were placed in the greenhouse to receive the same amount of light. All pots were watered once/twice a week as needed and common tap water was used for irrigation. The plant samples were collected during flowering period

Barnyard, chicken and vermicompost were purchased from a commercial company. Soil analysis was performed at the Tekirdağ Commodity Exchange Analysis Laboratory. According to the analysis the soil used in this study had acidic pH (5.90), humidity 48%, 0.87% organic matter, 0.52 dS/m electrical conductivity, 43 ppm potassium, 3.60 kg/da phosphorus, 3974.25 ppm calcium, 866 ppm magnesium.

Plant samples were dried at 25 °C and grinding using a blender for preparation of methanol and water extracts. Essential oil was obtained from the dry samples obtained in sufficient quantities by method water distillation method in Clevenger apparatus (Karaca et al. 2017). Antimicrobial activity of these oil samples was determined. Volatile organic compounds in dry leaf samples were determined by Shimadzu model - QP2010 Ultrain brand GC-MS (gas chromatography/mass spectrometry).

2.2. Antioxidant capacity

Total phenolic contents (TPC) of water and methanol extracts of *O. basilicum* leaves were determined by Folin-Ciocalteu method (Singleton & Rossi 1965). DPPH free radical scavenging activity test was performed using 1,1-diphenyl-2-picrylhydrazyl (DPPH) radical in methanol (Blois, 1958). The reducing antioxidant power was calculated following the method based on the reduction of the Fe(III)-TPTZ (2,4,6-tris(2-pyridyl)-S-triazine) complex to form the blue colored Fe(II)-TPTZ complex depending on the antioxidant content of the tested samples (Benzie & Strain 1999).

2.3. Antimicrobial activity

The antimicrobial activities of the studied samples were tested in vitro by paper disk diffusion method. For bactericidal and fungicidal studies, MHA, and PDA media, respectively, were used. MHA medium was used for bactericidal studies and PDA medium was used for fungicidal studies. Known antibiotic such as Gentamicin and Nystatin were used as a standard drug for Microorganism. Antimicrobial activity was measured according to method followed by Ronald (1990).

2.4. Statistical analysis

The Kolmogorov-Smirnov test was used to test whether the data obtained in the study were normally distributed. The homogeneity of the variances of the groups was determined using Levene's test statistic. Considering the existing assumptions, the differences between the group averages were revealed using two-way and three-way ANOVA models. All the results obtained in the study were evaluated under 5% significance level and analyses were carried out through SPSS v24 and R package program.

3. Results and Discussion

The data of volatile component analyses of the leaf parts of basil (*O. basilicum*) grown in different organic fertilizer environments are shown in Table 1. Three volatile organic compounds were identified with a composition percentage above 10 to the control group. These are respectively; linalool (27.28%), Bergamotene <alpha-trans-> (22.19%) and eugenol (14.75%). The highest amounts of Linalool (48.45%-41.08%) were found in the samples where chicken manure was applied at low and medium doses to the control and other groups. However, the amounts of Bergamotene <alpha-trans-> (11.29%-10.44%) in the same doses of chicken manure applications were lower than the control and other groups. Analysis could not be performed due to insufficient sample at 10% dose. Linalool, Bergamotene <alpha-trans-> and Eugenol were the highest volatile compounds among others at all doses of barnyard and vermicompost treatments.

Table 1- Distribution of volatile organic compounds of plant leaf samples

Application	*Composition %*	*Volatile organic compounds*
	27.28	Linalool
Control	22.19	Bergamotene <alpha-trans->
	14.75	Eugenol
Low	48,45	Linalool
	11,29	Bergamotene <alpha-trans->
(Chicken) Medium	41.08	Linalool
	10.44	Bergamotene <alpha-trans->
Low	30.19	Linalool
	18.86	Bergamotene <alpha-trans->
	14.12	Eugenol
(Barn) Medium	34.75	Linalool
	14.47	Bergamotene <alpha-trans->
	12.01	Eugenol
High	38.13	Linalool
	15.54	Bergamotene <alpha-trans->
	10.87	Eugenol
Low	28.91	Linalool
	16.47	Bergamotene <alpha-trans->
	15.25	Eugenol
(Worm) Medium	33.81	Linalool
	17.85	Bergamotene <alpha-trans->
	10.98	Eugenol
High	39.21	Linalool
	17.2	Bergamotene <alpha-trans->

In addition to the quality and yield of medicinal and aromatic plants, the type, and the number of active substances they contain are also important (Özer 2022). Therefore, in the production of medicinal and aromatic plants, cultivation and fertilization techniques should be determined and correct applications should be made. Although basil species have different chemotypes in terms of essential oil constituents, there are 4 different classifications in terms of main constituents as they are generally rich in linalool, methyl cinnamate, methyl chavicol and eugenol (Vernin et al. 1984; Simon et al. 1999). In studies carried out in our country, it was determined that 7 different chemotypes were found in basil species collected from Turkey (Telci 2005). The quality of natural products is determined by their composition and functional properties. It is therefore important to identify how the characteristics of plants change depending on the use of different fertilizers (Özbucak et al. 2014). In present study, Linalool, Bergamotene and Eugenol were the most abundant volatile compounds in basil leaves grown in the environments where different organic fertilizers were applied. However, it is observed that the amounts of these components vary depending on the fertilizer application and dose. Linalool was the component with the highest amount in all experimental groups compared to the control. The high proportion of linalool explains that the variety we studied is in the Linalool chemotype. Studies have supported that chemotypes rich in linalool are known as European chemotypes (Vernin et al. 1984; Simon et al. 1999; Raghavan 2006). In the study, it was determined that the Linalool ratio of basil varied between 27.28% and 48.45%. These results are like to some other studies (Milenković et al. 2019; Yaldız et al. 2019). The highest amount of Linalool was observed in 2.5% and 5% chicken manure applications in present study. Linalool results of present study are like Yaldız et al. (2019).

When the volatile components of 10% vermicompost leaf samples are examined, it is seen that the amounts of Linalool and Eugenol are higher than the control. It is noteworthy that the only experiment in which the amount of Eugenol was higher than the control was the 10% vermicompost application. The antimicrobial activity of essential oils generally depends on the chemical composition and quantity of each component. Scientific studies have shown that Eugenol shows excellent antimicrobial activity against a wide variety of Gram (+) and (-) bacteria and fungi (Marchese et al. 2017). Most of the researchers think that the hydroxyl groups of Eugenol bind to proteins and prevent enzyme activation (Burt 2004). However, there are also different mechanisms explaining the activity of Eugenol in the bacterial cell (Devi et al. 2010).

3.1. Antioxidant capacity results

3.1.1. Total phenolic content (TPC) (mg GAE/g sample)

TPC of the methanol and water extracts of leaf parts of the basil samples grown under different doses of chicken, barn, and worm applications was determined. According to the three-way-ANOVA results, interaction of fertilizer type*fertilizer dose*extraction solvent was found statistically significant ($P<0.01$).

The highest TPC was the samples from the pot where worm manure was applied at 30% (83.29 ± 9.66 mgGAE/g extract) and the pot where chicken manure was applied at 2.5% (83.11 ± 8.36 mgGAE/g extract). In the pots where barnyard manure was used, the relatively highest phenolic substance content was determined in the samples at the 20% dose level (43.28 ± 4.67 mgGAE/g extract). On the other hand, the lowest phenolic substance content was obtained for the water extract samples of the control group (0.57 ± 0.1) and the pot where 30% farmyard manure was applied (0.97 ± 0.4) (Table 2).

Table 2- Descriptive statistics and multiple comparison results for TPC values

Fertilizer											
Dose	***Extract***	***Barn***				***Worm***				***Chicken***			
		M ± SD				*M ± SD*				*M ± SD*			
Control	Methanol	67.36 ± 9.67	A	a	x	67.36 ± 9.67	A	a	x	67.36 ± 9.67	A	a	x
	Water	0.57 ± 0.1	A	a	y	0.57 ± 0.1	A	a	y	0.57 ± 0.1	A	a	y
Low	Methanol	4.47 ± 0.44	A	b	x	2.22 ± 0.78	B	b	x	4.15 ± 0.79	A	b	x
	Water	4.21 ± 0.53	A	a	x	1.78 ± 0.48	A	a	x	83.11 ± 8.36	B	b	y
Medium	Methanol	13.65 ± 0.71	A	b	x	2.08 ± 0.82	B	b	x	23.21 ± 3.69	C	c	x
	Water	43.28 ± 4.67	A	b	y	2.25 ± 0.49	B	a	x	4.47 ± 3.3	B	a	y
High	Methanol	4.26 ± 0.66	A	b	x	1.8 ± 0.1	B	b	x	-			
	Water	0.97 ± 0.4	A	a	y	83.29 ± 9.66	B	b	y	-			

There is a difference between fertilizer averages without a common capital letter in the same dose and extract (P<0.05)
There is a difference between dose averages that do not have a common lower case letter (a,b,c) at same extract and fertilizer (P<0.05)
There is a difference between extract averages that do not have a common lower case letter (x,y,z) at same dose and fertilizer (P<0.05)

TPC was found to be high and significant in the case of methanol extracts of control group pots (67.36 ± 9.67) ($P<0.05$). Therefore, any type or dose of fertilizer intervention in the pots led to a decrease in all methanol-extracted phenolic contents. in pots where barn manure was applied as 20%, worm manure as 30% and chicken manure as 2.5%, the phenolic substance content was higher in the case of water extract than in methanol extract.

TPC values calculated for methanol extracts of leaf samples from pots containing low levels of vermicompost were found to be statistically different ($P<0.05$) from the TPC values of the samples obtained with the other two types of fertilizer ($P<0.05$). Similar statistical difference was observed in the case of water extract of chicken manure treated samples. In the case of water extract, it was determined that only the phenolic content of barnyard manure differed.

The main objective of the production of medicinal and aromatic plants is to increase the yield of secondary metabolites. Organic fertilizers are known to improve secondary metabolism in essential oil-containing plants by increasing the amount of available nutrients (Emami Bistgani et al. 2018; Rostaei et al. 2018; Rao et al. 2022).

At the same time, organic fertilizer applications can play an important role in increasing plant resilience to stress conditions. In fact, based on this information, it may lead to a decrease in the production of secondary metabolites, as observed by Jakovljević et al. (2017) because it can be interpreted that the plant may not choose the secondary metabolite production route to cope with stress. Numerous studies have noted the effectiveness of organic fertilizers, compost, and compost teas in increasing vegetative growth, biomass, and essential oil yield of sweet marjoram, cumin, fennel, and basil. Improvements in yield and quality following the application of these organic-based materials may be due to increased beneficial microbial communities in the soil, improved nutrient absorption conditions for plants, and induction of plant defence compounds, growth regulators, or phytohormones (Javanmardi & Gjhorbani 2012).

3.1.2. Ferring reducing antioxidant power (FRAP)

It was concluded that the FRAP test results performed on extracts prepared from samples obtained by exposing them to different conditions to evaluate antioxidant activity were statistically significant (P<0.01). According to the three-way ANOVA test in terms of fertilizer type*fertilizer dose*extraction solvent, just like the TPC results (Table 3).

Table 3- Descriptive statistics and multiple comparison results for FRAP values

Fertilizer													
Dose	***Extract***	***Barn*** *M ± SD*				***Worm*** *M ± SD*				***Chicken*** *M ± SD*			
Control	Methanol	851.8 ± 10.16	A	a	x	851.8 ± 10.16	A	b	x	851.8 ± 110.16	A	a	x
	Water	4.32 ± 1.96	A	a	y	4.32 ± 1.96	A	a	y	4.32 ± 1.96	A	a	y
Low	Methanol	78.94 ± 3.5	A	b	x	31.95 ± 3.67	B	a	x	27.97 ± 5.38	B	b	x
	Water	29.17 ± 7.13	A	a	y	15.29 ± 5.43	A	a	y	625.07 ± 109.99	B	b	y
Medium	Methanol	181.49 ± 9.58	A	b	x	28.29 ± 2.64	B	a	x	213.75 ± 29.83	A	c	x
	Water	297.69 ± 6.87	A	b	y	15.99 ± 6.57	B	a	y	54.08 ± 20.32	B	a	y
High	Methanol	67.66 ± 2.96	A	b	x	19 ± 1.69	B	a	x	-			
	Water	8.44 ± 0.48	A	a	y	633.41 ± 4.12	B	b	y	-			

There is a difference between fertilizer averages without a common capital letter in the same dose and extract (P<0.05)
There is a difference between dose averages that do not have a common lower case letter (a,b,c) at same extract and fertilizer (P<0.05)
There is a difference between extract averages that do not have a common lower case letter (x,y,z) at same dose and fertilizer (P<0.05)

The highest and lowest FRAP values were observed in the methanol and water extracts of the control pots, respectively. Therefore, any dose of fertilizer intervention in the pots causes a decrease for methanol extracts and an increase for water extracts in terms of FRAP values FRAP values obtained for water extract were higher than those obtained for methanol extract in the samples treated with 20% barnyard manure, 30% vermicompost and 2.5% chicken manure, unlike the control samples. FRAP values of water extract samples were the highest in all pots except the control pot. However, the highest FRAP values were also determined in the methanol extracts of the samples treated with 20% barnyard manure and 5% chicken manure.

3.1.3. DPPH free radical scavenging activity

DPPH radical scavenging activity results as ascorbic acid equivalent were evaluated according to three-way ANOVA model (Table 4). Fertilizer type*fertilizer dose*extraction solvent interaction effect was found statistically significant (P<0.01). Statistically significant difference was observed in the DPPH radical scavenging activity values in the case of methanol and water extracts in all samples except the samples in which 20% stable manure was applied (P<0.05). The highest DPPH scavenging activity was obtained in the case of water extract of the samples obtained from the control pot. It can be said that applying any dose of fertilizer to the pots causes a decrease in the DPPH radical scavenging activity in the water extract samples.

Table 4- Descriptive statistics and multiple comparison results for DPPH values

Fertilizer													
Dose	***Extract***	***Barn*** *M ± SD*				***Worm*** *M ± SD*				***Chicken*** *M ± SD*			
Control	Methanol	0.17 ± 0.09	A	a	x	0.17 ± 0.09	A	a	x	0.17 ± 0.09	A	a	x
	Water	16.09 ± 4.85	A	a	y	16.09 ± 4.85	A	a	y	16.09 ± 4.85	A	a	y
Low	Methanol	0.62 ± 0.01	A	b	x	1.19 ± 0.08	B	b	x	1.22 ± 0.05	B	b	x
	Water	2.39 ± 0.48	A	bc	y	4.32 ± 0.29	B	b	y	0.1 ± 0.03	C	b	y
Medium	Methanol	0.23 ± 0.05	A	a	x	1.47 ± 0.06	B	b	x	0.19 ± 0.07	A	a	x
	Water	0.25 ± 0.11	A	b	x	3.93 ± 0.19	B	b	y	1.54 ± 0.45	C	b	y
High	Methanol	0.61 ± 0.03	A	b	x	2.09 ± 0.19	B	c	x	-			
	Water	8.11 ± 3.06	A	c	y	0.1 ± 0.02	B	b	y	-			

There is a difference between fertilizer averages without a common capital letter in the same dose and extract (P<0.05)
There is a difference between dose averages that do not have a common lower case letter (a,b,c) at same extract and fertilizer (P<0.05)
There is a difference between extract averages that do not have a common lower case letter (x,y,z) at same dose and fertilizer (P<0.05)

In barnyard manure applications, regardless of the solvent, the lowest DPPH activity was reached at 20% dose level. It was observed that DPPH activity reached the highest value in the water extract samples of the pot samples in which the same fertilizer was applied at 30%, and this value was statistically significantly different from all other units except the control pot. It was

observed that the DPPH radical scavenging activity in the water extract of the samples applied with 10% and 20% vermicompost was statistically significant and higher compared to other fertilizer types corresponding to the same dose levels (Table 4).

In accordance with the FRAP test results, a statistically significant difference was observed in terms of DPPH radical scavenging activity values in the case of methanol and water extracts in all samples except the samples in which 20% barnyard manure was applied ($P<0.05$).

Total phenolic content and antioxidant activity values decreased in the case of methanol extraction after all fertilizer applications, while an increase was observed in all 3 parameters in the case of water extraction. It has been previously reported that fertilizers can have positive and negative effects on phenolic compounds in plants (Gavrić et al. 2021). According to the results, the production of secondary metabolites in the leaves of basil plants was greatly affected by different types of fertilizers. In cases where low and moderate fertilizer applications were made, worm manure was less effective in terms of phenolic content than the other two types of fertilizers, while in cases of high dose fertilizer application, it was concluded that worm manure was more beneficial. The determined FRAP values of the samples prepared by extracting the basil samples grown with water and methanol after the application of control and all fertilizer types at different doses showed statistical differences. In this way, the importance of the effect of the extraction solvent on antioxidant activity was clearly observed. A similar study was carried out on *Ocimum* × *citriodorum* Vis. using worm and chicken manure teas in the ratios of 1:5 and 1:10, and the highest phenolic content was determined when worm manure tea was used in the ratio of 1:5. In this study, where the samples were extracted with acetone, the total antioxidant activity was found to be higher for the same group of extracts compared to others (Javanmardi & Ghorbani 2012).

The correlation coefficient between the total phenolic content and FRAP values of the extracts of the samples obtained after barn and worm manure applications was calculated as 0.91, while in the case of chicken manure application, this value was 0.89. The highest value of the correlation coefficient calculated between phenolic content and DPPH activity values is in the case of the sample exposed to worm manure, while it is 0.76 and 0.88 in the case of samples grown in the presence of barnyard and chicken manure, respectively. In another study where the effect of farm, worm, and chicken manure on the antioxidant activity of basil was compared with chemical fertilizer, the application of poultry manure achieved the highest result among organic fertilizers. Although chemical fertilizer did not cause any change in antioxidant activity compared to the control, the highest DPPH activity data was obtained when chemical fertilizer was used in combination with organic fertilizers (Pandey et al. 2016). Moreover, in the study, the statistical difference arising from the extraction solvent is generally noticeable in all cases. Studies have shown that extraction efficiency is strongly affected by the extraction solvent (Filip et al. 2017; Złotek et al. 2016) It was observed that the total phenolic content values calculated after the extraction of 1 g of basil plant material with 30 mL acetone, water, methanol, and ethanol for 40 minutes were highest in methanol (Do et al. 2020). In the current study, we detected higher phenolic content for methanol extracts, especially in the control group, which is compatible with the literature.

It was concluded that TPC and antioxidant capacity values of basil plants grown with organic, mineral and organomineral fertilizer applications were not significantly affected. This may be due to the weather conditions at the time of the trial. Because weather conditions can have a greater impact on the content of bioactive compounds than the effect of fertilizers (Gavrić et al. 2021). In another study examining the effects of four types of fertilizers, such as biosolids, microorganisms, organic fertilizers, and mineral fertilizers, on the bioactive compounds and antioxidant capacity of basil, it was found that, on the contrary, the bioactive compound content and antioxidant capacity value in basil increased in all applications. In the said study, the highest concentration of total polyphenols was recorded for biosolids fertilization, while total flavonoids and total anthocyanins were better increased by organic and microorganism treatments. However, the total amount of bioactive compounds recorded in dry basil leaves was detected at the highest level in organic and chemical fertilization and microorganism application (Teliban et al. 2020).

Jakovljević et al. (2017) reported that excessive application of chemical fertilizers stimulated the growth of basil plants, but also reduced the content of secondary compounds. To overcome these limitations, future researchers should more carefully consider the potential impact of different fertilizer types and doses, for example, on the composition of essential oils and phenols (Gavrić et al. 2021). In the essential oil analysis of the current study, the highest percentage of linalool was obtained when chicken manure was applied at a rate of 2.5%, and the fact that the highest phenolic content and highest frap values were calculated in the case of the same dose and type of fertilizer application is good evidence of the contribution of the essential oil composition to antioxidant activity.

3.1.4. Antimicrobial activity analysis results

Antimicrobial activities of the leaf samples for the paper disc diffusion method assay propose that the EOs essential oils and methanolic fractions had secondary antibacterial activities. Increasing doses of chicken, worm and barnyard manure applied in the study showed a progressively greater prohibitive influence on the antimicrobial fortune of *O. basilicum* leaf extracts (Table 5). The antimicrobial efficiencies of essential oils and methanolic extracts from *O. basilicum* treated with chicken, worm and barn manure were more powerful against bacteria than fungi and may be potential resources of new antimicrobial agents. The essential oils obtained from the *O. basilicum* treated with barn manures exhibited the highest antibacterial activity against Gram-

negative bacterial strains, such as *E. coli, C. freundii* and Gram-positive bacterial strains, such as *B. cereus*. Especially *O. basilicum* treated with barn manures showed the most effective antifungal effect against *S. cerevisiae* yeast. The EOs treated with worm demonstrated a more powerful zone demonstration against only Gram-positive bacterial strains, such as *B. cereus* and *S. aureus.* In this work, the excessive antibacterial activity was acquired with a small amount of vermicompost applied. At the same time, this application, which investigated the antimicrobial activity of *O. basilicum* against pathogenic bacteria, obtained the most effective antifungal activity. Right now, investigation we utilization Gram-negative bacteria that are contemplated more indestructible to essential oils than Gram-positive homologues (Nazzaro et al. 2013). This action is imputed to the complicated cell-wall structures of Gram-negative organisms, which do not allow a simple influence of antibiotics and drugs, with the inclusion of phenolic compounds, such as eugenol, Linalool, and other essential oils (Tiwari et al. 2009). The accepted mechanism of activity of essential oils depends on their capability to disturb the bacterial cell wall and the cytoplasmic membrane, so producing a cell lysis and loss of intracellular compounds (Lopez-Romero et al. 2015). The essential oils and the main components of these oils, such as carvacrol, citronellol, geraniol and neroli, exhibited antibacterial activities (Janssen et al. 1986). The essential oils acquired from the *O. basilicum* L. cured with chicken manure demonstrated the greatest antibacterial activity against *S. aureus, L. monocytogenes* and *E. coli* while that acquired by handling of chicken manures did not show the power antifungal efficiency against *C. albican* and *S. cerevisae.* Essentially, EOs can diffuse, the phospholipid bilayer of the bacterial cell wall, bind to proteins, and arrest their normal duties owing to their interplays with phenolic molecules (Sakkas et al. 2016). In contrast with beforehand conclusions (Morelli et al. 2017), who did not find any discoverable influence of soil HS practice neither on plant development, nor on *O. basilicum* antibacterial activity, our conclusions remark that. With chicken, worm, and barn fertilizer compost do impact the subsidiary metabolism of an aromatic plant, such as basil, and promote larger yields of phenolic EO components, such as Linalool, Bergamotene, Eugenol, Cadinene, Muurolol and Terpineol which are included in their antimicrobial activity. *O. basilicum* L. treated with middle chicken manure application of vermicompost, oil acquired from basil leaves was found to be most influential on microorganisms EO obtained from control leaves showed average effect on micro-organisms, while this value O. *basilicum* L. treated with little worm and middle barn manure application of showed power effect on microorganisms.

Among Gram (-) bacteria, the highest antibacterial effect on *E. coli* was determined in the essential oil extract of 20% barnyard manure leaf samples. The essential oil extract of 20% barnyard manure leaf samples also showed significant and high antifungal effect on *S. cerevisae*. When the volatile components of 20% samples of barn manure are examined, it is seen that the amount of Linalool is higher than the control. It has been determined by in vivo studies that linalool and linalool-rich essential oils have various biological activities such as antimicrobial, anti-inflammatory, anticancer and antioxidant. The repellent effect of linalool on insects damaging crops has also been investigated, emphasizing the importance of the application of this molecule in environmentally friendly insect pest management (Kamatou &Viljoen 2008). However, scientific studies have determined that linalool has many therapeutic effects (Sezen et al. 2021).

Table 5- Antimicrobial Activity Analysis Results

Fertilizer	Dose	Extract	Microorganism							
			S. aureus	*B. cereus*	*L. monocytogenes*	*E. coli*	*C. freundii*	*P. aerusinasa*	*C. albicans*	*S. cerevisae*
Barn	Low	M	**17.05 ± 0.01Abx**	6 ± 0Aa	**15.75 ± 0.01Abx**	15.25 ± 0.01Abx	12 ± 0.01Aax	12 ± 0.01abx	13 ± 0.01Abx	12.5 ± 0.01Abx
		EO	6 ± 0Aby	6 ± 0Aa	6 ± 0Aby	6 ± 0Aby	6 ± 0Aay	6 ± 0Aby	6 ± 0Aby	6 ± 0Aby
	Medium	M	12.34 ± 0.01Acx	-6 ± 0Aa	14 ± 0.01Acx	13.5 ± 0.01Acx	10 ± 0.01Abx	10.5 ± 0.01Ac	12.75 ± 0.01Acx	9.75 ± 0.01Acx
		EO	6 ± 0Aby	**17.5 ± 0.01Ab**	6 ± 0Aby	**20.5 ± 0.01Acy**	9 ± 0.01Aby	6 ± 0Ab	14 ± 0.01Acy	**21.25 ± 0.01Acy**
	High	M	13.25 ± 0.01Adx	6 ± 0Aa	11 ± 0.01Adx	11.25 ± 0.01Adx	9.75 ± 0.01Acx	8 ± 0.01d	13 ± 0.01Abx	9 ± 0.01Adx
		EO	6 ± 0Aby	13.75 ± 0.01Ac	6 ± 0Aby	12 ± 0.01Ady	**19.4 ± 0.01Acy**	6 ± 0Ab	9.25 ± 0.01Ady	12.75 ± 0.01Ady
Worm	Low	M	12.64 ± 0.01Bbx	6 ± 0Aa	10 ± 0.01Bbx	13 ± 0.01Bax	11.25 ± 0.01Bax	10.75 ± 0.01Bbx	11.5 ± 0.01Bbx	13.25 ± 0.01Bbx
		EO	6 ± 0Aby	**26 ± 0.01Ba**	6 ± 0Aby	13.5 ± 0.01Bby	13.75 ± 0.01Bay	6 ± 0Aby	10.13 ± 0Bby	**20 ± 0.01Bby**
	Medium	M	6.94 ± 0.01Bcx	-6 ± 0Aa	6 ± 0Aa	8.95 ± 0.78Bbx	11 ± 0.01Bbx	6 ± 0Aa	12.75 ± 0.01Acx	10 ± 0.01Bcx
		EO	6 ± 0Aby	20 ± 0.01Bb	6 ± 0Ab	11.75 ± 0.01Bcy	10 ± 0.01Bby	6 ± 0Ab	12.5 ± 0.01Bcy	14 ± 0.01Bcy
	High	M	**15.09 ± 0.01Bdx**	6 ± 0Aa	14 ± 0.01Bcx	9 ± 0.01Bbx	10.75 ± 0.01Bcx	6 ± 0Aa	14 ± 0.01Bdx	8 ± 0.01Bdx
		EO	8.15 ± 0.01Bcy	8.75 ± 0.01Bc	11.5 ± 0.01Bcy	6 ± 0Bdy	**21.75 ± 0.01Bcy**	6 ± 0Ab	7.25 ± 0.01Bdy	10.5 ± 0.01Bdy
Chicken	Low	M	9.24 ± 0.01Cbx	6 ± 0Aa	-	14.5 ± 0.01Abx	11.25 ± 0.01Bax	6 ± 0Aa	13.75 ± 0.01Cbx	12 ± 0.01Cbx
		EO	11.59 ± 0.01Bby	6 ± 0Aa	6 ± 0Ab	6 ± 0Aby	6 ± 0Aay	6 ± 0Ab	7 ± 0.01Cby	10 ± 0.01Cby
	Medium	M	**16.57 ± 0.01Ccx**	6 ± 0Aa	**15 ± 0.01Bb**	**15 ± 0.01Ccx**	11.75 ± 0.01Cbx	11.75 ± 0.01Bb	12 ± 0.01Bax	9 ± 0.01Ccx
		EO	6 ± 0Acy	7 ± 0.01Cb	6 ± 0Ab	8.75 ± 0.01Ccy	9.75 ± 0.01Cby	6 ± 0Ab	9 ± 0.01Ccy	14 ± 0.01Ccy
	Gentamicin		41,9± **0.01BAc**	40,99± **0.0BAc**	49,99± **0.01 BAc**	41,9± **0.01 BAc**	37,99± **0.01 BAc**	34,76± **0.01 BAc**	NT	NT
	Nystatin		NT	NT	NT	NT	NT	NT	17.67± **0.01**	17.45± **0.01**
ANOVA (3-way) results, only for interaction term		**F. Value**	F(5,44)=507691.16	F(3,16)=2208055.56	F(3,40)=75775.79	F(5,44)=1141.01	F(3,32)=131138.1	F(2,38)=13494.32	F(5,44)=84648.37	F(5,44)=367485.12
		P. Value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

M: Methanol, EO: Essential oil, NT: Not tested, *Listeria monocytog*enes ATCC®7677, *Staphylococcus aureus* ATCC 6538, *Bacillus cereus* ATCC®10876) *Pseudomonas aeruginosa* ATCC®27853, *Citrobacter freundii* ATCC® 43864 (-), *Escherichia coli,* yeast *Saccharomyces cerevisiae* ATCC 976, and fungi, *Candida albicans* ATCC®10231). The significant values are indicated in red.

4. Conclusions

Few investigations have been reported on behavior regarding the organic chicken, worm, and barn fertilization effect on essential oil content and the quality of basil. Particularly, medium chicken, low worm, and medium barn manure doses had the largest antibacterial efficiency, as well as antioxidant activity, total phenolics, flavonoid and essential oil components. These fertilizers could thence be contemplated as appropriate substitutes for chemical fertilizers when expanding medicinal plants that are gaining both an incremented significance and requisition. We recommend that the effective chemical compounds existing in *O. basilicum* L could fluctuation a role in the treatment of various bacterial contaminations. Hence, wider, and detailed research may be requisitioned to discover the potential use of *O. basilicum* L. for treatment of contagious diseases.

Acknowledgments

We would like to thank Mr. Mehmet Muharrem Özcan for his help in obtaining essential oils from plant samples.

References

Abou-Sreea A I, Rady M M, Roby M H, Ahmed S M, Majrashi A & Ali E F (2021). Cattle manure and bio-nourishing royal jelly as alternatives to chemical fertilizers: Potential for sustainable production of organic *Hibiscus sabdariffa* L. *Journal of Applied Research on Medicinal and Aromatic Plants* 25: 100334

AcıbucaV & Budak D B (2018). The place and importance of medicinal and aromatic plants in the world and Turkey. *Çukurova Journal of Agricultural and Food Sciences* 33(1): 37-44

Arslan N, Baydar H, Kızıl S, Karık Ü, Şekeroğlu N & Gümüşçü A (2015). Changes and New Searches in Medicinal and Aromatic Plants Production. *VII. Turkey Agricultural Engineering Technical Congress* pp. 483-507

Bączek K, Kosakowska O, Gniewosz M, Gientka I & Weglarz Z (2019). Sweet basil (*Ocimum basilicum* L.) productivity and raw material quality from organic cultivation. *Agronomy* 9(6): 279

Benzie I F & Strain J J (1999). The ferric reducing ability of plasma (FRAP) as a measure of "antioxidant power": the FRAP assay. *Analytical Biochemistry* 15: 239(1):70-6

Bernal M P, Alburquerque J A & Moral R (2009). Composting of animal manures and chemical criteria for compost maturity assessment. A review. *Bioresource technology* 100(22): 5444-54532

Blois M S (1958). Antioxidant Determinations Using a Stable Free Radical. *Nature* 181: 1199-1200

Burt S (2004). Essential oils: their antibacterial properties and potential applications in foods—a review. *International journal of food microbiology* 94(3): 223-253

Dehghani-Samani J A, Ghasemi Pirbalouti F, Malekpoor & Rajabzadeh F (2021). Chemical composition and yield of essential oil from two sweet basil species (*Ocimum ciliatum* L. and *O. basilicum* L.) under different fertilizers. *Journal of Medicinal Herbs* 12(1): 27-33

Devi K P, Nisha S A, Sakthivel R & Pandian S K (2010). Eugenol (an essential oil of clove) acts as an antibacterial agent against *Salmonella typhi* by disrupting the cellular membrane. *Journal of ethnopharmacology* 130(1): 107-115

Do T H, Truong H B & Nguyen H C (2020). Optimization of extraction of phenolic compounds from *Ocimum basilicum* leaves and evaluation of their antioxidant activity. *Pharmaceutical Chemistry Journal* 54(2): 162-169

Egata D F (2021). Benefit and use of sweet basil (*Ocimum basilicum* L.) in Ethiopia: A review. *J. Nutr. Food Proces* 4(5): 57-59

Emami Bistgani Z, Siadat S A, Bakhshandeh A, Pirbalouti A G, Hashemi M, Maggi F & Morshedloo M R (2018). Application of combined fertilizers improves biomass, essential oil yield, aroma profile, and antioxidant properties of *Thymus daenensis* Celak. *Ind Crop Prod* 121:434–440

Feledyn-Szewczyk B, Stalenga J & Harasim E (2020). Soil Science and Plant Cultivation in Organic Farming. MDPI2.

Filip S, Pavlić B, Vidović S, Vladić J & Zeković Z (2017). Optimization of microwave- assisted extraction of polyphenolic compounds from *Ocimum basilicum* by response surface methodology. *Food Analytical Methods* 10(7): 2270-2280

Gavrić T, Jurković J, Gadžo D. Čengić L, Sijahović E & Bašić F (2021). Fertilizer effect on some basil bioactive compounds and yield. *Ciência e Agrotecnologia*, 45, e003121.

Janssen A M, Chin N L J, Scheffer J J C & Svendsen AB (1986). Screening for antimicrobial activity of some essential oils by the agar overlays technique. Pharmaceutisch Weekblad 8(6): 289-292

Kamatou G P & Viljoen A M (2008). Linalool–A review of a biologically active compound of commercial importance. Natural Product Communications 3(7): 1183-1192

Karaca M, Kara Ş M & Özcan M M (2017). Determination of herb yield and essential oil content of some basil (*Ocimum basilicum* L.) populations. *Ordu University Journal of Science and Technology* 7(2): 160-169

Jakovljević D, Stanković M, Bojović & Topuzović M (2017). Regulation of early growth and antioxidant defence mechanism of sweet basil seedlings in response to nutrition. *Acta Physiol. Plant*, 2017, 39: 243

Javanmardi J & Ghorbani E (2012). Effects of chicken manure and vermicompost teas on herb yield, secondary metabolites, and antioxidant activity of lemon basil (*Ocimum x citriodorum* Vis.). *Advences in Horticultural Science* 26(3-4): 151-157

Lopez-Romero J C, González-Ríos H, Borges A & Simões M (2015). Antibacterial effects and mode of action of selected essential oils components against *Escherichia coli* and *Staphylococcus aureus*. *Evid Based Complementary Altern Med*. https://doi.org/10.1155/2015/795435.

Loss A, Couto R D R, Atanassov A, Vlahova, M, Carlier L, Christov N & Esatbeyoglu T (2020). Animal manure as fertilizer: changes in soil attributes, productivity, and food composition. *Scientific reports* 10(1): 1-161

Marchese A, Barbieri R, Coppo E, Orhan I E, Daglia M, Nabavi S, Izadi M, Abdollahi M, Nabavi S & Ajami M (2017). Antimicrobial activity of eugenol and essential oils containing eugenol: A mechanistic viewpoint. *Critical reviews in microbiology* 43(6): 668-689

Milenković, Stanojević J, Cvetković D, Stanojević L, Lalević D & Šunić L (2019). New technology in basil production with high essential oil yield and quality. *Industrial Crops and Products* 140: 111718

Mohamed A A & Alotaibi B M (2023). Essential oils of some medicinal plants and their biological activities: a mini review. *Journal of Umm Al-Qura University for Applied Sciences* 9(1): 40-49

Morelli F, Ferarrese L, Munhoz C L & Alberton O (2017). Antimicrobial activity of essential oil and growth of *Ocimum basilicum* (L.) inoculated with mycorrhiza and humic substances applied to soil. *Genet Mol Res*. 16(3):16039710

Murillo-Amador B, Nieto-Garibay A, Troyo-Diéguez E, Flores-Hernández A, Cordoba-Matson MV & Villegas-Espinoza A (2013). Proximate analysis among 24 Ocimum cultivars under two cultivation environments: A comparative study. *J. Food Agric. Environ* 11: 2842-2848.

Nadeem H R, Akhtar S, Sestili, P, Ismail T, Neugart S, Qamar M & Esatbeyoglu T (2022). Toxicity, antioxidant activity, and phytochemicals of basil (*Ocimum basilicum* L.) leaves cultivated in Southern Punjab, Pakistan. Foods, 11(9), 1239.

Nazzaro F, Fratianni F, De Martino L, Coppola R & De Feo V (2013). Effect of essential oils on pathogenic bacteria. *Pharmaceuticals* 6(12): 1451-1474

Özbucak T B, Erturk Ö, Yıldız O, Bayrak A, Kara M, Şahin H & Kiralan M (2014). The effect of different manures and synthetic fertilizer on biochemical and antimicrobial properties of *Mentha piperita* L. *Journal of Food Biochemistry* 38(4): 424-432.

Özbucak T & Alan H (2024). Influence of Solid and Liquid Red California Vermicompost (*Eisenia foetida*) on Growth and Yield of Lettuce (*Lactuca sativa* var. *crispa* L.). *Journal of Agricultural Sciences* 30(2): 336-344

Özer P C (2022). The effects of different organic and inorganic fertilizers on agronomic characteristics, essential oil content and composition of basil (*Ocimum basilicum* L.) in Bursa ecological conditions. Master's Thesis, *Bursa Uludag University, Institute of Science and Technology*, Department of Field Crops, Bursa

Pandey V, Patel A & Patra D D (2016). Integrated nutrient regimes ameliorate crop productivity, nutritive value, antioxidant activity and volatiles in basil (*Ocimum basilicum* L.). *Industrial Crops and Products* 87: 124-131

Raghavan S (2006). Handbook of spices, seasonings, and flavorings. *CRC press*.

Raina A P & Gupta V (2018). Chemotypic characterization of diversity in essential oil composition of *Ocimum* species and varieties from India. *Journal of Essential Oil Research* 30(6): 444-456

Rao K S, Haran R H& Rajpoot V S (2022). Value addition: A novel strategy for quality enhancement of medicinal and aromatic plants. *Journal of Applied Research on Medicinal and Aromatic Plants* 31: 100415

Ronald M A & Microbiologia (1990). Compania Editorial Continental S.A. de C.V., Mexico: 505

Rostaei M, Fallah S, Lorigooini Z & Surki A A (2018). The effect of organic manure and chemical fertilizer on essential oil, chemical compositions, and antioxidant activity of dill (*Anethum graveolens*) in sole and intercropped with soybean (Glycine max). *Journal of Cleaner Production* 199: 18-26

Sakkas H, Gousia P, Economou V, Sakkas V, Petsios S & Papadopoulou C (2016). In vitro antimicrobial activity of five essential oils on multidrug resistant gram-negative clinical isolates. *Journal of Intercultural Ethnopharmacology* 5: 212-218

Santos B C S, Pires A S, Yamamoto C H, Couri M R C, Taranto A G, Alves M S, De Matos Araújo A L & De Sousa O V (2018). Methyl chavicol and its synthetic analogue as possible antioxidant and anti-lipase agents based on the in vitro and in silico assays. *Oxid. Med. Cell. Longev*.

Sezen S, Özer Ö S & Çınar F (2021). Chemical constituents, antioxidant and antibacterial activities of Citrus (*Citrus aurantium* L.) leaf, peel and juice essential oils. *Mediterranean Fisheries and Aquaculture Research* 4(3): 58-73

Shahrajabian M H, Sun W & Cheng Q (2020). Chemical components and pharmacological benefits of Basil (*Ocimum basilicum*): A review. *International Journal of Food Properties* 23(1): 1961-1970

Sharma S, Kumari A, Dhatwalia J, Guleria I, Lal S, Upadhyay N & Kumar A (2021). Effect of solvents extraction on phytochemical profile and biological activities of two *Ocimum* species: A comparative study. *Journal of Applied Research on Medicinal and Aromatic Plants* 25: 100348

Simon J E, Quinn J & Murray R G (1999). Basil: a source of essential oils. In: Janick, J., Simon, JE. (Eds.), *Advanced in New Crops*. Timber Press, Portland, OR, pp. 484-489

Singleton V L & Rossi J A (1965). Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents. *American journal of Enology and Viticulture* 16(3): 144-158

Telci İ (2005). Determination of suitable mowing heights in Reyhan (*Ocimum basilicum* L.) genotypes. *Gaziosmanpasa University Journal of Faculty of Agriculture* 22(2): 77-83

Teliban G C, Stoleru V, Burducea M, Lobiuc A, Munteanu N, Popa L D & Caruso G (2020). Biochemical, physiological and yield characteristics of red basil as affected by cultivar and fertilization. *Agriculture* 10(2): 48

Tenore G C, Campiglia P, Ciampaglia R, Izzo L & Novellino E (2017). Antioxidant and antimicrobial properties of traditional green and purple "Napoletano" basil cultivars (*Ocimum basilicum* L.) from Campania region (Italy). *Natural Product Research* 31(17): 2067-2071

Tiwari B K, Valdramidis V P, Donnel C P O, Muthukumarappan K, Bourk P & Cullen P J (2009). Application of natural antimicrobials for food preservation. *Journal of Agricultural Food Chemistry* 2009; 57: 5987–6000

Tsvetkov I, Atanassov A, Vlahova M, Carlier L, Christov N, Lefort F & Atanassov I (2018). Plant organic farming research–status and opportunities for future development. *Biotechnology & Biotechnological Equipment* 32(2): 241-260

Vernin G, Metzger J, Fraisse D, Suon K N& Scharff C (1984). Analysis of basic oils by GC-MS data bank. *Perfumer & flavorist*, 9(5): 71-86

Yaldiz G, Camlica M & Ozen F (2019 a). Biological value and chemical components of essential oils of sweet basil (*Ocimum basilicum* L.) grown with organic fertilization sources. *Journal of the Science of Food and Agriculture* 99(4): 2005-2013

Yaldız G, Çamlıca M, Özen F & Eratalar S A (2019 b). Effect of poultry manure on yield and nutrient composition of sweet basil (*Ocimum basilicum* L.). *Communications in soil science and plant analysis* 50(7): 838-852

Yolcu H (2011). The effects of some organic and chemical fertilizer applications on yield, morphology, quality, and mineral content of common vetch (*Vicia sativa* L.). *Turkish Journal of Field Crops* 16(2): 197-202

Złotek U, Szymanowska U, Karaś M & Świeca M (2016). Antioxidative and anti-inflammatory potential of phenolics from purple basil (*Ocimum basilicum* L.) leaves induced by jasmonic, arachidonic and β-aminobutyric acid elicitation. *International Journal of Food Science and Technology* 51(1): 163–170

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Possibilities of Using Regional Index-Flood Method with Annual Maximum and Partial Duration Series: A Case Study of Susurluk River Basin, Turkey

Ayşe Doğanülker[a], Alper Serdar Anli[a*], Havva Eylem Polat[a]

[a]*Ankara University Agricultural Faculty Agricultural Structures and Irrigation Department, 06110, Diskapi, Ankara, TÜRKIYE*

ARTICLE INFO
Research Article
Corresponding Author: Alper Serdar Anli, E-mail: asanli@agri.ankara.edu.tr
Received: 01 July 2024 / Revised: 26 August 2024 / Accepted: 29 August 2024 / Online: 14 January 2025

Cite this article
Doğanülker A, Anli A S, Polat H E (2025). Flood Method and Annual Maximum and Partial Duration Series: A Case Study of Susurluk River Basin, Turkey. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):161-181. DOI: 10.15832/ankutbd.1508286

ABSTRACT
Among the natural disasters experienced in Turkey, floods, which cause the most loss of life and property after the earthquake, have increased their impact and frequency of occurrence over time, as well as unplanned urbanization caused by the increasing population, uncontrolled construction in stream beds, and changing climate. Therefore, it is important to accurately predict the magnitude and frequency of floods. This study investigated the possibilities of using the regional index-flood method and annual maximum series (AMS) and partial duration series (PDS) in the Susurluk River basin. Annual maximum flood series provided homogeneity in the Susurluk basin as a single region, and the Generalized Logistic (GLO) distribution fits the AMS. PDS was extracted according to the threshold levels determined using the variance-mean ratio and frequency factors. The PDS's most appropriate frequency factors (k) were determined according to the Poisson distribution, which makes the variance-mean ratio equal. k=3.5 was determined for only two stations. k=4 was suitable for seven stations, and k=5 was suitable for thirteen stations. The average number of peaks over the threshold level (λ`) varies between 1.26 and 5.31. Since PDS is not homogeneous in a single region, cluster analysis divided the basin into three regions. After homogeneity was achieved, Pearson Type 3 (PE3) and Generalized Pareto (GPA) distributions were suitable with the PDS. The study concluded that instead of annual maximum flood series, partial duration flood series can be used in many stations in the short and medium term but can be used in fewer stations in the long-term estimations. Since Regions I and II are relatively lower and flatter areas than Region III, it was observed that the flows started to accumulate at the stations in these regions, and larger floods were predicted. Region III is close to the basin upstream, and smaller floods were predicted at the stations in this region. Since partial duration flood estimations are lower than annual maximum flood ones, they can provide advantages to engineering projects with lower costs. In addition, PDS can be useful in regionalizing floods, which are very common due to the data extraction process.

Keywords: Peaks-over-threshold, frequency factor, L-moments, regional frequency analysis, design flood, Susurluk basin

1. Introduction

Hydrology studies the cycle, distribution, properties, and environmental effects of water on the atmosphere, surface, and underground. As the need for water increases over time, it has become an increasingly important branch of science. Therefore, water distribution, protection, and control have become vital. Flood is a situation where the amount of water in stream beds suddenly increases due to the excess rainfall in the basin, overflows from the stream bed, and damages the living beings, soil, and property around it. The acceleration of the hydrological cycle due to global climate change and designs according to incorrect flow rates have increased the frequency of floods today (Doğanülker 2022).

Many precautions have been taken against floods: Preventing settlements in areas with stream beds, reducing damage to the design of water structures, and controlling floods. Strong forecasts of floods are important to be able to take these precautions. It is impossible to have precise information about when and how much a future flood will occur. For this reason, floods can be estimated using appropriate statistical methods using data obtained from flow observation stations established at certain points on the stream (Anlı et al. 2007). One of these methods is frequency analysis (how much and at what intervals a flood will occur in the future), and the data obtained from the selected region for analysis must cover a period long enough to make the event sufficient. Flood frequency is an analysis performed by finding the most compatible probability distribution with the peak flow values occurring in a region. The expected result of this frequency analysis is that the estimated flood magnitude is close to the desired recurrence period (Fill & Steiner 2003). Two types of series are frequently used in flood frequency analysis: Annual maximum & partial-duration flood series. The series created from the maximum flows occurring once in a water year is called the annual maximum flood series. The series extracted with data higher than a specified threshold value instead of the maximum value taken yearly is called a partial-duration series (Seçkin 2009). Flood frequency analysis can be applied with two different techniques: at-site and regional. Since at-site flood frequency analysis is obtained using hydrological data at a single flow station,

reliable results cannot be obtained when similar situations are calculated from different stations. In basins where no measurements or insufficient measurements have been made; the regional flood frequency analysis method is used to estimate flood flows. Previous studies on at-site and regional frequency analysis of annual maximum and partial flood series have been attempted to be summarized chronologically.

Cunnane (1979) investigated the fit of the Poisson distribution to partial duration series using data from 20 stations in 20 basins in Great Britain. Ben-Zvi (1991) tested the fit of negative Binomial and Poisson distributions to partial duration flood series using data from 8 hydrometric stations in temporary rivers in Israel. A range of threshold levels were considered at each station. Positive results were obtained at the 5% significance level in 30 out of 53 tests for the Poisson distribution and 22 out of 28 for the negative Binomial distribution. Rasmussen & Rosbjerg (1991) stated that dividing the year into a certain number of seasons is important to correctly represent the data above the threshold value determined in the partial duration series. However, they stated that the seasonal models extracted are unsuitable for predictions because they require many parameters. Their study estimated seasonal and non-seasonal return periods by applying the Exponential and Poisson distributions to those exceeding the threshold value. As a result, they stated that the most suitable estimates were mostly obtained in the non-seasonal model. Rosbjerg et al. (1992) stated in their studies that the assumption that the data above the threshold value determined in the partial duration series is suitable for the Exponential distribution is suitable for the Generalized Pareto distribution. Wilks (1993) performed frequency analysis using annual maximum and partial duration data taken from stations in the USA. He analysed the performance of 8 three-parameter probability distributions and stated that the beta-k distribution most accurately represented the rightmost tail of the annual maximum data. He also observed that the Beta-P distribution best represented the partial duration series. The two-parameter Gumbel distribution gave below-prediction probabilities in annual maximum and partial duration data at high rainfall amounts. Madsen et al. (1997a) used the Generalized Pareto distribution for the partial duration series and the Generalized Extreme Values distribution for the annual maximum series in the study. Madsen et al. (1997b) stated that in the regional analysis where partial duration and annual maximum series were compared, partial duration series gave the best results with Generalized Pareto distribution and annual maximum series with Generalized Extreme Values distribution. The accuracy of the partial duration series Generalized Pareto, and annual maximum series Generalized Extreme Values regional index-flood models were compared using Monte Carlo simulation, and it was revealed that the partial duration series data of the defined regions were more homogeneous than the annual maximum series. Lang et al. (1999) emphasized that annual maximum flood continues to be the most well-known approach in frequency analysis. They stated that the difficulties associated with using an alternative, the peaks over threshold approach, are the selection of thresholds and flood peak protection criteria. They argued that the literature on the peaks over threshold model is limited and inconsistent. Anlı (2009), in his regional frequency analysis study focusing on precipitation in Ankara using the L-moment method, used annual maximum and partial duration series and stated that generally more accurate results were obtained with high probabilities in the partial duration data set and that the partial duration data set could be used as an alternative to the annual maximum data set in at-site regional precipitation estimates. Rahman et al. (2013), in a regional flood frequency analysis study for semiarid and arid regions of Australia, found that the Generalized Pareto distribution was preferable to the Exponential distribution for the application of partial duration flood data in arid regions and determined that arid regions exhibited a much steeper flood frequency growth curve than humid regions. Bezak et al. (2014) stated that the peaks over the threshold data set provided more accurate results than the annual maximum data set for data obtained from a measurement station in Slovenia. They stated that the Binomial distribution did not provide a noticeable improvement over the Poisson distribution in modelling the number of annual exceedances of the threshold. Pham et al. (2014) examined the performance of the partial duration series and aimed to determine the regional λ (The average number of peaks over the threshold) value. According to the results obtained, they observed that the Generalized Pareto distribution best explains the partial duration series in the region and stated that they achieved the best partial duration series/Generalized Pareto performance when the λ value was equal to 5. Guru (2016) observed that the Generalized Pareto distribution in India best describes the partial duration series in the region and also observed that the λ' value in the most accurate partial duration series/Generalized Pareto relationship was found at 2, 2.5, and 3. Karim et al. (2017) showed in their study that frequency estimates based on partial series were better than those based on annual series for small and medium-sized floods, and both methods gave similar results for large floods. In their study, Zadeh et al. (2019) stated that the technique of regionalizing peaks over threshold series generally improved flood estimation compared to regionalizing annual maximum series. Ahmad et al. (2019) observed that Generalized Extreme Values with L-moment parameter estimation methods in partial duration and annual maximum series were the most suitable statistical distributions for partial duration series. Generalized Logistic was the most suitable statistical distribution for the annual maximum series. They also concluded in the study that the partial duration series performed better than the annual maximum series in estimating amounts for the most suitable probability distributions. Swetapadma & Ojha (2021) applied maximum entropy theory in partial duration series modelling of flood frequency analysis to find the appropriate threshold level and related distribution patterns in New Zealand. They concluded that it can be used as a good method for threshold determination in partial duration series of flood frequency studies.

Other important similar studies on partial duration series in recent years are as follows: Askhar & Ba (2017) applied the four-parameter Kappa distribution to partial duration flood series, Durocher et al. (2019) studied non-stationary frequency analysis on partial duration flood series with the help of the regional index-flood method, Agilan et al. (2020) used the Generalized Pareto distribution as non-stationary frequency analysis on partial duration series, Van Campenhout et al. (2020) emphasized that hourly data should be investigated in partial duration series, not daily, Kiran & Srinavas (2021) applied regional frequency analysis to both annual maximum and partial duration flood series and obtained various regression models, Pan & Rahman (2021)

investigated the differences in annual maximum and partial duration flood series estimates with the help of basin physiographic features, Guru (2022) mentioned the difficulty in finding the average number of peaks over threshold (λ') and tried to estimate with various methods, Pan et al. (2022) investigated the superiority of partial duration flood series over annual maximum series, Yue et al. (2022) examined partial duration rainfall series with non-stationary frequency analysis, Amorim & Villarini (2024) performed trend analysis with parametric and non-parametric methods in partial duration series with Generalized Pareto distribution.

There are few studies in Turkey where partial duration flood series and regional analysis are applied. There is also no study in the Susurluk basin, where floods cause damage. This study aims to investigate the possibilities of using partial duration flood series on a regional and at-site basis as an alternative to annual instantaneous maximum flood data in the Susurluk River basin using the regional index-flood approach and to assist flood risk management.

2. Material and Methods

2.1. Study area and data

The Susurluk basin, a key area of more than 2 million hectares in the west of Turkey, is a significant player in water resources management. Situated between 39°-40° north latitudes and 27°-30° east longitudes, the basin's total rainfall area is 22 399 km^2, with an average annual flow of 5.43 km^3. In the south of the Marmara Region, the basin is home to numerous large and small rivers that flow continuously or for short periods, making it a crucial resource for the region's water needs (Gürler et al. 2024). Susurluk Basin covers some of the rapidly growing provinces such as Bursa, Kütahya, Balıkesir, Çanakkale, Bilecik, Manisa, and İzmir. Floods are even more dangerous for those living in these cities, where the population increases daily. Especially in the Osmangazi district of Bursa province, located in the north-east of the basin, flood events have occurred from time to time, and various precautions and evacuation plans have been made in the reports prepared by the General Directorate of Water Management in this region for environmental impact assessment, human health and protection against floods. Other areas of the basin are affected by floods at medium and low risk (SYGM 2022). Since cities are located in areas with high flood risk, the risk of damage resulting from floods and possible major accidents is also high. This research used annual maximum and partial duration flood series among daily flow rates for flood frequency analysis and for determining flood risk. Susurluk Basin was chosen for regional flood frequency analysis due to its location. Daily streamflow discharges obtained from 22 streamflow observation stations were used in the study. Some characteristics of the streamflow observation stations operated by the General Directorate of State Hydraulic Works (DSI) and the previously closed Electrical Works Survey and Development Administration (EIE) in the basin are given in Table 1. The observation periods of the stations are from the date of commencement of operation to the last evaluation year. Streamflow rates were naturalized by the institutions from which they were obtained (DSI and EIE). The study consists of a master's thesis and started in 2021. Therefore, the streamflow data from the stations used are until 2016 and 2017. The observation period of the stations varies between 25-68 years. Streamflow data was obtained continuously for the specified years. The data set of the stations operated by DSI is shorter, while the data set of the stations operated by EIE is longer. While DSI stations are generally operated to determine the capacity of water storage structures, EIE stations are operated for both this purpose and research purposes. In addition, the streamflow data of both institutions are uninterrupted in the Susurluk basin. Many small and large streams in the basin flow continuously or for short periods. While the flows of these streams contain very little water, especially in the summer and early autumn months due to agricultural water use, they reach quite high levels in the winter and spring with the amount of melting snow. The locations of the streamflow observation stations in the Susurluk basin, where the daily flow rates used as material in the research were obtained, are shared in Figure 1.

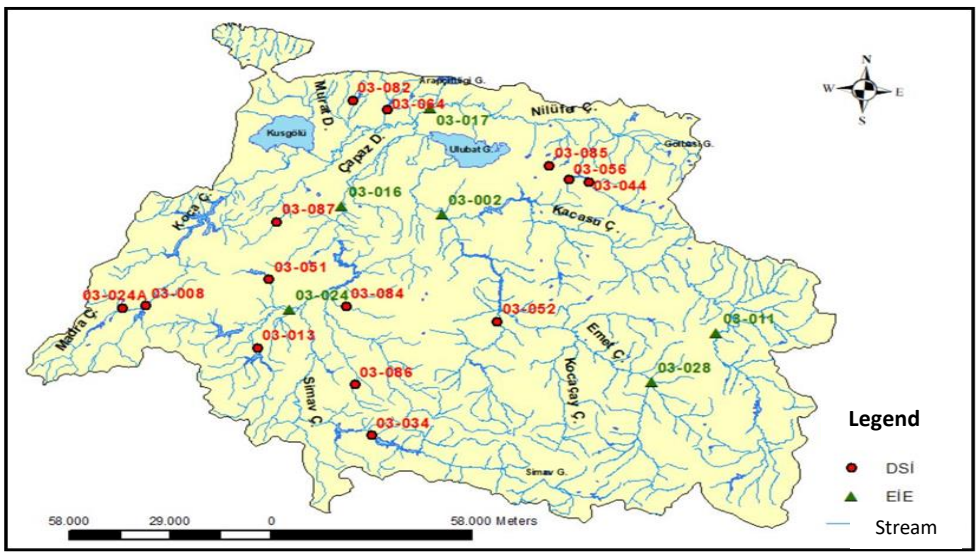


Figure 1- The location of the streamflow observation stations in the Susurluk Basin

Table 1- Characteristics of the streamflow observation stations in the Susurluk Basin

No	*Station code*	*Station name*	*Longitude–Latitude (°)*				*Precipitation area (km²)*	*Elevation (m)*	*Observation period*	*Sample size*
1	D03A008	Kahve	27.54	East	39.61	North	741	190	1969–2016	48
2	D03A013	İkizcetepeler	27.92	East	39.50	North	467	128	1972–2017	46
3	D03A024	Ayaklı	27.36	East	39.52	North	115	250	1967–2016	50
4	D03A034	Osmanlar Köp.	28.32	East	39.25	North	1266	277	1979–2017	39
5	D03A044	S.Saygı Brj. Gir.	29.00	East	40.08	North	377	341	1987–2017	31
6	D03A051	Değirmenboğazı	27.95	East	39.71	North	84	192	1981–2017	37
7	D03A052	Sinderler	28.72	East	39.62	North	965	294	1983–2017	35
8	D03A056	Sultaniye	28.94	East	40.09	North	50	368	1989–2017	29
9	D03A064	Gölecik	28.28	East	39.61	North	111	27	1993–2017	25
10	D03A082	Keçiler	28.18	East	40.30	North	21	65	1990–2017	28
11	D03A084	Eyüpbükü	28.23	East	39.65	North	241	945	1986–2017	32
12	D03A085	İnegazi	28.87	East	40.13	North	15	306	1989–2017	29
13	D03A086	Adalı	28.26	East	39.39	North	66	375	1987–2017	31
14	D03A087	Yeşilova	27.96	East	39.90	North	141	250	1991–2017	27
15	D03A096	Okçular	28.30	East	39.40	North	35	405	1991–2017	27
16	E03A002	Döllük	28.51	East	39.62	North	9617	40	1950–2017	68
17	E03A011	Küçükilet	29.86	East	39.12	North	1642	795	1955–2017	63
18	E03A016	Yahyaköy	28.17	East	39.98	North	6376	32	1954–2017	64
19	E03A017	Akçasusurluk	28.40	East	40.26	North	20	2	1957–2017	61
20	E03A024	Balıklı	28.02	East	39.63	North	244	94	1958–2017	60
21	E03A028	Dereli	29.25	East	39.46	North	1165	557	1969–2017	49
22	E03A031	Dağgüney	29.06	East	39.92	North	3493	365	1992–2017	26

2.2. Methods

2.2.1. Selection of annual maximum flood series

The annual maximum flood series used in the research are meticulously selected, considering the largest values from the data obtained from the time series of a hydrological event within a year. This careful selection process ensures the reliability of our findings.

2.2.2. Extraction of partial duration flood series

A partial duration flood series to be used in the research is a series consisting of independent events (m) that all exceed the selected threshold value (x_0) at a station and are subtracted from the observation period (n). Adamowski et al. (1998) stated that non-annual exceedance series were selected. In extracting the partial duration series, it is essential to choose the threshold value, specify the peak flow rates that exceed the predicted threshold value, and determine parameter estimates and probability distributions for modelling the size of these peaks. A series larger than the estimated threshold value and more ($m > n$) values than the observation year is selected. The selection of the threshold value when creating a non-annual exceedance series is obtained by the variance-mean ratio (σ^2/μ) method. The purpose of this method is explained by the fact that the mean and variance of the Poisson distribution are equal. Therefore, if the distribution of peak numbers fits with the Poisson distribution, the ratio of the variance (σ^2) to the mean (μ) of the peak numbers occurring each year (exceeding the threshold value) during the observation period is expected to be equal to or close to one (Adamowski 2000). In recent years, partial duration series models derived from this method have been used by some researchers (Pan et al. 2022; Yue et al. 2022; Amorim & Villarini 2023).

When selecting the threshold value, it will be assumed that the number of values exceeding the threshold value determined as a trial from the beginning for each year fits with the Poisson distribution, which is a discrete distribution given in Equation 1 (Cunnane 1979);

$$P(z\ pik > x_0) = P_z = e^{-\lambda}\ \lambda^z / z! \quad (1)$$

Where: P_z= probability of the number of peaks over a threshold in each year, z= the number of peaks over a threshold value in each year, x_0 = threshold value, λ = mean of the distribution.

When determining non-annual exceedance series, first, various values at a station are tried until a condition is obtained where the variance-mean ratios (σ^2/μ) are one or very close to one throughout the observation period. Then, flood discharges are selected among the observations according to this condition. As a result, it is reached that the average number (λ') of the peaks exceeding the experimentally selected threshold value at the stations throughout the observation period is greater than one. A flood data lower than the largest flood data chosen each year in the annual maximum series may also cause floods, and the results obtained in a design based on annual maximum flood discharges may not be sufficient. For this reason, the partial duration series (peaks-over-threshold level) method is used. One of the most important points in the partial duration series is the threshold level (x_0) selection. For this purpose, the frequency factor equation 2 given by Rosbjerg & Madsen (1992) was used in the study.

$$x_0 = E\{Q\} + kS\{Q\} \qquad k \geq 1 \tag{2}$$

Where: $E\{Q\}$= average daily flow rates (m^3/s), $S\{Q\}$= standard deviation of daily flow rates (m^3/s), k= frequency factor, x_0= threshold level (m^3/s)

2.2.3. Parameter estimation methods

Given that it's not feasible to determine all aspects of a random variable, it becomes the responsibility of researchers, statisticians, and students in probability and statistics to estimate the parameters of the probability distribution function using the sample of the variable. The shape, scale, skewness, kurtosis, and symmetry in the probability distribution function are all tied to these parameters. They are crucial in defining the statistical characteristic of a random variable in a given distribution (Önöz 1994). This research focused on the sequence models' probability-weighted moments, L-moments, and L-moment ratios, as outlined by Kjeldsen et al. (2002).

2.2.4. Probability-weighted moments

This method, developed by Greenwood et al. (1979), was later investigated by Hosking (1986). As a result of the research, it was observed that central statistics have similar properties to moments. It has been concluded that while the estimates of the mentioned moments are unbiased, especially for small samples, they are sensitive to values at data points that deviate significantly from the trend of the existing data. Modifications and distortions in inappropriate data significantly affect the statistical parameters calculated for small data. In addition, this technique shows a much lower effect from sampling changes than other moments since they are a linear data function. This feature provides superiority over other methods (Önöz 1994). Probability weighted moments are;

$$M_{p,r,s} = E\,[X^p\{F(X)\}^r\{1 - F(X)\}^s] \tag{3}$$

In Equation 3, X represents the statistical data, and $F(X)$ represents the cumulative distribution function of X. Probability-weighted moments, $\alpha_r = M_{1,0,r}$, and $\beta_r = M_{1,r,0}$ are used for the minimum probability of occurrence (exceedance) and the maximum probability of occurrence (non-exceedance), respectively. Probability-weighted moments of α_r and β_r are used as the basic method in parameter estimation of probability distributions. The probability-weighted moment of β_r is used in the case of ascending order of data, and a weighted moment of α_r is used in the case of descending order. In this research, the weighted moment of β_r is used and given in Equation 4 (Hosking 1986);

$$\beta_r = E[X\{F(X)\}^r] \qquad r = 0, 1, 2, \ldots \tag{4}$$

In Equation 4, the probability-weighted moment β_r is equal to the product of the powers (r) of the cumulative distribution function $F(X)$ of the X data. Here, the $F(X)$ function represents the probability function where X is weighted differently for different r values. For the value r= 0, the β_0 value is equal to the population mean ($E(X)$) (Haktanır & Bozduman 1995).

However, since it is difficult to specify the scale and shape of a probability distribution with probability-weighted moments directly, some linear combinations of these moments were created using the ordinal statistics explained below (Haktanır 1991; Seçkin & Topcu 2016).

2.2.5. L-moments method

Greenwood et al. (1979) express a linear function of the probability-weighted moments. This technique was developed through Hosking (1990). The observation statistics of the method, widely used in solving various problems such as regionalization and distribution, can be easily calculated without the need to square and cube the data. Although this technique is less sensitive to long-term data than normal product moments, it is generally similar to known moments. Therefore, the L-moment function of a data X is expressed as probability-weighted moments. From the observations here, X (n:j) is an *unbiased sample estimate* of the probability weight moments obtained by Greenwood et al. (1979), which is expressed in the following equation;

$$b_r = j^{-1} \sum_{n=1}^{j} \frac{(n-1)(\mathrm{n}-2)\ldots(\mathrm{n}-\mathrm{r})}{(j-1)(j-2)\ldots(j-r)} x_{n:j} \quad (5)$$

First four b_r values (r= 0, 1, 2, 3) and probability weighted moments (b_0, b_1, b_2, and b_3) are found, then L-moment statistics for any distribution,

$$\ell_1 = b_0, \quad (6)$$

$$\ell_2 = 2b_1 - b_0, \quad (7)$$

$$\ell_3 = 6b_2 - 6b_1 + b_0, \quad (8)$$

$$\ell_4 = 20b_3 - 30b_2 + 12b_1 - b_0 \quad (9)$$

The first L-moment, ℓ_1 is a measure of central tendency and is equivalent to the mean of the distribution. The measure of distribution is ℓ_2. The estimated dimensionless L-moment ratio;

$$t = \frac{\ell_2}{\ell_1} (L - coefficient\ of\ variation, L - Cv) \quad (10)$$

$$t_3 = \frac{\ell_3}{\ell_2} (L - skewness, L - Cs) \quad (11)$$

$$t_4 = \frac{\ell_4}{\ell_2} (L - kurtosis, L - Ck) \quad (12)$$

The above-mentioned b_1 and b_2, L-coefficient of variation, t and t_3 and t_4, in other words, L-moment ratios, are the most useful measures that briefly express probability distributions. L-moments of different distributions can be easily represented with the L-moment ratio diagram. In two-parameter distributions, only two values expressed as measurement and position parameters are distributed in any way. L-kurtosis and L-skewness values are equal and are indicated in the diagram with only one dot. In a three-parameter distribution, the parameters included are scale, location, and shape. In these three-parameter distributions, different points are determined with varying shape parameters to define them with a line (Hosking & Wallis 1997, Topcu & Seçkin 2016).

2.2.6. Index-flood method

In the index-flood method, the hypothesis is that the recurrence distributions of the data at all stations are the same, except for a scale factor. In the process, stations are methodically and rigorously divided into homogeneous regions by various analyses, ensuring the reliability of the results. It is extremely effective in combining summary statistics of separate data sets. The index-flood approach (Dalrymple, 1960) and its variants (Basu & Srinavas 2016, Stedinger 1983, Sveinsson et al. 2003, Öney & Anlı 2023) are widely used for regional frequency analysis. If there is a station i with N data in a basin with n_i stations and it is assumed that these data are defined in the form Q_{ij}, j= 1,..., n_i; Q_i (F); The assumption expressed as a quantile function of the non-exceedance probability of station's data i (F);

$$Q_i(F) = \mu_i\, q(F), \qquad i = 1,\ldots, N. \quad (13)$$

In this equation, μ_i; the index-flood value represents the average probability distribution at the station. This value represents the rainfall and surface flow characteristics in each basin and. it defines the regional growth curve of the non-exceedance probability of the q (F) value (F), which is equal for all stations. By multiplying the q (F) value obtained in the regional frequency analysis with the average of the desired station (F), the Q_i (F) value of the hydrological variable at the station to which it belongs is reached for the return period. Common regional frequency distribution function applied to dimensionless data;

$$q(F) = Q_{ij}/\mu_i \quad (14)$$

2.3. Stages of regional frequency analysis

Hosking & Wallis (1993) define the stages used in regional frequency analysis as follows: Preliminary statistical analysis of the data (Discordancy measure), determination of hydrologically homogeneous regions (Heterogeneity measure), determination of the best regional probability distribution (Goodness-of-fit measure), and development of the regional probable flood discharges (Regional L-moment algorithm).

2.3.1. Discordancy measure

The Discordancy measure (D_i) plays a crucial role in ensuring the accuracy of the data in regional frequency analysis. By reviewing the collected values, this measure enables the elimination of major inaccuracies and the detection of incompatibilities, thereby ensuring the compatibility of the stations separated as homogeneous regions.

$$D_i = \frac{1}{3} N(u_i - \bar{u})^T K^{-1}(u_i - \bar{u}) \tag{15}$$

According to the equation, u_i represents the vector of L-moment ratios at a particular station, $\bar{u}$ represents the mean of the vector, and K represents the covariance matrix of this vector. The fact that the D_i is higher than the critical tabular value, which varies depending on the number of stations in the region, leads to the conclusion that the station is completely discordant (Hosking & Wallis 1997).

2.3.2. Heterogeneity measure

When determining the homogeneity of a region based on the discordancy measure, it's crucial to calculate the heterogeneity of the groups at the stations. This involves comparing the L-moment variations of regions that are likely to be particularly homogeneous at stations. To calculate the heterogeneity measure, we first assume the basins to be homogeneous. We then compute the mean and standard deviations of the selected dispersion measure by simulating the values of a station in a homogeneous region in similar observations. The H statistic is then used to compare the dispersion measures obtained.

$$H = \frac{(V_{obs} - \mu_v)}{\sigma_v} \tag{16}$$

$$V_{obs} = \left\{ \frac{\sum_{i=1}^{N} n_{i\left(\tau_2^i - \tau_2^R\right)^2}}{\sum_{i=1}^{N} n_i} \right\}^{\frac{1}{2}} \tag{17}$$

While the weighted standard deviation calculated from regional data by looking at various L-moment ratios is expressed with the V_{obs} statistic, μ_v and σ_v represent the average and standard deviation values of the number of simulations of this statistic. N is the number of stations; n_i is the sample size of any station, τ_2^i is the sample L-coefficient of variation of any station, and τ_2^R is the regional sample L-coefficient of variation. The difference between the L-moment ratios at the stations and the regional L-moment ratios calculates the V_{obs} statistic. In Equation 17, only V_{obs} related to the H_1 statistic is given. Equations H_2 and H_3 are calculated similarly. In this research, the four-parameter Kappa distribution was used. Compared to two- and three-parameter distributions, this distribution is stronger since it expresses more than one distribution in the frequency analysis of hydrological events. For μ_v and σ_v values to provide more accurate results, the number of simulations will be considered 500 for a region. According to all these statements, the basin: If $H < 1$, it is acceptably homogeneous; if $1 < H < 2$, it may be heterogeneous; and if $H > 2$, it is not homogeneous. An attempt is made to achieve homogeneity by dividing a un homogeneous region into sub-regions.

2.3.3. Goodness-of-fit measure

Only a probability distribution shows the best fit for the data obtained from homogeneous stations. A Z^{DIST} statistic measure has been recommended for the goodness-of-fit criterion related to the L-kurtosis ratio and a random probability distribution. The equation of this method is as follows;

$$Z^{DIST} = \frac{(\tau_4^{DIST} - t_4^R + B_4)}{\sigma_4} \tag{18}$$

$$B_4 = N_{sim}^{-1} \sum_{m=1}^{N_{sim}} \left(t_4^{(m)} - t_4^R \right) \tag{19}$$

$$\sigma_4 = \left[\left(N_{sim} - 1 \right)^{-1} \left\{ \sum_{m=1}^{Nsim} \left(t_4^{(m)} - t_4^R \right)^2 - N_{sim} B_4^2 \right\} \right]^{1/2} \tag{20}$$

In this equation, t_4^R is defined as the L-kurtosis ratio of the regional mean value of the sample, while B_4 is the bias value and σ_4 is its standard deviation. In Equations 19 and 20, N_{sim} is defined as the number of simulations made using the Kappa four-parameter distribution, and m is defined as the number of areas simulated. In this study, Generalized Extreme Values (GEVs), Generalized Logistic (GLO), Generalized Pareto (GPA), Pearson Type 3 (PE3), and Generalized normal (GNO) distributions were used. Their cumulative distribution F (x) and quantile x (F) functions are given in Table 2. If the absolute $Z^{DIST} \leq 1.64$ in

any distribution, this distribution is considered suitable for regional distribution. This value corresponds to a 90% confidence level. However, among the distributions examined, the absolute Z^{DIST} value closest to zero is determined as the best-fit distribution.

Table 2- Cumulative distribution *F* (*x*) and quantile *x* (*F*) functions of the regional frequency distributions used in the study

Distribution	*Code*	*F (x), x (F)*
Generalized Extreme Values	GEVs	$F = \exp[-\{1-k(x-\xi)/\alpha\}^{1/k}]$ $x = \xi + \alpha\{1-(-\log F)^k\}/k$
Generalized Logistic	GLO	$F = 1/[1+\{1-k(x-\xi)/\alpha\}^{1/k}]$ $x = \xi + \alpha[1-\{(1-F)/F\}^k]/k$
Generalized Normal	GNO	$F = \Phi[-k^{-1}\log\{1-k(x-\xi)/\alpha\}]$ $x(F)$ not precisely defined
Generalized Pareto	GPA	$F = 1-\{1-k(x-\xi)/\alpha\}^{1/k}$ $x = \xi + \alpha\{1-(1-F)^k\}/k$
Pearson type 3	PE3	$F = G\left((x-\mu+2\sigma/\gamma)/\lvert\frac{1}{2}\sigma\gamma\rvert, 4/\gamma^2\right), \quad \gamma > 0$ $F = 1-G\left(-(x-\mu+2\sigma/\gamma)/\lvert\frac{1}{2}\sigma\gamma\rvert, 4/\gamma^2\right), \quad \gamma < 0$ $x(F)$ not precisely defined
Kappa	KAP	$F = [1-h\{1-k(x-\xi)/\alpha\}^{1/k}]^{1/h}$ $x = \xi + \alpha[1-\{(1-F)^h/h\}^k]/k$
Wakeby	WAK	$F(x)$ not precisely defined $x = \xi + \frac{\alpha}{\beta}\{1-(1-F)^{\beta}\} - \frac{\gamma}{\delta}\{1-(1-F)^{-\delta}\}$

$G(x,\alpha) = \{\Gamma(\alpha)\}^{-1}\int_0^x t^{\alpha-1}e^{-t}dt$ missing gamma integral.

$\Phi(x) = (2\pi)^{-1/2}\int_{-\infty}^{x}\exp(-t^2/2)dt$ standard normal cumulative distribution function.

2.3.4. Regional L-moment algorithm

The study used a regional algorithm that combines at-site statistics of L-moments with index-flood and weighted average methods. In the regional L-moment algorithm stage, a probability distribution adapted to the homogeneous region value is selected. The means of the probability distributions at all stations, accepted as index-flood data, are reliable and provide a confident basis for the data at the stations obtained by the sample mean of the at-site value.

The sample mean of a station with n_i data in an area with N stations is stated as ℓ_1^i. Sample L-moment ratios are obtained as $t^{(i)}, t_3^{(i)}, t_4^{(i)}$. The average L-moment ratio of each station during the observation period is calculated in the form t^R, t_3^R, t_4^R;

$$t^R = \frac{\sum_{i=1}^{N} n_i t^{(i)}}{\sum_{i=1}^{N} n_i} \tag{21}$$

Taking the regional average $\ell_1^R = 1$

$$t_r^R = \frac{\sum_{i=1}^{N} n_i t_r^{(i)}}{\sum_{i=1}^{N} n_i} \qquad r = 3, 4,\ldots \tag{22}$$

and from here the regional population (λ_i and τ_i) and sample L-moment ratios (ℓ_1^R, t_1^R) are equalized;

$$\lambda_1 = \ell_1^R$$
$$\tau = t^R \quad (23)$$
$$\tau_3 = t_3^R$$

Flood quantiles at the desired probability and return period along with dimensionless regional growth curves;

$$\hat{Q}_i(F) = \ell_1^i q(F; \ell_1^R, t^R, t_3^R, t_4^R) \quad (24)$$

FORTRAN 77 source codes written through Hosking (2005) were used in all these calculations, and (l-moments, version 3.04) routines were used. The routines collected under a basic program were arranged and executed.

3. Results and Discussion

3.1. Regional frequency analysis with annual maximum flood series

The annual maximum flood series were selected using the annual instantaneous maximum flood data in the DSI flow observation yearbook of the stations in the research. As a result of the discordancy measure (D_i) test determined using this data set, no discordant stations were detected in the annual maximum series within the 22 stations; in other words, it was concluded that the basin was homogeneous in a single region case (Table 3). The discordancy measure test consists of a test that determines the harmony between stations' L-moment ratios. In this study, among the annual maximum flood series obtained from 22 stations, no discordancy was detected between the L-moment ratios of any station flood data, based on the $D_{critical\ value}$ given in Hosking & Wallis (1997) and varying according to the number of stations in the region (All $D_i < D_{critical\ value}$: 3.00). This shows that the annual maximum flood data in the Susurluk basin does not have any discordant conditions in a single region. After this stage, whether the basin was homogeneous or not was tested.

On the other hand, when the L-moment ratios are examined in Table 3, the annual maximum flood data is skewed to the right at all stations. L-kurtosis ratios show that the flood data is flat at some stations and sharp at others. Accordingly, the annual maximum flood data is not normally distributed (Ahmad et al. 2019).

After no discordant station was achieved, the heterogeneity measure (H) was calculated. The heterogeneity measure compares the inter-site variations in sample L-moments for the group of sites with what would be expected of a homogeneous region. While H_1 and H_2 measures provided homogeneity (< 1.00) in the calculations, the H_3 measure showed possible heterogeneity case (> 1.00). However, since H_1 and H_2 measures provide the homogeneity conditions, the basin is accepted as homogeneous (Table 4). When Table 4 is examined, although the H_3 test value shows that the region is probably heterogeneous, it is clear that the region is homogeneous since the H_1 and H_2 test values are very close to zero. Negative heterogeneity measures (H_1 and H_2) indicate that the spread between stations is quite low and homogeneity is strong. It has been stated that regional frequency analysis is more effective than at-site frequency analysis, even if the basin considered in many regionalization studies is somewhat heterogeneous (Hosking & Wallis 1997; Anlı 2009; Van Campenhout et al. 2020; Kiran & Srinavas 2021; Pan & Rahman 2021).

Susurluk basin was obtained as a homogeneous region according to the annual maximum flood series as a single region, generalized logistic (GLO), generalized extreme values (GEVs), generalized normal (GNO), generalized Pareto (GPA) and Pearson type 3 (PE3) distributions were selected as candidates to estimate frequency distributions suitable for annual instantaneous flood discharges. The generalized logistic (GLO) distribution emerged as the only suitable distribution for annual maximum flood discharges among these distributions (Ahmad et al. 2019). Other frequency distributions still need to be able to describe the annual maximum flood data. L-kurtosis and Z^{DIST} (Goodness-of-fit measure) values found according to the candidate distributions are shown in Table 5. Regional L-kurtosis ratios of frequency distributions show that annual maximum floods are flat. Design flood discharges estimated by the regional L-moment algorithm obtained according to various recurrence probabilities and return periods are shown in Table 6. According to Table 6, the average (Q_2) design flood flow discharge is approximately 128 m^3/s. The spillway design discharge of small water storage structures ($Q_{50\text{-}100}$) is approximately 422-520 m^3/s, the diversion channel discharge (Q_{25}) is approximately 341 m^3/s, and the main irrigation channel project (Q_{10}) is approximately 254 m^3/s. It can be approximately 80-88 m^3/s for secondary and tertiary channel projects ($Q_{1.25\text{-}1.33}$) and 21-60 m^3/s for city drainage networks ($Q_{1.01\text{-}1.11}$). Since the basin is homogeneous according to the annual maximum flood data, design flood flow values can be considered for the entire Susurluk basin, but in the following section, at-site frequency analysis was carried out for each station separately with the L-moment parameter method.

Table 3- Mean, L-moment ratios, and discordancy measures (D_i) calculated for annual maximum flood series

N	*Station code*	*Mean*	*L-coefficient of variation*	*L-skewness*	*L-kurtosis*	D_i
1	D03A008	378.32	0.2543	0.0945	0.1316	0.91
2	D03A013	69.36	0.6664	0.5676	0.3566	1.40
3	D03A024	132.53	0.3590	0.3663	0.2389	0.32
4	D03A034	166.35	0.3821	0.2519	0.1684	0.47
5	D03A044	84.21	0.5180	0.4428	0.2348	1.19
6	D03A051	52.97	0.4480	0.3853	0.3967	1.49
7	D03A052	165.84	0.4002	0.4371	0.3609	0.77
8	D03A056	13.13	0.4936	0.4926	0.2923	1.01
9	D03A064	56.59	0.3263	0.2473	0.2002	0.26
10	D03A082	27.58	0.3619	0.3076	0.3098	0.66
11	D03A084	57.82	0.3064	0.1160	0.0454	2.17
12	D03A085	4.79	0.3793	0.3840	0.3660	1.15
13	D03A086	11.78	0.2459	0.1683	0.1785	0.79
14	D03A087	62.70	0.3110	0.1778	0.1524	0.56
15	D03A096	10.50	0.3156	0.2827	0.2343	0.39
16	E03A002	435.79	0.3304	0.2711	0.2159	2.16
17	E03A011	48.16	0.4295	0.3469	0.2488	0.19
18	E03A016	518.34	0.2653	0.1392	0.2493	2.21
19	E03A017	410.88	0.2550	0.0999	0.1681	0.98
20	E03A024	221.36	0.2314	0.0040	0.1552	1.56
21	E03A028	90.60	0.5138	0.4848	0.3694	0.77
22	E03A031	83.42	0.3230	0.2251	0.2716	0.58
	Weighted mean	148.44	0.3697	0.2852	0.2427	

Table 4- Heterogeneity measure results for annual maximum flood series

Heterogeneity measure	*Value*
H_1	-0.0777
H_2	-0.0553
H_3	1.1551*

*Possibly heterogeneous

Table 5- Regional L-kurtosis and goodness-of-fit (Z^{DIST}) results for annual maximum flood series

Regional frequency distribution	*L-kurtosis*	Z^{DIST}
GLO	0.234	-1.49*
GEVs	0.206	-2.72
GNO	0.187	-3.56
PE3	0.153	-5.02
GPA	0.131	-5.97

*Suitable distribution

Table 6- Regional flood discharges estimated index-flood method according to various recurrence probabilities and return periods according to the generalized logistic frequency distribution for annual maximum flood series (m^3/s)

***P* %**	1	5	10	20	25	50	80	90	96	98	99
T_{year}	1.01	1.05	1.11	1.25	1.33	2	5	10	25	50	100
***Q* (*F*)**	20.99	44.54	59.43	79.62	87.98	127.8	197.44	253.53	341.22	422.35	520.18

3.2. Regional frequency analysis with partial duration flood series

While partial duration flood series were extracted using the variance-mean ratio method, the mean and standard deviation of daily flow discharges measured at flow observation stations were calculated. According to the values obtained from daily flow discharges at flow observation stations, the standard deviation of daily flow discharges at all stations was greater than the means. If the standard deviation of a data is greater than its mean, it indicates that the data is heterogeneously distributed. While determining the threshold level, a threshold value was determined for each station using various frequency factor values (k=1, 2, 3, 3.5, 4.0, 4.5, and 5). When the frequency factor was used with a value greater than 5, the selected threshold level values were quite high. Hence, the number of above-threshold flood data was lower than the annual maximum flood series, and the $m > n$ condition was not provided. Since the threshold levels remained very low at k=1 and k=2, this caused much more flood data to be selected and made the study difficult as the number of samples increased greatly.

To determine the most suitable frequency factors according to stations, discrete probability distributions, Poisson, Binomial, and negative Binomial, were applied, and the Kolmogorov-Smirnov goodness-of-fit test determined suitable distributions. The aim is to ensure that the ratio of the variance of the peak numbers that exceed a certain threshold level and occur each year to their average is equal to or very close to one. As a result of the analyses, a frequency factor was determined for each station, prioritizing compliance with the Poisson distribution (Cunnane 1979; Ben-Zvi 1991; Rasmussen & Rosbjerg 1991; Bezak et al. 2014), Table 7 shows the Kolmogorov-Smirnov test results obtained from trial-and-error to choose the most appropriate discrete distribution for the frequency factors used in selecting the threshold level. After determining the threshold level for each station, the flood discharges above the threshold level were determined, and partial duration flood series were extracted. The Binomial distribution applied for partial duration flood discharges above the threshold levels obtained by the frequency factors providing the variance-mean ratio did not fit any data set in Table 7. The negative Binomial distribution fits very few data sets (Bezak et al. 2014). However, since the priority was the Poisson distribution, the analysis progressed according to this distribution's goodness of fit test values. The frequency factor (k=3) was not determined for any station, while k=3.5 was determined for only two stations. k=4 was suitable for seven stations, and k=5 was suitable for thirteen stations. On the other hand, the k=5 value chosen at station E03A017, the total m value was smaller than the annual maximum flood series ($m < n$), so a frequency factor of k=4.5 was considered. Since high-frequency factors are considered in most of the stations, the number of partial-duration flood discharges obtained is at a reasonable level. Threshold levels are determined using daily flow discharges, the number of partial duration flood data extracted according to these threshold levels, the average number of peaks over the threshold level, their means, L-moment ratios, and discordancy measures (D_i) are given in Table 8. Figure 2 also compares the sample size of the annual maximum and partial-duration flood series.

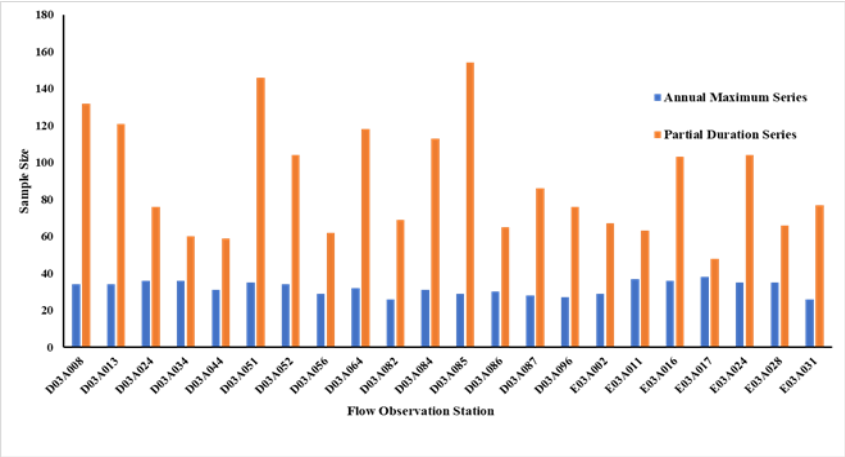


Figure 2- Sample size comparison of annual maximum and partial duration flood series.

The lowest threshold level was determined at station D03A085 with 1.64 m^3/s, and the highest threshold level was determined at station E03A017 with 612.6 m^3/s. E03A017 station is at the basin outlet point, and its elevation is 2 m. When the sample size of partial-duration flood extracted is examined, the average number of peaks over the threshold level (λ`) varies between 1.26 and 5.31 (Pham et al. 2014). Of course, the lowest and highest threshold levels are inversely proportional to ($\lambda$`) values. Therefore, the lowest flood data (m= 48) was extracted at station E03A017, and the highest flood data (m= 154) was extracted at station D03A085. The mentioned situation can also be seen in Figure 2. L-moment ratios show that the partial duration flood data is mostly right-skewed, flattened, and not normally distributed. When all partial duration series obtained from 22 stations in the Susurluk basin were subjected to regional analysis as a single region, station E03A002 was found to be discordant with the test value 4.52. L-moment ratio diagrams are given in Figure 3. In Figure 3, the discordant station E03A002 is circled. Since this station is important due to its observation period and location, it was not removed from the data set, and no outliers were sought. Therefore, according to Ward's Linkage Squared Euclidean Distance method, the basin was divided into sub-regions with cluster analysis.

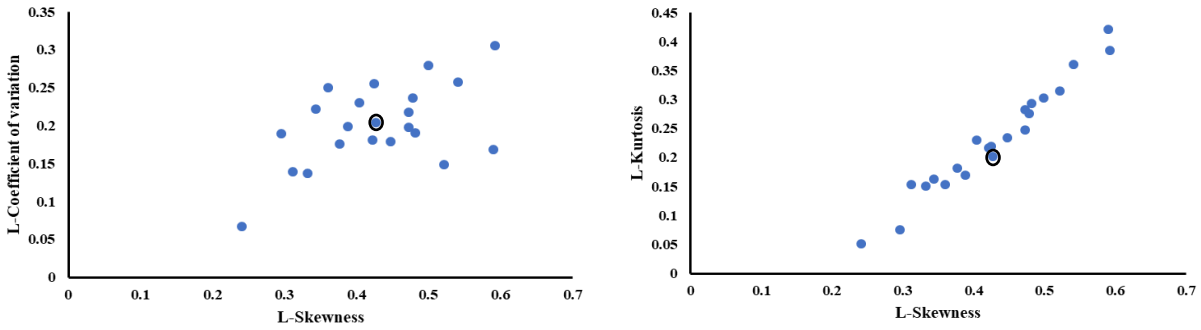


Figure 3- L-moment ratio diagrams with L-skewness versus L-coefficient of variation and L-kurtosis for E03A002 discordancy situation

The basin is subjectively divided into three sub-regions according to the cluster dendrogram in Figure 4. Sub-basin physiographic-hydrological characteristics used in the vectors of cluster analysis: elevation, latitude, longitude, precipitation area, maximum and minimum elevation of watersheds, watershed slope, and long-term average flow values. There were five stations in Region I, six in Region II, and eleven in Region III. While Regions I and II had lower-elevation flat stations, Region III had stations in the mountainous parts of the basin. This situation facilitated partial duration flood series for examining the basin in two geographical ways: flat and mountainous (Pan & Rahman 2021). Figure 5 shows the location of the stations in the homogeneous regions obtained according to the partial-duration flood series.

On the other hand, L-moment ratios show that the data is mostly right-skewed, flattened, and not normally distributed. The basin was subjected to regional analysis in its current form, and no discordant station was found in the three regions (Table 9). By performing cluster analysis, small sub-regions with fewer stations were obtained instead of having stations in a single region as a large cluster. Therefore, the harmony between fewer stations was investigated. Thus, no discordant station was detected in all three regions obtained. After no discordant station was achieved for each of the three regions, the heterogeneity measure values (H) were calculated and given in Table 10. While H_1 and H_2 measures provided homogeneity (< 1.00), the H_3 measure showed definitely heterogeneity case (> 2.00) for Region I. However, since H_1 and H_2 measures provide the homogeneity conditions, Region I is accepted as homogeneous. All heterogeneity measure results for Regions II and III revealed that these regions were acceptably homogeneous. Since all remaining H values, except the H_3 value in the first region, are very close to zero, it is clear that all three regions are satisfactorily homogeneous.

Three-parameter GLO, GEVs, GNO, PE3, and GPA probability distributions were applied to three separate homogeneous regions according to partial duration flood series, and the regional L-kurtosis and goodness-of-fit test (Z^{DIST}) results for estimating frequency distributions are given in Table 11. In Table 11, PE3 and GPA distributions in the first and second regions fit the data as expected. In contrast, GPA and GNO distributions in the third region fit the partial duration flood series. Rosbjerg (1992), Madsen et al. (1997a), Madsen et al. (1997b), Rahman et al. (2013), Guru (2016), Agilan et al. (2020), Amorim & Villarini (2024) also fitted the GPA distribution to the partial duration series. Unlike these studies, the partial-duration flood series were also fitted to the PE3 distribution in our current study. Regional L-kurtosis ratios show that the data is almost flat. Design flood discharges estimated by regional L-moment algorithm obtained according to various recurrence probabilities and relevant return periods are shown in Table 12. In Figure 6, the regional design flood discharges of annual maximum and partial duration flood series are compared to various return periods.

Table 7- Kolmogorov-Smirnov test results according to certain frequency factors and discrete probability distributions in the selection of partial duration flood series for all stations

N	*Station code*	*Discrete Probability Distribution, K-S*											
		Poisson				*Binomial*				*Negative Binomial*			
		Frequency Factor (k)											
		3.0	*3.5*	*4.0*	*5.0*	*3.0*	*3.5*	*4.0*	*5.0*	*3.0*	*3.5*	*4.0*	*5.0*
1	D03A008	0.458	0.347	0.333*	0.355	NF	NF	NF	NF	NF	NF	NF	NF
2	D03A013	0.537	0.518	0.485*	0.554	NF	NF	NF	NF	NF	NF	NF	NF
3	D03A024	0.302	0.312	0.270	0.226*	NF	NF	NF	NF	0.204	0.194	0.219	NF
4	D03A034	0.455	0.405	0.423	0.294*	NF	NF	NF	NF	NF	NF	NF	NF
5	D03A044	0.436	0.432	0.394*	0.409	NF	NF	NF	NF	NF	NF	NF	NF
6	D03A051	0.256	0.232	0.204*	0.214	NF	NF	NF	NF	0.291	0.360	0.397	0.353
7	D03A052	0.337	0.298	0.205*	0.241	NF	NF	NF	NF	0.251	0.263	0.322	NF
8	D03A056	0.354	0.310	0.334	0.278*	NF	NF	NF	NF	NF	NF	NF	NF
9	D03A064	0.251	0.166*	0.188	0.253	NF	NF	NF	NF	0.373	0.393	0.440	0.283
10	D03A082	0.316	0.282	0.246	0.206*	NF	NF	NF	NF	NF	NF	NF	NF
11	D03A084	0.287	0.251	0.176*	0.188	NF	NF	NF	NF	0.283	0.243	0.332	0.327
12	D03A085	0.396	0.386*	0.475	0.420	NF	NF	NF	NF	NF	NF	NF	NF
13	D03A086	0.336	0.307	0.293	0.240*	NF	NF	NF	NF	0.313	0.252	NF	0.387
14	D03A087	0.308	0.296	0.350	0.291*	NF	NF	NF	NF	0.172	NF	NF	NF
15	D03A096	0.362	0.307	0.278	0.259*	NF	NF	NF	NF	0.238	NF	0.260	NF
16	E03A002	0.399	0.361	0.331	0.250*	NF	NF	NF	NF	NF	NF	NF	NF
17	E03A011	0.548	0.565	0.614	0.546*	NF	NF	NF	NF	NF	NF	NF	NF
18	E03A016	0.350	0.342	0.329	0.328*	NF	NF	NF	NF	NF	NF	NF	NF
19	E03A017	0.637	0.701	0.770	0.612**	NF	NF	NF	NF	NF	NF	NF	NF
20	E03A024	0.313	0.293	0.273	0.232*	NF	NF	NF	NF	NF	NF	NF	NF
21	E03A028	0.464	0.504	0.395	0.318*	NF	NF	NF	NF	NF	NF	NF	NF
22	E03A031	0.458	0.347	0.333*	0.355	NF	NF	NF	NF	NF	NF	NF	NF

K-S: Kolmogorov-Smirnov test value; *Best-fit discrete distribution and relevant frequency factor; **While determining the threshold level with a frequency factor of k = 5.0, since the number of peaks exceeding the threshold level was obtained less than the observation period, a frequency factor of k = 4.5 was considered; NF: Non-fit discrete probability distribution

According to Table 12a, the average (Q_2) design flood flow discharge is approximately 62.15 m^3/s. The spillway design discharge of small water storage structures ($Q_{50\text{-}100}$) is approximately 170-195 m^3/s, the diversion channel discharge (Q_{25}) is

approximately 146 m^3/s, and the main irrigation channel project (Q_{10}) is approximately 114 m^3/s. It can be approximately 50-51 m^3/s for secondary and tertiary channel projects ($Q_{1.25\text{-}1.33}$) and 45-46 m^3/s for urban drainage networks ($Q_{1.01\text{-}1.11}$) for the first region. For the second region, the average (Q_2) design flood discharge was estimated to be approximately 210.65 m^3/s. Spillway design flood discharge of small dam structures ($Q_{50\text{-}100}$) is approximately 442-495 m^3/s, and diversion channel flood discharge (Q_{25}) is approximately 390 m^3/s. The main irrigation canal project (Q_{10}) is approximately 322 m^3/s. It can be evaluated as approximately 184-187 m^3/s in secondary and tertiary channel projects ($Q_{1.25\text{-}1.33}$) and approximately 176-179 m^3/s in urban drainage networks ($Q_{1.01\text{-}1.11}$) (Table 12b). When the third region design flood rates were examined, the Q_2 design flood flow rate was estimated to be approximately 36 m^3/s. The spillway design flood discharge of pond structures ($Q_{50\text{-}100}$) is about 112-139 m^3/s, and the diversion channel flood discharge (Q_{25}) is about 89 m^3/s. The main irrigation canal project (Q_{10}) is approximately 65 m^3/s; in secondary and tertiary canal projects ($Q_{1.25\text{-}1.33}$), it is approximately 30 m^3/s, and in urban drainage networks ($Q_{1.01\text{-}1.11}$), it is approximately 27 m^3/s (Table 12c). Among the three regions, the lowest design flood discharges were obtained in the third region. This is because the third region is in the mountainous regions on the upstream side of the basin, and the streamflow discharges have not yet been collected. The second region with the highest flood values is near the basin outlet point (downstream) (Pan & Rahman 2021). Annual maximum flood discharges (AMS) obtained from a single region according to various return periods with the regional L-moment algorithm were compared with partial duration flood discharges (PDS-I, PDS-II, and PDS-III) determined separately for three regions. Among them, the highest design flood discharges were obtained in PDS-II during the remaining return periods, except for Q_{100}, the upper tail of the distribution. The highest flood discharge in Q_{100} was obtained in AMS. AMS, PDS-I, and PDS-III are estimated close to each other in the lower tail of the distribution at $Q_{1.01}$. In $Q_{1.05}$, AMS and PDS-I were almost the same, while PDS-III was slightly lower.

Table 8- Threshold levels determined by using daily flow rates (x_0), the number of partial duration flood data extracted according to these threshold levels (m), the average number of peaks over the threshold level (λ`), their means, L-moment ratios and discordancy measures (D_i)

N	*Station code*	*Sample size (n)*	*Threshold level (x_0)*	*Partial duration sample size (m)*	*λ` (m/n)*	*Mean*	*L-coefficient of variation*	*L-skewness*	*L-kurtosis*	*D_i*
1	D03A008	34	69.483	132	3.88	122.77	0.2214	0.3434	0.1622	0.56
2	D03A013	34	26.204	121	3.56	45.14	0.2369	0.4784	0.2760	0.14
3	D03A024	36	30.160	76	2.11	49.53	0.2174	0.4721	0.2819	0.15
4	D03A034	36	75.680	60	1.67	112.81	0.1759	0.3763	0.1821	0.12
5	D03A044	31	31.580	59	1.90	57.47	0.3057	0.5924	0.3844	1.21
6	D03A051	35	5.640	146	4.17	10.12	0.2553	0.4251	0.2198	0.30
7	D03A052	34	55.950	104	3.06	93.45	0.2570	0.5413	0.3601	1.04
8	D03A056	29	5.060	62	2.14	6.74	0.1689	0.5906	0.4215	2.02
9	D03A064	32	8.920	118	3.69	16.43	0.2498	0.3602	0.1540	0.61
10	D03A082	26	7.080	69	2.65	10.28	0.1905	0.4820	0.2938	0.29
11	D03A084	31	13.790	113	3.65	24.18	0.2297	0.4046	0.2299	0.60
12	D03A085	29	1.640	154	5.31	2.36	0.1810	0.4219	0.2167	0.36
13	D03A086	30	5.380	65	2.17	7.88	0.1790	0.4479	0.2336	0.59
14	D03A087	28	14.910	86	3.07	23.11	0.1989	0.3877	0.1691	0.91
15	D03A096	27	2.940	76	2.81	4.05	0.1389	0.3119	0.1530	2.17
16	E03A002	29	329.630	67	2.31	499.24	0.2044	0.4274	0.2020	4.52**
17	E03A011	37	38.770	63	1.70	73.70	0.2797	0.4994	0.3023	0.24
18	E03A016	36	350.600	103	2.86	473.07	0.1371	0.3321	0.1512	1.83
19	E03A017	38	612.60	48	1.26	707.25	0.0669	0.2408	0.0513	1.63
20	E03A024	35	93.490	104	2.97	149.11	0.1898	0.2961	0.0757	1.37
21	E03A028	35	37.490	66	1.89	54.38	0.1973	0.4728	0.2481	0.93
22	E03A031	26	63.570	77	2.96	85.04	0.1486	0.5213	0.3148	0.42

λ`: average number of peaks over the threshold level; ** *discordant station*

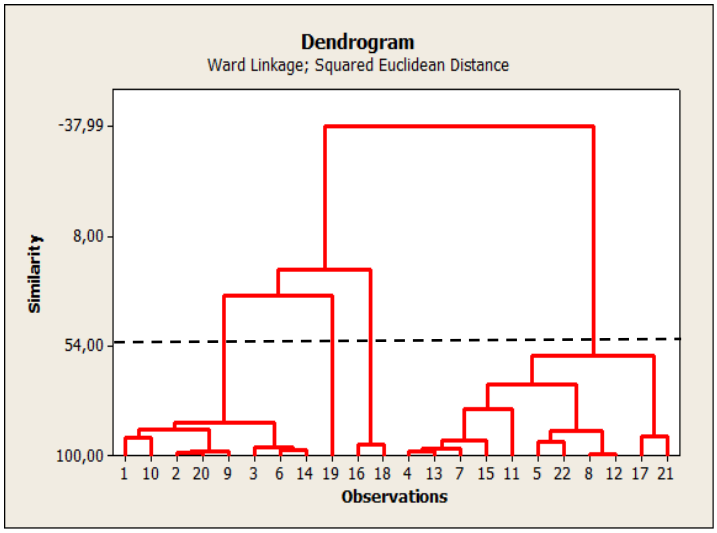


Figure 4- Dendrogram of cluster analysis

Table 9- Means, L-moment ratios, and discordancy measures (D_i) of partial duration flood series according to sub-regions

N	*Station code*	*Sample size*	*Mean*	*L-coefficient of variation*	*L-skewness*	*L-kurtosis*	D_i
			Region I				
1	D03A008	132	122.77	0.2214	0.3434	0.1622	1.27
2	D03A013	121	45.14	0.2369	0.4784	0.2760	0.90
9	D03A064	118	16.43	0.2498	0.3602	0.1540	1.13
10	D03A082	69	10.28	0.1905	0.4482	0.2938	0.59
20	E03A024	104	149.11	0.1898	0.2961	0.0757	1.10
			Region II				
3	D03A024	76	49.53	0.2174	0.4721	0.2819	1.08
6	D03A051	146	10.12	0.2553	0.4251	0.2198	0.38
14	D03A087	86	23.11	0.1989	0.3877	0.1691	0.95
16	E03A002	67	499.24	0.2044	0.4274	0.2020	1.62
18	E03A016	103	473.07	0.1371	0.3321	0.1512	0.82
19	E03A017	48	707.25	0.0669	0.2408	0.0513	1.15
			Region III				
4	D03A034	60	112.81	0.1756	0.3763	0.1821	1.25
5	D03A044	59	57.47	0.3057	0.5924	0.3844	0.76
7	D03A052	104	93.45	0.2570	0.5413	0.3601	1.23
8	D03A056	62	6.74	0.1689	0.5906	0.4215	2.31
11	D03A084	113	24.18	0.2297	0.4046	0.2299	0.44
12	D03A085	154	2.36	0.1810	0.4219	0.2167	0.71
13	D03A086	65	7.88	0.1790	0.4479	0.2236	0.90
15	D03A087	76	4.05	0.1389	0.3119	0.1530	1.84
17	E03A011	63	73.70	0.2797	0.4994	0.3023	0.47
21	E03A028	66	54.38	0.1973	0.4728	0.2481	0.87
22	E03A031	77	85.04	0.1486	0.5213	0.3148	0.20

Table 10- Heterogeneity measure (*H*) results for partial duration flood series for each region

Heterogeneity measure	*Region I*	*Region II*	*Region III*
H_1	-0.1432	-0.0740	-0.1414
H_2	-0.1347	-0.0497	-0.1453
H_3	2.6896**	0.7506	0.7787

**: Definitely heterogeneous

Table 11- Regional L-kurtosis and goodness-of-fit test (Z^{DIST}) results for estimating regional frequency distributions

Regional frequency distribution									
GLO		***GEVs***		***GNO***		***PE3***		***GPA***	
Region I									
L-kurtosis	Z^{DIST}	**L-kurtosis**	Z^{DIST}	**L-kurtosis**	Z^{DIST}	**L-kurtosis**	Z^{DIST}	**L-kurtosis**	Z^{DIST}
0.290	4.79	0.272	3.93	0.240	2.45	0.187	-0.07**	0.210	1.02*
Region II									
0.294	4.52	0.276	3.72	0.244	2.29	0.189	-0.14**	0.215	1.00*
Region III									
0.346	2.64	0.334	2.16	0.294	0.54*	0.226	-2.22	0.282	0.04**

*: Suitable distribution; **: Best-fit distribution

Table 12- Flood discharges estimated index-flood method for homogeneous regions according to various recurrence probabilities and return periods according to suitable distributions for partial duration flood series (m^3/s)

Region I (a)											
PE3 distribution											
***P* %**	1	5	10	20	25	50	80	90	96	98	99
T_{year}	1.01	1.05	1.11	1.25	1.33	2	5	10	25	50	100
***Q* (*F*)**	45.46	45.99	46.93	49.46	51.01	62.15	91.00	114.21	145.77	170.04	194.55
GPA distribution											
***P* %**	1	5	10	20	25	50	80	90	96	98	99
T_{year}	1.01	1.05	1.11	1.25	1.33	2	5	10	25	50	100
***Q* (*F*)**	44.18	45.25	46.67	49.79	51.51	62.64	89.77	112.24	144.78	171.73	200.89
Region II (b)											
PE3 distribution											
***P* %**	1	5	10	20	25	50	80	90	96	98	99
T_{year}	1.01	1.05	1.11	1.25	1.33	2	5	10	25	50	100
***Q* (*F*)**	176.03	177.05	178.92	184.07	187.28	210.65	272.13	321.93	389.83	442.16	495.06
GPA distribution											
***P* %**	1	5	10	20	25	50	80	90	96	98	99
T_{year}	1.01	1.05	1.11	1.25	1.33	2	5	10	25	50	100
***Q* (*F*)**	173.08	175.32	178.29	184.81	188.43	211.82	269.23	317.79	389.23	441.55	494.99
Region III (c)											
GPA distribution											
***P* %**	1	5	10	20	25	50	80	90	96	98	99
T_{year}	1.01	1.05	1.11	1.25	1.33	2	5	10	25	50	100
***Q* (*F*)**	26.11	27.16	28.10	29.80	30.67	36.13	50.83	65.14	89.23	112.03	139.37
GNO distribution											
***P* %**	1	5	10	20	25	50	80	90	96	98	99
T_{year}	1.01	1.05	1.11	1.25	1.33	2	5	10	25	50	100
***Q* (*F*)**	26.91	27.42	28.09	29.58	30.42	36.05	51.23	65.45	88.79	110.65	136.96

In the estimates of $Q_{(1.11,\ 1.25,\ 1.33)}$, which are the main body of the distribution, AMS, PDS-I, and PDS-III are listed from largest to smallest. From Q_2 to Q_{100}, AMS, PDS-I, and PDS-III had very high estimates. PDS-II flood discharges were estimated almost the same from $Q_{1.01}$ to $Q_{1.33}$, and increased rapidly from the main body of the distribution to the upper tail. From Q_{10} to Q_{100}, AMS and PDS-II produced similar predictions.

The values obtained as a result of regional analyses vary. However, as an alternative to annual maximum floods, the PDS-I and PDS-III data sets in the lower tail of the distribution can be used. The PDS-II data set can be used in the upper tail of the distribution (Karim et al. 2017, Zadeh et al. 2019, Ahmad et al. 2019). Since the partial duration, flood estimations are lower than annual maximum estimations, they can provide advantages for engineering projects at lower costs (Pan et al. 2022). For the main body of the distribution, the PDS-II data set is not recommended in terms of cost since it gives higher estimates than the AMS data set, except for $Q_{(25,50,100)}$ (Figure 6).

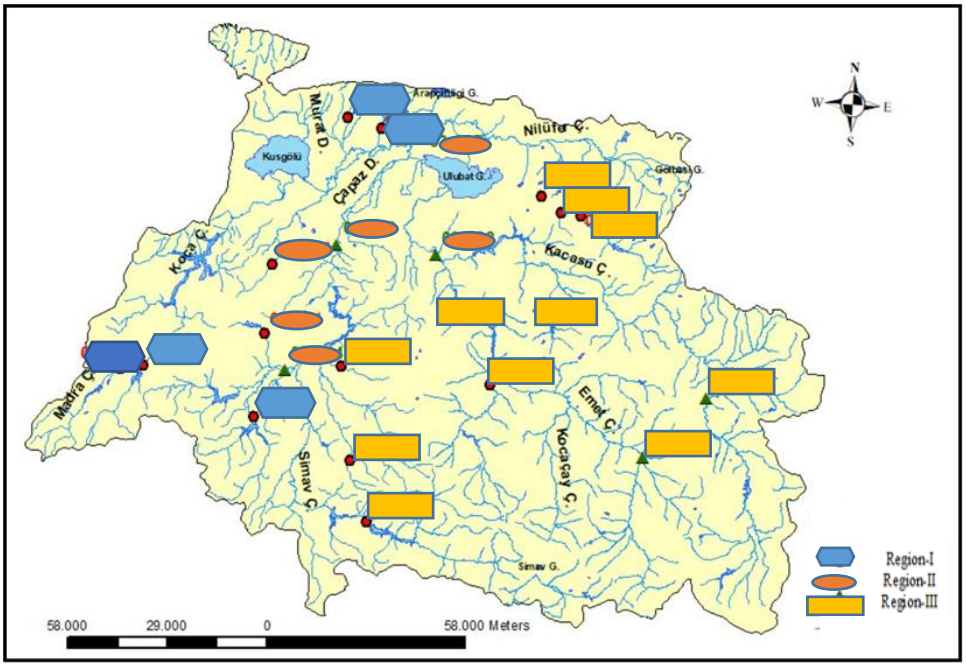


Figure 5- Location of stations in homogeneous regions obtained according to partial duration flood series

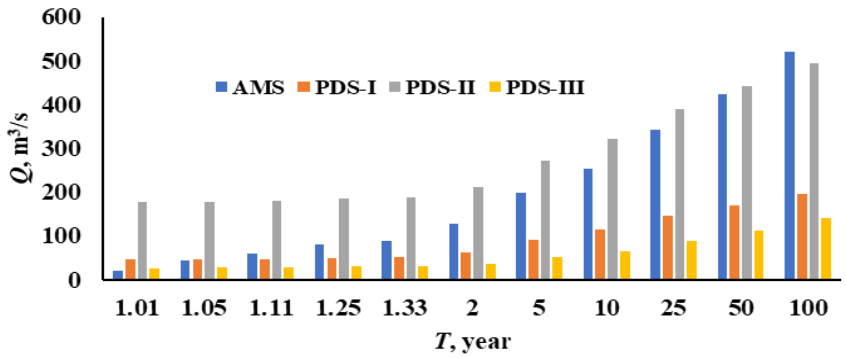


Figure 6- Comparison of the regional design flood discharges of annual maximum and partial duration flood series
***AMS: Annual maximum series, PDS-I: Partial duration series for Region I, PDS-II: Partial duration series for Region II, PDS-III: Partial duration series for Region III**

In our study, we performed regionalization of flood frequency analysis and used L-moments as the parameter estimation method. L-moments have a significant advantage over other ordinary product moments because they are calculated by linear combinations of data without squaring and cubing them. It is unbiased in its predictions of quantiles in the lower and upper tails of the distribution and provides strong estimates of the probability distribution parameters. On the other hand, it makes unbiased predictions by using L-moment ratios to measure the discordancy, heterogeneity, and goodness of fit tests. Finally, it produces highly advantageous results over at-site estimates through the regional L-moment algorithm (Anlı 2009, Zadeh et al. 2019). According to the regional homogeneity results, the design of flood discharges for both data sets obtained in various return periods according to the purpose in terms of the flood risk of the basin can be used in practice, as mentioned above. The flood discharges can be used appropriately in city drainage networks and infrastructure projects from the design of the elements of the water storage (hydraulic) structures (Durocher et al. 2019). However, it's important to note that there are a large number of flow observation stations (more than 80) in the basin, but only 22 of them could be selected because they have sufficient observation periods and regular streamflow data. Some of the remaining stations have been closed, and some have become unable to perform their duties due to urbanization. Therefore, only 22 stations that were not intervened could be selected.

3.3. At-site flood frequency analysis

The at-site frequency analysis was a crucial part of our study, as it was performed for flood data sets at all stations for annual maximum and partial duration series. The analysis revealed that the partial duration flood series fits all distributions. In contrast, the data from D03A013 and D03A085 stations do not fit the logarithmic gamma distribution for the annual maximum flood series. The most appropriate frequency distributions and relevant Kolmogorov-Smirnov test values for annual maximum and

partial duration flood series are shown in Table 13. According to Table 13, GPA was the most dominant distribution for both data sets among the at-site frequency distributions. This situation has also been seen in partial-duration flood series with very high suitability and other distributions adapted to less data and in the annual maximum flood series, GLO, FRE3, LP3, and LLO3 distributions provided the best fit to the data, while other distributions adapted to less data. Using the most appropriate distributions for both data sets, the at-site design flood discharges for certain return periods were obtained with the L-moment parameter estimation method and are given for comparison in Figure 7. According to the at-site frequency analysis results, in the lower tail of the distribution (short-term return periods), at D03A008, D03A013, D03A024, D03A044, D03A051, D03A052, D03A056, D03A064, D03A082, D03A084, D03A086, D03A087, D03A096, and E03A028 stations, both data sets produced estimates of design flood discharges close to each other. In the main body of the distribution (medium-term return periods), design flood discharges of both data sets produced close estimates to each other at stations D03A013, D03A024, D03A044, D03A056, E03A002, E03A011, E03A016, E03A024, E03A028 and E03A031. In the upper tail of the distribution (long-term return periods), design flood discharge estimates of both data sets were close to each other at stations D03A086, E03A002, E03A011, E03A016, E03A017, and E03A031. The stations whose estimations are close to each other in both short-term and medium-term return periods are D03A013, D03A024, D03A044, D03A056, E03A024, and E03A028. The stations whose estimations are close to each other in both medium-term and long-term return periods are E03A011, E03A016, and E03A031. Station data whose forecasts for all three periods were close to each other could not be determined. As can be understood from the expressions, using partial duration flood series as an alternative to annual maximum flood series in at-site frequency analysis is more appropriate, especially for the short and medium term. In long-term estimations, partial duration series at fewer stations can be an alternative to the annual maximum flood series (Kiran and Srinivas 2021).

Table 13- The best-fit frequency distributions and relevant Kolmogorov-Smirnov test values for both data series

		Annual Maximum Series		*Partial Duration Series*	
N	***Station Code***	***K-S***	***Best-fit Frequency Distribution***	***K-S***	***Best-fit Frequency Distribution***
1	D03A008	0.093	GG4	0.054	GPA
2	D03A013	0.107	LP3	0.079	FRE3
3	D03A024	0.065	LP3	0.067	GPA
4	D03A034	0.062	WE3	0.069	GPA
5	D03A044	0.070	LP3	0.285	LN3
6	D03A051	0.128	LLO3	0.063	GPA
7	D03A052	0.077	LN	0.055	GPA
8	D03A056	0.079	FRE3	0.090	G3
9	D03A064	0.073	FRE3	0.606	WE3
10	D03A082	0.109	LLO3	0.088	GPA
11	D03A084	0.070	GPA	0.059	FRE3
12	D03A085	0.110	WE	0.100	GPA
13	D03A086	0.073	GG4	0.062	LN
14	D03A087	0.081	G3	0.075	GPA
15	D03A096	0.093	GLO	0.053	GPA
16	E03A002	0.106	WE	0.063	G3
17	E03A011	0.101	GLO	0.069	GPA
18	E03A016	0.806	LLO3	0.058	GG4
19	E03A017	0.061	GLO	0.083	GPA
20	E03A024	0.086	G	0.081	LLO3
21	E03A028	0.085	FRE3	0.068	GG4
22	E03A031	0.142	GLO	0.051	FRE3

*K-S: Kolmogorov-Smirnov test value, EXP: Exponential, EXP2: 2-parameter exponential, FRE: Frechet, FRE3: 3-parameter Frechet, G: Gamma, G3: 3-parameter gamma, GEVs: Generalized extreme values, GG: Generalized gamma, GG4: 4-parameter generalized gamma, GLO: Generalized logistic, GPA: Generalized Pareto, LLO: Logarithmic logistic, LLO3: 3-parameter logarithmic logistic, LP3: 3-parameter logarithmic Pearson, LO: Logistic, LN: Logarithmic normal, LN3: 3-parameter logarithmic normal, N: Normal, WE: Weibull, WE3: 3-parameter Weibull.

Figure 7- Comparison of the at-site design flood discharges at certain return periods for both series

4. Conclusions

This study investigated the possibilities of using partial duration flood series as an alternative to annual instantaneous maximum flood series on an at-site and regional basis using the regional index-flood approach in the Susurluk river basin. The annual maximum flood series provided homogeneity in the Susurluk basin as a single region, and regionally designed flood discharges were estimated for hydraulic structures and flood offset in short, medium- and long-term return periods according to the one suitable GLO distribution. According to annual maximum flood data, since the basin is homogeneous, it has been revealed that design flood discharges can be considered for the entire Susurluk basin. The variance-mean ratio method was used while selecting partial duration flood series, and various frequency factors were used to extract peaks over the threshold. The most appropriate frequency factors for the data sets were determined according to the Poisson distribution, which makes the variance-mean ratio equal. k=3.5 was determined for only two stations. k=4 was suitable for seven stations, and k=5 was suitable for thirteen stations. For the k=5 value chosen at station E03A017, the total m (number of partial duration flood) value was smaller than the n (number of annual maximum flood) ($m < n$), so a frequency factor of k=4.5 was considered. The average number of peaks over the threshold level (λ`) varies between 1.26 and 5.31. The lowest flood data ($\lambda$`=1.26, m=48) was extracted at station E03A017, and the highest (λ`=5.31, m=154) was extracted at station D03A085. The basin was divided into three sub-regions by cluster analysis according to partial duration series. After ensuring homogeneity in three separate regions, regionally designed flood discharges were estimated for hydraulic structures and flood offset in short- medium- and long-term return periods according to PE3 and GPA distributions. In a regional comparison of data sets, it can be considered that PDS-I and PDS-III data sets can be used as an alternative to annual maximum floods for the lower tail of the distribution, and PDS-II data sets can be used for the upper tail of the distribution. The PDS-II data set for the main body of the distribution, except for the upper tail of the distribution, is not recommended from a cost perspective as it gives higher estimates than the AMS data set. In at-site frequency analysis, using partial duration flood series as an alternative to annual maximum flood series is more appropriate, especially in the short and medium term. In long-term estimations, a partial duration series at fewer stations can be an alternative to the annual maximum flood series. The study concluded that partial duration flood series can be used in many stations in the short and medium term instead of annual maximum flood series but can be used in fewer stations in the long term. Since partial duration flood estimations are lower than annual maximum ones, they can provide advantages to engineering projects with lower costs (Yue et al. 2022).

At the end of the study, some important points where partial duration series are advantageous in flood frequency analysis are listed below:

- Due to the extracting process (peaks-over-threshold), the partial duration series is not limited to smaller sample sizes than the annual maximum series, as the overall data length is flexible.
- Due to the width of the data set, partial duration series are effective in estimating frequent floods needed by the ecosystem.
- Partial duration series can be useful in regionalizing floods, which are common due to data extraction.
- Due to the controllability of the time series, partial duration series are more suitable for performing non-stationary flood frequency analysis.

Acknowledgments

This study was produced from the MSc thesis of Ayşe Doğanülker under the supervision of Alper Serdar Anli.

References

Adamowski K (2000). Regional analysis of annual maximum and partial duration flood data by nonparametric and L-moment methods. *Journal of Hydrology* 229(3): 219-231 doi:10.1016/s0022-1694(00)00156-6

Adamowski K, Liang G & Patry G G (1998). Annual maxima and partial duration flood series analysis by parametric and non-parametric methods. Hydrological Processes 12: 1685-99

Agilan V, Umamahesh N V & Mujumdar P P (2020). Influence of threshold selection in modeling peaks over threshold-based nonstationary extreme rainfall series. *Journal of Hydrology* 593: 125625

Ahmad I, Laksacia A, Chikr-Elmezouar Z, Almanjahie Ibrahim M & Khan DA (2019). At-site rainfall frequency analysis of partial duration series and annual maximum series using L-moments in Rawalpindi city of Pakistan. Applied Ecology and Environmental Research 17(4): 8351-8367. http://dx.doi.org/10.15666/aeer/1704_83518367.

Amorim R & Villarini G (2024). Assessing the performance of parametric and non-parametric tests for trend detection in partial duration time series. *Journal of Flood Risk Management* 17(1): e12957

Anlı A S (2009). Regional frequency analysis of rainfall data in Ankara province using L-moment methods. Ankara University, PhD thesis (In Turkish).

Anli A S, Apaydin H & Ozturk F (2007). Regional flood frequency estimation for the Göksu river basin through L-moments. In International River Basin Management Conference, State Hydraulic Works (pp. 22-24).

Ashkar F & Ba I (2017) Selection between the generalized Pareto and kappa distributions in peaks-over-threshold hydrological frequency modelling. *Hydrological Sciences Journal* 62(7): 1167–1180.

Basu B & Srinivas V V (2016). Evaluation of the index-flood approach related regional frequency analysis procedures. *Journal of Hydrologic Engineering* 21(1): 04015052. https://doi.org/10.1061/(asce)he.1943-5584.0001264

Ben-Zvi A (1991). Observed advantage for negative binomial over Poisson distribution in partial duration series. Stochastic Hydrology and Hydraulics 5(2): 135–146

Bezak N, Brilly M & Šraj M (2014). Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis. *Hydrological Sciences Journal* 59(5): 959–977

Cunnane C (1979). A note on the Poisson assumption in partial duration series models. Water Resources Research 15: 489-94

Dalrymple T (1960). Flood frequency analyses. Water Supply Paper 1543-A, U.S. Geological Survey, Reston, Va.

Doğanülker A (2022). Flood Frequency Analysis in Susurluk River Basin. Ankara University, MSc. thesis (In Turkish).

Durocher M, Burn D H & Ashkar F (2019). Comparison of estimation methods for a nonstationary index-flood model in flood frequency analysis using peaks over threshold. Water Resources Research 55(11): 9398–9416

Fill H D & Steiner A A (2003). Estimating instantaneous peak flow from mean daily flow data. *Journal of Hydrologic Engineering*, Vol. 8, No. 6, November 1, 203: 365-369

Greenwood J A, Landwehr J M & Matalas N C (1979). Probability weighted moments: Definition and relation of parameters of several distributions expressible in inverse form. Water Resources Research 15: 1049-1054

Guru N (2016). Flood frequency analysis of partial duration series using soft computing techniques for Mahanadi River basin in India. National Institute of Technology Rourkela-769008. Doctor of Philosophy

Guru N (2022). Implication of partial duration series on regional flood frequency analysis, *International Journal of River Basin Management*, 22:2, 167-186, DOI: 10.1080/15715124.2022.2114486.

Gürler Ç, Anli AS & Polat H E (2024). Developing regional hydrological drought risk models through ordinary and principal component regression using low-flow indexes in Susurluk basin, Turkey. Water 16, no. 11: 1473. https://doi.org/10.3390/w16111473.

Haktanır T (1991). Practical computation of gamma frequency factors. Hydrological Sciences 36,6: 559-610.

Haktanır T & Bozduman A (1995). A study on sensitivity of the probability weighted method on the choice of the plotting position formula. *Journal of Hydrology* 168: 265-281

Hosking J R (1986). Estimation of the generalized extreme value distribution by the method of probability-weighted moments. Technometrics 27: 251-261

Hosking J R M (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society*. Series B 52(1): 105-124

Hosking J R M & Wallis J R (1993). Some statistics useful in regional flood frequency analysis. Water Resources Research 23: 271-281.

Hosking J R M & Wallis J R (1997). Regional frequency analysis an approach based on L-moments, Cambridge University Press.

Hosking J R M (2005). Fortran routines for use with the method of L-moments, Version 3.04. Research Report RC 20525, IBM Research Division, T.C. Watson Research Center, Yorktown Heights, N.Y.

Karim F, Masud H & Marvanek S (2017). Evaluating annual maximum and partial duration series for estimating frequency of small magnitude floods, Water 9, no. 7: 481. https://doi.org/10.3390/w9070481.

Kiran K G & Srinivas VV (2021). Distributional regression forests approach to regional frequency analysis with partial duration series. Water Resources Research, 57, e2021WR029909. https://doi.org/10.1029/2021WR029909.

Kjeldsen T R, Smithers J C & Schulze R E (2002). Regional flood frequency analysis in the KwaZulu-Natal province, South Africa, using the index-flood method. *Journal of Hydrology* 255: 194-211

Lang M, Ouarda T B M J & Bobée B (1999). Towards operational guidelines for over-threshold modeling, *Journal of Hydrology* 225(3-4): 103- 117

Madsen H, Rasmussen P F & Rosbjerg D (1997a). Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events. 1. At-site modeling. Water Resources Research, 33: 747-57

Madsen H, Pearson C P & Rosbjerg D (1997b). Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events. 2. Regional modeling. Water Resources Research 33: 759-769

Öney M & Anlı A (2023). Regional drought analysis with standardized precipitation evapotranspiration index (SPEI): Gediz basin, Turkey. *Journal of Agricultural Sciences* 29(4): 1032-1049. https://doi.org/10.15832/ankutbd.1030782.

Önöz B (1994). A new parameter estimation method, probability weighted moment method. DSI Technical Bulletin, 81: 49-54 (In Turkish).

Pan X & Rahman A (2021). Comparison of annual maximum and peaks-over-threshold methods with automated threshold selection in flood frequency analysis: a case study for Australia. Natural Haz. https://doi.org/10.1007/s11069-021-05092-y.

Pan X, Rahman A, Haddad K & Ouarda T B (2022). Peaks-over-threshold model in flood frequency analysis: a scoping review. Stochastic Environmental Research and Risk Assessment 36(9): 2419-2435

Pham H X, Shamseldin A Y & Melville B (2014). Statistical properties of partial duration series: Case study of North Island, New Zealand. *Journal of Hydrologic Engineering* 19(4): 807-815

Rahman A S, Rahman A, Zaman M A, Haddad K, Ahsan A & Imteaz M (2013). A study on selection of probability distributions for at-site flood frequency analysis in Australia. Natural Hazards 69(3): 1803-1813

Rasmussen P F & Rosbjerg D (1991). Prediction uncertainty in seasonal partial duration series. Water Resources Research 27: 2875-83.

Rosbjerg D & Madsen H (1992). On the choice of threshold level in partial duration series, Proc. Nordic Hydrological Conference, Alta (ed. G. Østrem), NHP Rep. no. 30: 604-615

Rosbjerg D, Madsen H & Rasmussen P F (1992). Prediction in partial duration series with generalized Pareto-distributed exceedances. Water Resources Research 28: 3001-3010

Seçkin N (2009). Regional flood frequency analysis with index-flood method based on L-moments, Çukurova University, PhD. Thesis, Adana.

Seckin N & Topcu E (2016). Regional frequency analysis of annual peak rainfall of adana and the vicinity. Journal of the Faculty of Engineering and Architecture of Gazi University 31(4): 1049-1062

Stedinger J R (1983). Estimating a regional flood frequency distribution. Water Resources Research 19(2): 503-510

Sveinsson O G B, Salas J D & Boes D C (2003). Uncertainty of quantile estimators using the population index flood method. Water Resources Research, 39(8): 1206. https://doi.org/10.1029/2002wr001594.

Swetapadma S & Ojha C (2021). Technical Note: Flood frequency study using partial duration series coupled with entropy principle. Hydrology and Earth System Sciences. https://doi.org/10.5194/hess-2021-570

SYGM (2022). Susurluk basin flood management planning report, General Directorate of Water Management, Ankara (In Turkish).

Topcu E & Seckin N (2016). Drought analysis of the Seyhan Basin by using standardized precipitation index SPI and L-moments. *Journal of Agricultural Sciences* 22(2): 196-215

Van Campenhout J, Houbrechts G, Peeters A & Petit F (2020). Return period of characteristic discharges from the comparison between partial duration and annual series: Application to the Walloon Rivers (Belgium). Water 12: 792. https://doi.org/10.3390/w12030792.

Wilks D S (1993). Comparison of three-parameter probability distributions for representing annual extreme and partial duration precipitation series. Water Resources Research 29: 3543-49.

Yue Z, Xiong L, Zha X, Liu C, Chen J & Liu D (2022). Impact of thresholds on nonstationary frequency analyses of peak over threshold extreme rainfall series in Pearl River Basin, China. Atmospheric Research 276: 106269

Zadeh S, Durocher M, Burn & Ashkar F (2019). Pooled flood frequency analysis: a comparison based on peaks-over threshold and annual maximum series, *Hydrological Sciences Journal* 64(2): 121-136, DOI:10.1080/02626667.2019.1577556.

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

A New Innovative Approach with Revised Pythagorean Fuzzy SWARA in Assessing of Soil Erodibility Factor

Aykut Çağlar[a], Barış Özkan[b], Orhan Dengiz[c*]

[a]*Republic of Türkiye Ministry of Agriculture and Forestry, Black Sea Agricultural Research Institute, Samsun, TÜRKIYE*
[b]*Department of Industrial Engineering, Faculty of Engineering, Ondokuz Mayıs University, Samsun, TÜRKIYE*
[c]*Department of Soil Science and Plant Nutrition, Faculty of Agriculture, Ondokuz Mayıs University, Samsun, TÜRKIYE*

ARTICLE INFO
Research Article
Corresponding Author: Orhan Dengiz, E-mail: odengiz@omu.edu.tr
Received: 15 June 2024 / Revised: 28 August 2024 / Accepted: 10 September 2024 / Online: 14 January 2025

Cite this article
ÇAĞLAR A, ÖZKAN B, DENGIZ O (2025). A New Innovative Approach with Revised Pythagorean Fuzzy SWARA in Assessing of Soil Erodibility Factor. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):182-195. DOI: 10.15832/Ankutbd.1501907

ABSTRACT
Soil erosion is a significant issue that threatens to soil in land degradation processes. The soil erodibility factor is a crucial tool for assessing the susceptibility of soils to erosion. The main aim of this study was to compare the results obtained using the Pythagorean Fuzzy-SWARA method which evaluates the impact weights of the criteria considered for the soil erodibility factor of the soils in the micro-basins located in the district of Çarşamba district of Samsun province, with the results obtained using the formula developed by Wischmeier and Smith. To achieve this case, 78 surface soil samples were collected from micro basins and analyzed for organic matter, clay, sand, silt, very fine sand, degree of structure, and hydraulic conductivity parameters. The erodibility factor was then calculated using these data, and spatial distribution maps were created for both methods. In this study, a revised of the Pythagorean Fuzzy-SWARA approach is proposed to calculate the weight values of the criteria. The values were 0.418 for organic matter, 0.227 for clay, 0.120 for degree of structure, 0.100 for hydraulic conductivity, 0.058 for sand, 0.053 for silt, and 0.039 for very fine sand. Soil erodibility values were determined using a linear combination approach, which normalized all parameter values by a standard scoring function. In estimating soil erodibility, our revised Pythagorean Fuzzy-SWARA approach was found to have a significant relationship with the soil erodibility factor method (R^2 = 0.691 at the 1% level) compared to the soil erodibility factor method in estimating soil erodibility. Consequently, the method developed here suggests that fuzzy multi-criteria decision-making methods can be an alternative approach for determining the soil erodibility factor.

Keywords: Soil erodibility, Pythagorean Fuzzy Sets, SWARA, RUSLE-K

1. Introduction

Soil erosion can negatively affect the sustainability of soils in terms of environmental processes, water retention capacities and crop yields, and it is a very critical process in terms of land degradation (Pimentel & Burgess 2013). The intensification of agricultural activities and increasing population can lead to a decrease in soil production capacity and an increase in soil erosion susceptibility rates (Yang et al. 2003; Dotterweich 2013). Soil erosion is influenced by physical factors such as soil sensitivity and properties that affect the separation of soil particles, as well as infiltration, permeability, and water holding capacity (Wischmeier & Smith 1965). The realization of erosion processes due to these properties depends on various soil and environmental factors, including soil organic matter level, texture, slope, and rainfall (Dede et al. 2022).

The Universal Soil Loss Equation (USLE) (Wischmeier & Smith 1978) and the Revised Universal Soil Loss Equation (RUSLE) (Renard et al. 1997) are commonly used to estimate soil erosion worldwide. The erodibility of soils in both equations is determined by the K factor, which is strongly influenced by the physical, hydrological, chemical, mineralogical, and biological properties of soils (Perez-Rodriguez et al. 2007). The consistency of the calculating the soil erodibility factor (RUSLE-K) factor by GIS and geostatistical analyses is based on its potential to represent characteristic variations of soil properties and processes (Saygın et al. 2011). Many studies have used the RUSLE-K factor to the calculate of soil erodibility (Gao et al. 2023; Başkan & Dengiz 2008; Başkan 2022; Beretta & Carrasco-Letelier 2017; Parlak et al. 2014).

Multi-criteria decision-making (MCDM) methods have become increasingly popular in the fields of spatial planning and management. They are now a critical tool for decision-makers, particularly in multi-factor evaluations (Valkanou et al. 2021). Multi-criteria decision-making methods have a wide range of applications and have been used in many studies. They have been used for prioritizing erosion risk areas (Zhang et al. 2020), predicting erosion risk (Demirağ Turan & Dengiz 2017), mapping potential soil erosion (Cartwright et al. 2022), and conducting land suitability studies (Mercan 2023). Zadeh (1965) introduced

the concept of fuzzy sets to better express uncertainty, inadequacy, and complexity, which are significant issues in decision-making methods. Yager (2013) continued the development process with Pythagorean Fuzzy Sets (PFSs). When defining PFSs, it was determined that the sum of the squares of the degrees of membership and non-membership of the set should not exceed 1. This structure enables decision-makers to make evaluations from a broader perspective. The SWARA method, developed by Kerŝulienė et al. (2010), enables the determination of criteria weights by considering the priority rankings determined by decision-makers. Pairwise comparisons are made between the criteria by taking into account expert opinions. Several studies in the literature integrate PFSs to SWARA method. Rani et al. (2020) used the Pythagorean Fuzzy-SWARA (PF-SWARA) and VIKOR integration for performance evaluation of solar panel selection, Kamali Saraji et al. (2022) used PF-SWARA and TOPSIS integration in the evaluation of sustainable energy development processes of European Union countries, Rani et al. (2020) used PF-SWARA ARAS method to evaluate healthcare waste treatment, Saeidi et al. (2022) used PF-SWARA and TOPSIS integration in the evaluation of sustainable human resources management.

In the above-mentioned studies, experts were asked to rate a set of pre-defined criteria using a scale to determine the weighting in the PF-SWARA approaches. The resulting score function values were ranked from highest to lowest and then used in the SWARA method by subtracting the criteria from each other, rather than through pairwise comparisons. However, it is important to note that the SWARA method ranks criteria based on their importance degrees within the scope of expert opinions, without making pairwise comparisons for each criterion. This may lead to different interpretations of the results obtained. To prevent such interpretation differences, this study proposes a new innovative approach to the PF-SWARA method.

In this context, the present study uses the revised PF-SWARA approach to calculate the soil erodibility of the micro-basins located in Çarşamba district of Samsun province This approach evaluates the formula and criteria developed by Wischmeier and Smith through pairwise comparisons. The current study also aims to create spatial distribution maps and compare the obtained results. Furthermore, the revised PF-SWARA approach was applied for the first time to the soil erodibility factor, thus it has the potential to make a significant contribution to this field of study and the existing literature.

2. Material and Methods

The study area is situated in the eastern part of Samsun Province of Türkiye and the western part of Ordu Province, covering an area of 1037.7 km^2. It is located between the east coordinates of 296 000-316 000 and the north coordinates of 4 549 000-4 577 000 (Universal Transverse Mercator-UTM, WGS-84, Zone-37), with an elevation ranging from 0 to 416 m above sea level (Figure 1). The selected study area lies south of the town Samsun-Çarşamba district on the edge of the Çarşamba delta plain, including both a part of the lowland and hills and extends around the accumulation lake on the Yeşilırmak River that passes through the middle of our study.

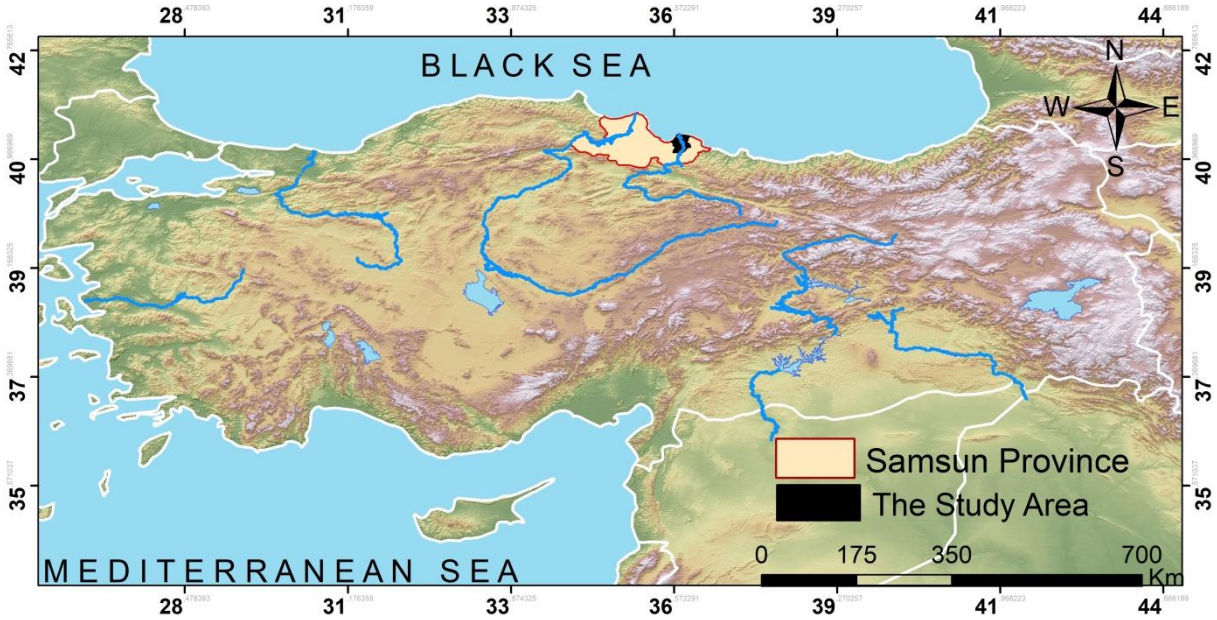
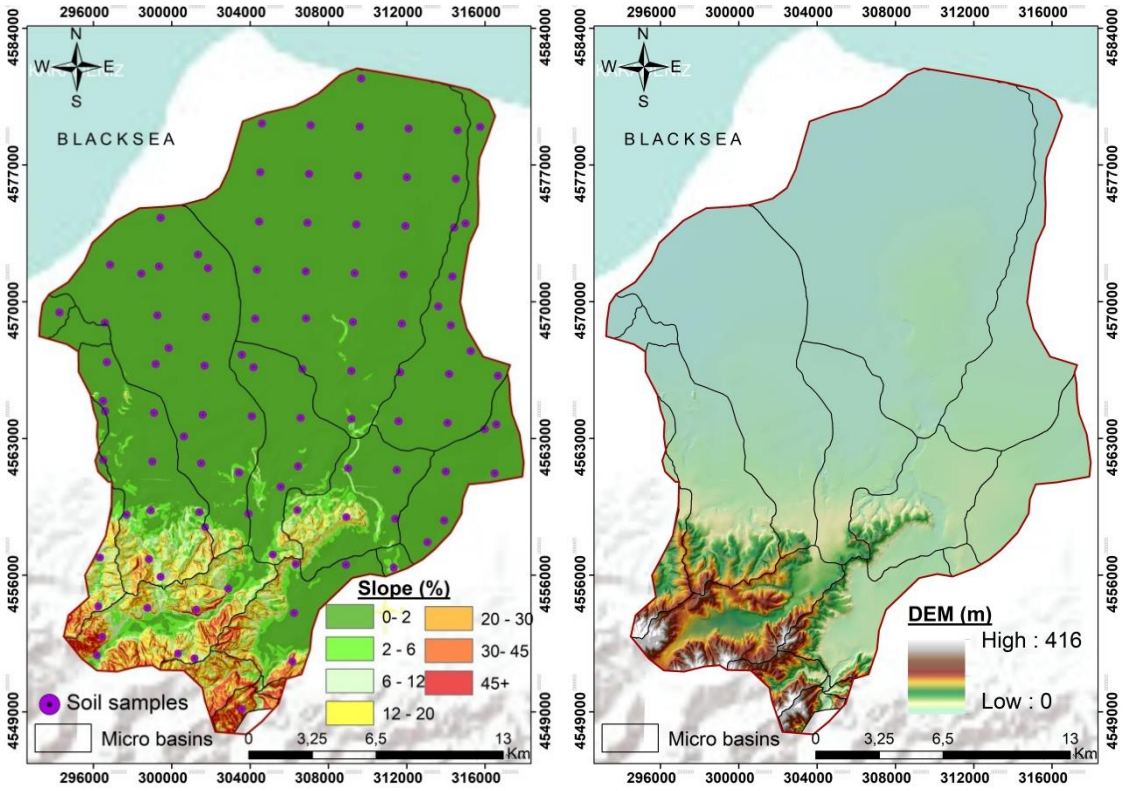


Figure 1- Location map of the study area

According to meteorological data collected by the Turkish State Meteorological Service (DMI) between the years 1960 and 2017, the climate in Çarşamba district is semi-humid with average annual precipitation and evapotranspiration levels being 1023 mm and 772 mm. The average annual temperature is 14.4 °C, average temperature in July is 23.3 °C and in January 6,4 °C. As calculated in the study by Miháliková & Dengiz (2019), the Newhall simulation model classifies the soil climate of the study area

as having a mesic soil temperature regime and udic (subgroup: dry tempudic) moisture regime. Besides, Bölük (2016) reports that the study area is classified as "very moist," with a precipitation activity index of 66.64 scores based on the macroclimate regions of Erinc in Türkiye.

The general slopes of the study area's the flat and nearly flat lands of the study area are in the south-north direction and average 0.1%. This slope decreases to 0.0 - 0.02% as it approaches the seaside. On the slope lands towards the south, the slope varies between 2 - 40% (Figure 2). Although the climatic conditions of the plain are suitable for cultivating many crops, the high level of groundwater, surface drainage needs, lack of irrigation water, leveling disorders, and consolidation needs have an adverse effect on the crop pattern and yield.



,

Figure 2- Slope, elevation and soil sampling pattern maps

2.1. Soil sampling and analysis

Within the scope of the study, 78 soil samples were collected from micro basins located within the borders of Çarşamba district of Samsun province. The samples were taken from 0-30 depth cm in an area with a semi-humid ecosystem (refer to Figure 2). This study used Bouyoucos (1962) to determine the textural content of the soils, the Walkley-Blake method (Jackson 1958) to determine the organic matter content, Klute & Dirksen's (1986) method to determine saturated hydraulic conductivity, and the wet sieving method under laboratory conditions to determine the very fine sand content. To perform Very fine sand analysis, 10 g of soil must first be crushed with chemical (calgon) and mechanical (mixer). It must then be sieved through a 0.105 mm, after which the percentage of coarse sand can be determined from the remaining dry particles. The percentage of fine sand is calculated by subtracting the percentage of coarse sand from the total percentage of sand obtained by texture analysis (Wischmeier & Smith 1978). The RUSLE-K factor was then determined, and the soil structure classes were determined according to Wischmeier & Smith (1978).

2.2. Soil erodibility (RUSLE-K)

The following formula developed by Wischmeier & Smith (1978) is used to determine the susceptibility of soils to erosion.

$$100 * K = ((2.1 \times 10^{-4}) \times (12 - OM) \times M^{1.14} + 3.25 \times (S - 2) + 2.5 \times (P - 3))/d \quad (1)$$

In the equation; K: Soil erodibility factor, M: Particle size parameter, OM: Organic matter content, %, S: Structure type code, P: Hydraulic conductivity, d: Conversion coefficient to metric system (d=7.59)

The following equation was used to determine the particle size (M) parameter in the equation.

$$M = (\%\, Very\ fine\ sand + \%\, Silt) \times (100 - \%\, Clay) \quad (2)$$

The RUSLE-K equation used in this study is valid only when the organic matter is less than 4%. For this reason, instead of excluding the data of 12 soil samples in our study, it was used by limiting it to 4% (Efthimiou 2018).

2.2. Pythagorean fuzzy set theory preliminary definition

PFSs were first introduced by Yager in 2013 and were developed as an extension of Intuitionistic Fuzzy Sets (IFSs) to offer decision-makers a wider scope in defining uncertainties (Yager 2013). When determining IFSs, the sum of the membership and non-membership degrees cannot be greater than 1. Still, when when defining uncertainty in PFSs, the sum of the squares of the membership and non-membership degrees is determined to not exceed 1. This condition enables PFSs to cover a much wider area in defining uncertainties (Fgure 3).

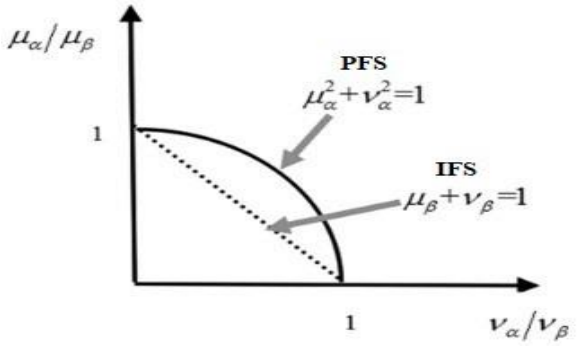


Figure 3- Formal expression of PFSs and IFSs (Peng and Yang, 2016)

The following equations are used to define PFSs (Yager 2013).

$$P = \{x, \mu_p(x), \nu_p(x);\ x \in X\} \quad (3)$$

Where, $\mu_p: X \to [0,1]$ defined with the degree of membership, $\nu_p: X \to [0,1]$ defined with the degree of non-membership. In here, $\mu_p(x)$, $\nu_p(x) \in [0,1]$ and is described as

$$0 \le (\mu_p(x))^2 + (\nu_p(x))^2 \le 1 \quad (4)$$

Eq. 5 is used to determine the hesitation degrees of PFSs.

$$\pi_p(x) = \sqrt{1 - (\mu_p(x))^2 + (\nu_p(x))^2} \quad (5)$$

Where, πp is hesitation degrees.

Arithmetic operations for two PFSs are stated in the equations below. It is assumed that $\beta_1 = P(\mu_{\beta 1}, \nu_{\beta 1})$ and $\beta_2 = P(\mu_{\beta 2}, \nu_{\beta 2})$ two PFSs.

$$\beta_1 \oplus \beta_2 = P\left(\sqrt{\mu_{\beta 1}^2 + \mu_{\beta 2}^2 - \mu_{\beta 1}^2 \mu_{\beta 2}^2}, \nu_{\beta 1} \nu_{\beta 2}\right) \quad (6)$$

$$\beta_1 \otimes \beta_2 = P\left(\mu_{\beta 1} \mu_{\beta 2}, \sqrt{\nu_{\beta 1}^2 + \nu_{\beta 2}^2 - \nu_{\beta 1}^2 \nu_{\beta 2}^2}\right) \quad (7)$$

$$\lambda\beta = P\left(\sqrt{1 - (1 - \mu_\beta^2)^\lambda}, (\nu_\beta)^\lambda\right) \quad (8)$$

$$\beta^\lambda = P\left((\mu_\beta)^\lambda, \sqrt{1 - (1 - \nu_\beta^2)^\lambda}\right) \quad (9)$$

The equations used to determine the score and uncertainty functions of PFSs are (Zhang & Xu 2014);

$$S(\beta) = (\mu_\beta)^2 - (v_\beta)^2, \hbar(\beta) = (\mu_\beta)^2 + (v_\beta)^2, \text{ here} S(\beta) \in [-1,1] \text{ and } \hbar(\eta) \in [0,1] \quad (10)$$

Where; S(β): the score function and $\hbar(\eta)$: uncertainty function.

Where; since the score (S) function is between [-1, 1], equation 11 is used when calculating the normalized form of the values and the degree of uncertainty (Wu & Wei 2017).

$$S*(\beta) = \frac{1}{2}(S(\beta) + 1), \hbar^o(\beta) = 1 - \hbar(\beta), \text{ so that } S*(\beta), \hbar^o(\beta) \in [0,1]. \quad (11)$$

Where; S*(β): normalized score function, $\hbar^o(\beta)$: degree of uncertainty.

2.3. The why proposed pythagorean fuzzy SWARA approach

It is well known that MCDM methods are an important tool that takes into account the opinions of decision makers in solving any complex problem in last decades. It can be also mentioned that SWARA method is used by Kerŝulienė et al. (2010) as a method in which we can obtain weight by ranking the criteria according to their importance in solving critical problems. The method allows decision-makers to choose their own priorities by taking into account the conditions in terms of the problem (Zavadskas et al. 2019). On the other hand, the area of use of the SWARA method has been expanded by using it in many fuzzy set studies used to identify uncertainties. In the PF-SWARA method, which is widely used in the literature, each criterion in the problem is evaluated by taking expert opinions with PFSs and the criteria are ranked from the highest to the lowest by calculating the score values. The score values obtained after the criteria ranking are not subjected to the pairwise comparison process, which is a stage of the classic SWARA method, and instead, the score values are subtracted from the score value below and the resulting value is processed in the SWARA method (Rani et al. 2020; Saeidi et al. 2022; Kamali Saraji et al. 2022). Here, although there is a score value initially obtained with PFSs, since each criterion is evaluated within each criterion instead of pairwise comparison, it causes the desired results for the problem not to be obtained in the criterion weights obtained with the SWARA method afterward.

This study was used the revised PF-SWARA method to determine the criterion weights. With the method we propose within the scope of our study, the criteria are ranked according to their importance, taking into account expert opinions. Scale evaluation is then performed using PFSs clusters. Here, pairwise comparisons in the evaluation of criteria are made using PFSs. The score function obtained by making pairwise comparisons is then used to calculate the criterion weights. Although it is similar to the traditional SWARA method (Kerŝulienė et al. 2010) in terms of method, the PFSs approach used in the evaluation of the criteria distinguishes the study from other methods (Rani et al. 2022; Kamali Saraji et al. 2022; Saeidi et al. 2022) and offers a new approach to the literature. The stages of the revised PF-SWARA method are stated below.

Step 1: First of all, experts are evaluated with the help of the scale in Table 1, taking into account their work history and expertise on the subject.

Table 1- Scale for evaluating the performance of experts (Cui et al. 2021)

Linguistic expressions	*PFNs* (μ, v, π)
Extremely important (EI)	(0.90, 0.15, 0.409)
Very very important (VVI)	(0.75, 0.40, 0.527)
Very important (VI)	(0.70, 0.55, 0.456)
Important (I)	(0.60, 0.70, 0.387)
Less important (LI)	(0.40, 0.80, 0.447)
Very less important (VLI)	(0.30, 0.90, 0.316)

μ: membership degree, v: non-membership degree, π: hesitation degree

Step 2: The following equation (Eq. 12) is used to calculate the weights of experts. Calculating expert weights is very important for MCDM processes. Here, let $E_k=\gamma(\mu_k, v_k)$ be in the PFN of k experts, while calculating the weight of this expert;

$$\omega_k = \frac{(\mu_k^2+\pi_k^2\times(\frac{\mu_k^2}{\mu_k^2+v_k^2}))}{\sum_{k=1}^{\ell}(\mu_k^2+\pi_k^2\times(\frac{\mu_k^2}{\mu_k^2+v_k^2}))}, k = 1(1)\ell; \omega_k \geq 0, \sum_{k=1}^{\ell}\omega_k = 1 \quad (12)$$

Where; k represents the decision expert, w: the weight of k^{th}, μ: membership degree, *v*: non-membership degree, π: hesitation degree and ℓ: total number of the experts.

Step 3: The criteria are ranked from highest to lowest according to their importance, taking into account expert opinions before pairwise comparisons.

Step 4: At this stage, normalized score function values were obtained by making a pairwise comparison of the criteria determined according to their importance level with the criterion in the lower row. The linguistic expressions used by experts to evaluate the criteria within the scope of the study are listed in Table 2.

Table 2- Linguistic expressions used for expert opinions (Saeidi et al. 2022)

Linguistic expressions	*PFNs* (μ, v, π)
Absolutely high (AH)	(0.95, 0.20, 0.387)
Very very high (VVH)	(0.85, 0.30, 0.433)
Very high (VH)	(0.80, 0.35, 0.487)
High (H)	(0.70, 0.45, 0.554)
Medium high (MH)	(0.60, 0.55, 0.581)
Medium (M)	(0.50, 0.60, 0.624)
Medium low (ML)	(0.40, 0.70, 0.592)
Low (L)	(0.30, 0.75, 0.589)
Very low (VL)	(0.20, 0.85, 0.487)
Absolutely low (AL)	(0.10, 0.95, 0.296)

μ: membership degree, *v*: non-membership degree, π: hesitation degree

Step 5: At this stage, the most important criterion in calculating the k_j value is determined as 1, and +1 is added to each score value below it (Eq. 13) (Kerŝulienė et al. 2010).

$$k_j = \begin{cases} 1 & ,j = 1 \\ s_j + 1 & ,j > 1 \end{cases} \quad (13)$$

Where; kj is relative coefficient and sj presents the comparative significance of score value.

Step 6: In this step, the importance vector is calculated and, as stated in the previous equation, the most important criterion is 1, and the importance vectors for other criteria are determined by dividing the previous importance vector to the criterion coefficient (Eq. 14).

$$q_j = \begin{cases} 1 & j = 1 \\ \frac{q_{(j-1)}}{k_j} & ,j > 1 \end{cases} \quad (14)$$

Where; q_j is the weight of the importance vector.

Step 7: The final weights of the criteria are obtained by dividing the weight of the importance vectors obtained in the previous step by the sum of the importance vector weights as stated in Eq. 15. The m value here represents the number of criteria.

$$w_j = \frac{q_j}{\sum_{j=1}^{m} q_j} \quad (15)$$

Step 8: Within the scope of the study, the criterion weights in the study are calculated for each expert by following the above steps. In the next stage, the sum of the values obtained by multiplying the expert weights obtained within the scope of evaluating the experts and the criterion weights obtained from the experts is taken and the criterion weights are calculated.

2.4. Normalization process with a standard scoring function

Soil erodibility factor is determined by considering soil parameters such as organic matter content, texture, degree of structure, permeability, and very fine sand properties (Wischmeier & Smith 1978; Renard et al. 1997). All parameters used in the calculation of soil erodibility factor were calculated under laboratory conditions in this study. At this stage, the standard scoring function was used to normalize the values of soil parameters and assign scores ranging from 0 to 1 due to differences in soil parameter units (Andrews et al. 2002). The scoring function enables the considers of low, high, and average levels of the desired feature within the study's scope (Liebig et al. 2001). Two different scoring functions were used to ensure a linear correlation between the score values used in estimating soil erodibility and the soil erodibility obtained by the formula developed by Wischmeier & Smith (1978). The 'Less is better' (LB) function was utilized to determine the soil's organic matter content, clay, structural stability, and saturated hydraulic conductivity values. For the sand, silt and very fine sand properties of the soils, the "more is better-MB" function was used (Table 3).

Table 3- Standard scoring functions for K parameters (Eqs. 16-17)

Parameters	*SF*	*L*	*U*	*Standard Scoring Function Equation*
Organic matter	LB	0.50	4.00	$f(x) = \begin{cases} 0.1 & , x \le L \\ (1-(0.9)*(\frac{x-L}{U-L})), & L \le x \le U \\ 1 & , x \ge U \end{cases}$ (16)
Clay	LB	6.44	63.53	
Structure degree	LB	1.00	3.00	
Hydraulic conductivity	LB	0.07	7.31	
Sand	MB	7.33	77.93	$f(x) = \begin{cases} 0.1 & , x \le L \\ (0.1+(0.9)*(\frac{x-L}{U-L})), & L \le x \le U \\ 1 & , x \ge U \end{cases}$ (17)
Silt	MB	13.81	53.9	
Very fine sand	MB	5.05	24.69	

SF: Score Function, MB: more is better, LB: less is better. L and U are the lower and upper thresholds, respectively

After the determination of the weights of the relative importance levels of soil erodobility parameters by the revised PF-SWARA method and the normalized values obtained from the standard scoring function, the Weighted Linear Combination (WLC) method was used to determine the erodobility levels of the soils. WLC is also known as simple additive weighting (SAW), weighted summation, weighted linear average and weighted overlay (Malczewski & Rinner, 2015). The WLC method calculates the soil erosion susceptibility value using the formula in Eq. (18).

$$\text{PF SWARA} - \text{K} = \sum_{k=1}^{l} \omega_k * a_{ik} \quad (18)$$

Where; ω_k is the criteria weights obtained from the revised PF-SWARA method, a_{ik} is the normalized standard value of the parameter k obtained from the standard scoring function, and l is the total number of criteria.

2.5. Determination of spatial distributions

After the calculating the criterion weights and score values obtained for each parameter using the weighted linear combination method, interpolation methods were used to generate a map of the estimated soil erodibility. The values were calculated using the RUSLE-K method. In mapping the obtained coordinated point values, the inverse distance weighting (IDW) method, which uses a linear combination of weightings at known points to estimate the value at an unknown point, was considered. Teegavarapu & Chandramouli (2005) found that IDW predicts attribute values at unsampled points by adding the values at sampled points, weighted by an inverse function of the distance from the point of interest to the sampled points. The IDW approach assumes that the value of a point is more influenced by closer points than by those further away. Predictions were determined using the Eq. (19):

$$Z = [\sum_{i=1}^{n}(Z_i/d_i^m)/\sum_{i=1}^{n}(1/d_i^m)] \quad (19)$$

Where; Z is the estimated value, Z_i is the measured sample value at point i, d_i is the distance between Z and Z_i, and m is the weighting power that describes the ratio at which weights fall off with d_i. The study compared the IDW predictions using the common 1, 2, and 3 weighting powers (Pirmoradian et al. 2010; Keshavarzı and Sarmadian 2012).

Comparative statistics (Spearman's correlation) between parameters and descriptive statistics such as average, minimum, maximum, standard deviation, coefficient of variation, kurtosis and skewness obtained within the scope of the study with using IBM SPSS Statistics 23v. software (IBM Corp 2015).

3. Results and Discussion

3.1. Investigation of soil erodibility parameters and correlation relationship

In a total of 78 coordinated soil samples taken from the study area, the parameters used to determine the erodibility of the soils were calculated and the descriptive statistics of RUSLE-K and PF SWARA-K obtained by using these parameters are given in Table 4. The clay content of the soils varied between 6.4 % and 63.5 %, sand content between 7.3 % and 77.9 % and silt content between 13.8 % and 53.9 % and they were generally classified as medium textured. In addition, very fine sand contents, which are used to determine the susceptibility of soils to erosion, were found to vary between 5.1 % and 24.7 %.

The RUSLE-K values were calculated using the formula developed by Wischmeier & Smith (1978), ranging between 0.008 and 0.049. Meanwhile, the K factor was calculated using the revised PF-SWARA method, ranging between 0.260 and 0.874. The coefficients of variation of the parameters used in the study were evaluated according to Wilding's (1985) classification system. It was emphasized that coefficients of variation are low for values less than 15%, medium for values between 15% and 35%, and high for values above 35%. Upon analyzing the coefficients of variation of the soil samples, it was determined that the degree of structure had a low coefficient of variation at 12%. In comparison, the percentage of sand and hydraulic conductivity values had high variability at 55.9% and 156.8%, respectively. The soils' organic matter content ranged from 0.5% to 4%, with moderate variability (coefficient of variation of 34.4%).

Table 4- Descriptive statistics values of the parameters

Parameters	*Mean*	*SD*	*Variation*	*CV.*	*Min.*	*Max.*	*Skewness*	*Kurtosis*
RUSLE-K	0.030	0.008	0.000	27.43	0.008	0.049	-0.137	0.369
PF SWARA-K	0.468	0.133	0.018	28.42	0.258	0.871	-0.328	0.557
OM (%)	2.678	0.922	0.849	34.41	0.503	4.000	-0.836	-0.191
Clay (%)	41.084	11.225	126.011	27.32	6.438	63.534	-0.234	-0.202
Sand (%)	23.762	13.282	176.412	55.89	7.333	77.928	2.737	1.446
Silt (%)	35.154	7.908	62.537	22.49	13.807	53.903	0.163	-0.199
H.C (cm/h)	0.524	0.822	0.675	156.78	0.071	7.307	62.215	7.534
S.D (unitless)	2.885	0.360	0.129	12.46	1.000	3.000	11.068	-3.275
Vfs (%)	13.842	4.664	21.749	33.69	5.048	24.689	-0.516	0.316

SD: standard deviation, Min.: minimum, Max.: maximum, n: sample number (78), CV (coefficient of variation): <15 = low variation, 15–35 = moderate variation, >35 = high variation, Skewness:< | +- 0.5 | = normal distribution, 0.5–1.0 = application of character changing for dataset, and >1.0 → application of Logarithmic change, Vfs: Very fine sand, OM: Organic matter, H.C: Hydraulic conductivity, S.D: Structure degree

In the present study, Spearman's correlation analyses were performed for all parameters as well as pairwise comparisons of RUSLE-K obtained from real values and K factor obtained by PF-SWARA method (Table 5). The statistically significant value of 0.691** (P<0.01) between the soil erodibility factor calculated using the PF-SWARA method and the values calculated with the RUSLE-K method is an important result in terms of the success of the applied method. Here, instead of looking at the pairwise comparisons between all parameters, examining the importance of the criteria used in the calculating soil erodibility in terms of PF SWARA-K and RUSLE-K would be more accurate. For instance, both methods yielded significant results at the 1% level for very fine sand, clay, and hydraulic conductivity values, and the correlation values were similar. The study's most significant difference is that organic matter is -0.579** for RUSLE-K and -0.917** for PF SWARA-K. Therefore, organic matter is a more critical parameter for our newly developed method. Previous studies have shown that organic matter plays a crucial role in supporting aggregate structures in soils, which results in a reduction in soil erodibility (Dede et al. 2022; İmamoğlu & Dengiz 2017).

Table 5- Correlation values between model outputs and parameters

Parameters	*RUSLE-K*	*PF SWARA-K*	*Vfs*	*Clay*	*Sand*	*Silt*	*OM*	*H.C*	*S.D*
RUSLE-K	1								
PF SWARA-K	0.691**	1							
Vfs	0.477**	0.336**	1						
Clay	-0.562**	-0.659**	-0.457**	1					
Sand	0.250*	0.416**	0.655**	-0.789**	1				
Silt	0.636**	0.251*	-0.147	-0.194	-0.346**	1			
OM	-0.579**	-0.917**	-0.193	0.351**	-0.153	-0.198	1		
H.C	0.320**	0.343**	0.299**	-0.874**	0.622**	0.234*	-0.021	1	
S.D	0.148	-0.322**	-0.037	0.457**	-0.444**	0.191	0.149	-0.431**	1

*: Correlation is significant at 0.05. **: Correlation is significant at 0.01 level. Vfs: Very fine sand, OM: Organic matter, H.C: Hydraulic conductivity, S.D: Structure degree

3.2. Weighting of soil erodibility criteria by revised PF SWARA method

The criteria used in the RUSLE-K calculation of soil erodibility, including organic matter, clay, sand, silt, very fine sand, hydraulic conductivity, and degree of structure, were also used to determine the criteria weights by the revised PF SWARA method. The opinions of three expert agricultural engineers experienced in soil science, particularly erosion, were consulted for this study. The weights of the experts were determined based on their level of experience in the field. Each expert was evaluated using the scale provided in Table 1, and the resulting expert weights are presented in Table 6.

Table 6- Weights of experts evaluating the criteria

Experts	*Linguistic Expressions*	μ	ν	π	*Expert Weights*
Expert 1	VI	0.7	0.55	0.456	0.372
Expert 2	VI	0.7	0.55	0.456	0.372
Expert 3	I	0.6	0.7	0.387	0.255

μ: membership degree, ν: non-membership degree, π: hesitation degree

In this study, the revised PF SWARA approach prioritize each criterion based on expert opinions using the traditional SWARA approach. Pairwise comparisons between each prioritised criterion and the one below are determined using PFNs. Expert opinions were obtained using the scale provided in Table 2, and the resulting linguistic expressions are presented in Table 7.

Table 7- Linguistic expression of the criteria in line with expert opinions

Criteria	*Expert 1*	*Expert 2*	*Expert 3*
OM (%)			
Clay (%)	AH	VVH	VH
S.D (Unitless)	VVH	AH	AH
H.C (cm/h)	VL	L	VL
Sand (%)	VH	VVH	H
Silt (%)	VL	AL	AL
Vfs (%)	M	ML	L

Vfs: Very fine sand, OM: Organic matter, H.C: Hydraulic conductivity, S.D: Structure degree; Absolutely high (AH), Very very high (VVH), Very high (VH), High (H), Medium high (MH), Medium (M), Medium low (ML), Low (L), Very low (VL), Absolutely low (AL)

The score function was calculated using the linguistic expressions obtained from pairwise comparisons, with Eqs. (10-11) (Table 8). Once the score function values were obtained through the use of PFSs in pairwise comparisons, the process of determining the criteria weights began.

Table 8- Expression of the score values of the criteria according to expert opinions

Criteria	*Expert 1* (μ, ν)	*Expert 2* (μ, ν)	*Expert 3* (μ, ν)	*1st. Crisp Values*	*2st. Crisp Values*	*3st. Crisp Values*
OM (%)						
Clay (%)	(0.95, 0.20)	(0.85, 0.30)	(0.80, 0.35)	0.931	0.816	0.759
S.D (Unitless)	(0.85, 0.30)	(0.95, 0.20)	(0.95, 0.20)	0.816	0.931	0.931
H.C (cm/h)	(0.20, 0.85)	(0.30, 0.75)	(0.20, 0.85)	0.159	0.264	0.159
Sand (%)	(0.80, 0.35)	(0.85, 0.30)	(0.70, 0.45)	0.759	0.816	0.644
Silt (%)	(0.20, 0.85)	(0.10, 0.95)	(0.10, 0.95)	0.159	0.054	0.054
Vfs (%)	(0.50, 0.60)	(0.40, 0.70)	(0.30, 0.75)	0.445	0.335	0.264

Vfs: Very fine sand, OM: Organic matter, H.C: Hydraulic conductivity, S.D: Structure degree. μ: membership degree, ν: non-membership degree, π: hesitation degree

The score functions obtained within the scope of the study were used to calculate the k_j value with the help of Eq. (13). Here, the most important criterion is accepted as 1 and k_j value is calculated by adding +1 to each criterion in the lower step. In the calculating the importance vector (q_j) value, the importance vectors of the other criteria were determined by dividing the previous importance vector by the k_j coefficient as shown in Eq. 14, again taking the most important criterion as 1. Due to the high number of steps, only the calculations made for expert 1 are shared in this section (Table 9).

Table 9- Calculation of criterion weights according to expert 1 opinions

Criteria	s_j* Crisp Values*	k_j	q_j	w_j
OM (%)		1.000	1.000	0.418
Clay (%)	0.931	1.931	0.518	0.216
S.D (Unitless)	0.816	1.816	0.285	0.119
H.C (cm/h)	0.159	1.159	0.246	0.103
Sand (%)	0.759	1.759	0.140	0.058
Silt (%)	0.159	1.159	0.121	0.050
Vfs (%)	0.445	1.445	0.084	0.035

Sj: the comparative significance of score value, kj: relative coefficient, qj: the weight of the importance vector, Wj: the final weight of criteria.

Then, the criteria weights were calculated for the other experts and the weights obtained as a result of the evaluations of all experts are given in Table 10. In the last column of Table 10, the aggregated criteria weights calculated by multiplying the expert weights of 0.372, 0.372 and 0.255 calculated in the first stage of the PF-SWARA method with the weights obtained from the criteria are given.

When the criteria weights calculated by the PF-SWARA method are examined, it is observed that the highest criterion weight is 0.418 for organic matter. Subsequently, the weights of criteria considered important for soil erodibility, namely clay, structure degree, and hydraulic conductivity, are obtained as 0.227, 0.120, and 0.100, respectively. In terms of the study, it is seen that the criteria levels of sand, silt, and very fine sand have less importance with weights of 0.058, 0.053, and 0.039, respectively. Pacci et al. (2023) conducted a study to calculate the criterion weights of erosion sensitivity parameters using the

fuzzy analytic hierarchy process approach. The study found that organic matter, bulk density, and clay had the highest criterion weights of 0.317, 0.224, and 0.264, respectively. These criterion weights were similar to those obtained with the revised PF-SWARA approach for organic matter and clay levels.

Table 10- Aggregated criterion weights of criteria

Criteria	*Expert 1*	*Expert 2*	*Expert 3*	*Aggregated Criterion Weights*
OM	0.418	0.418	0.418	0.418
Clay	0.216	0.230	0.238	0.227
S.D	0.119	0.119	0.123	0.120
H.C	0.103	0.094	0.106	0.100
Sand	0.058	0.052	0.065	0.058
Silt	0.050	0.049	0.061	0.053
Vfs	0.035	0.037	0.049	0.039

3.3. Spatial distributions of soil erodibility factor for both approaches

The normalized values of the soil properties obtained from the analysis results with the standard scoring function were used to calculate soil erodibility by the WLC technique, in line with expert opinions on determining of criteria weights for the parameters considered in the study. The calculation of soil erodibility was based on each criterion analyzed (Patrono 1998; Çakır & Dengiz 2021). Distribution maps were generated using the RUSLE-K method, which is utilized to obtain soil erodibility factor, and values obtained from our developed PF SWARA-K approach, as well as the IDW method. The distribution maps obtained from the traditional RUSLE-K method for estimating soil erodibility showed parallelism with those obtained from the PF SWARA-K method, with a statistically significant relationship at the 1% level (0.691**). When examining the spatial distributions of soil erodibility values, sensitive areas to soil erosion are observed in the central regions in both prediction methods (Figure 3). It is observed that similar results are obtained for both parameters and soil erodibility increases as the coastal areas are approached.

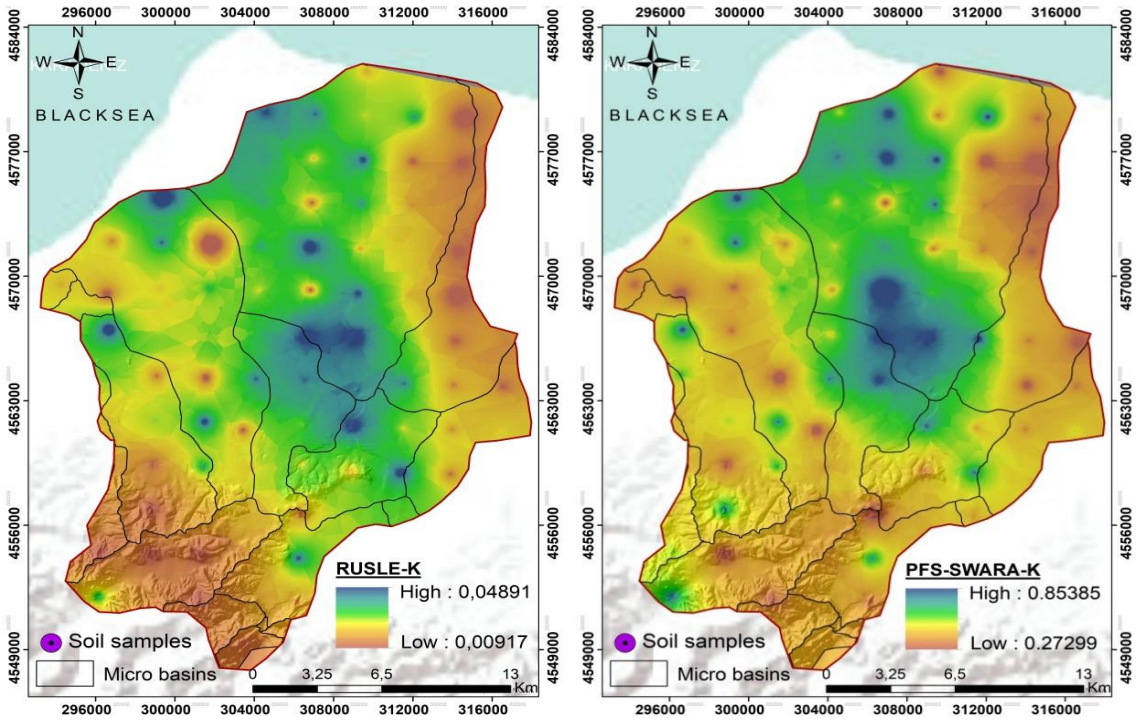


Figure 3- Distribution maps of RUSLE - K and PF SWARA-K

4. Conclusions

The soil erodibility factor is an important parameter used to determine the sensitivity of soils to erosion. It can be said that the erodibility factor in prediction models such as USLE and RUSLE used in erosion studies has been calculated with different

applications in different studies (Delgado et al. 2023; Pontes et al. 2022; Panagos et al. 2014) and the studies conducted in terms of literature are accepted.

The present study, it was aimed to compare the RUSLE-K factor calculated by traditional methods with PF-SWARA and distribution maps in an area located in the micro-basins of Çarşamba district of Samsun province of Central Black Sea. All parameters used in the traditional method were used as the basis for the determination of the criteria for the proposed approach. On the other hand, the main approach of in the PF SWARA-K method was considered expert knowledge approach considered the contribution and effect of the considered parameters on soil erosion susceptibility. As a result of the study, it was observed that the revised PF SWARA-K method obtained significant results at 1% level with 0.691** compared to the RUSLE-K factor calculated by traditional methods. As a result, it can be said that the fuzzy-multi criteria decision methods are an alternative approach that can be used in determining the soil erodibility factor. These results which we obtained with the expert opinions used in determining the criteria weights are likely to obtain much better results with the improvements to be made in the importance levels of the criteria and expert opinions in future studies. In addition, this study was carried out in an area with a semi-humid-humid ecological characteristic, and it is recommended to carry out studies in areas with different ecological characteristics.

Statements & Declarations

Ethical Approval:

We hereby would like to warrant that the manuscript represents original work that is not being considered for publication, in whole or in part, in another journal, book, conference proceedings, or government publication with a substantial circulation. I would like to warrant that all previously published work cited in the manuscript has been fully acknowledged. In addition, our manuscript has not been submitted to a preprint server prior to submission on any journal.

Competing Interest

Financial interests: All authors declare that they have no financial interests.

Conflict of Interest Statement:
The authors declare that they have no conflict of interest related to the content of this manuscript. The authors have no relevant financial or non-financial interests to disclose

Data Availability Statement:

The datasets generated and analysed during the current study are not publicly available but are available from the corresponding author on reasonable request.

Funding:

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Authors' contributions:

All authors contributed to the study conception and design. Methodology, Conceptualization and Resources were done by Orhan Dengiz, Aykut Çağlar and Barış Özkan. Supervision was performed by Orhan Dengiz. Material preparation, investigation and analysis were performed by Aykut Çağlar and Barış Özkan. All maps created by Orhan Dengiz. The first draft of the manuscript was written by Orhan Dengiz, Aykut Çağlar and Barış Özkan commented on previous versions of the manuscript. All authors read and approved the final manuscript

References

Andrews S S, Karlen D L & Mitchell J P (2002). A comparison of soil quality indexing methods for vegetable production systems in Northern California. *Agri Ecosyst Environ* 90: 25–45. doi:10.1016/S0167-8809(01)00174-8

Baskan O & Dengiz O (2008). Comparison of traditional and geostatistical methods to estimate soil erodibility factor. *Arid Land Research and Management* 22(1): 29–45. https://doi.org/10.1080/15324980701784241

Baskan O (2022). Analysis of spatial and temporal changes of RUSLE-K soil erodibility factor in semi-arid areas in two different periods by conditional simulation, *Archives of Agronomy and Soil Science* 68:12, 1698-1710. DOI: 10.1080/03650340.2021.1922673

Beretta A N & Carrasco Letelier L (2017). USLE/RUSLE K-factors allocated through a linear mixed model for Uruguayan soils. *Ciencia e investigación agraria: revista latinoamericana de ciencias de la agricultura* 44(1): 100-112. http://dx.doi.org/10.7764/rcia.v44i1.1622

Bouyoucos G J (1962). Hydrometer method improved for making particle size analyses of soils 1. *Agronomy journal* 54(5): 464-465

Bölük E (2016). According to Erinç Climate Classification Turkish Climate, Ministry of Forestry and Water Management General Directorate of Meteorology, Ankara

Cartwright J H, Shammi S A & Rodgers III J C (2022). Use of Multi-Criteria Decision Analysis (MCDA) for Mapping Erosion Potential in Gulf of Mexico Watersheds. *Water* 14(12): 1923. https://doi.org/10.3390/w14121923

Cui Y, Liu W, Rani P & Alrasheedi M (2021). Internet of Things (IoT) adoption barriers for the circular economy using Pythagorean fuzzy SWARA-CoCoSo decision-making approach in the manufacturing sector. *Technological Forecasting and Social Change* 171: 120951. https://doi.org/10.1016/j.techfore.2021.120951

Çakır M & Dengiz O (2021). Land Evaluation Study Using Linear Combination Technique, Case Study Sefali Village. *Journal of Soil Science and Plant Nutrition* 9(1): 43-56. https://doi.org/10.33409/tbbbd.899746 (In Turkish)

Dede V, Dengiz O, Demirağ Turan İ, Zorlu K, Pacci S & Serin S (2022). Determination of erosion susceptibilities of soils formed on the periglacial landforms of mount Ilgar and its estimation using artificial neural network (ANN). *International Journal of Geography and Geography Education (IGGE)* 47: 1-22. https://doi.org/10.32003/igge.1097942 (In Turkish)

Delgado D, Sadaoui M, Ludwig W & Mendez W (2023). Depth of the pedological profile as a conditioning factor of soil erodibility (RUSLE K-Factor) in Ecuadorian basins. *Environmental Earth Sciences* 82(12): 286. https://doi.org/10.1007/s12665-023-10944-w

Demirağ Turan İ & Dengiz O (2017). Erosion Risk Prediction Using Multi-Criteria Assessment in Ankara Güvenç Basin. *Journal of Agricultural Sciences* 23(3): 285-297. https://doi.org/10.15832/ankutbd.447600 (In Turkish)

Dotterweich M (2013). The history of human-induced soil erosion: Geomorphic legacies, early descriptions and research, and the development of soil conservation—A global synopsis. *Geomorphology* 201: 1–34. https://doi.org/10.1016/j.geomorph.2013.07.021

Efthimiou N (2018). The importance of soil data availability on erosion modeling. *Catena* 165:551–566. https://doi.org/10.1016/j.catena.2018.03.002

Gao G, Liang Y, Liu J, Dunkerley D & Fu B (2023). A modified RUSLE model to simulate soil erosion under different ecological restoration types in the loess hilly area. *International Soil and Water Conservation Research* https://doi.org/10.1016/j.iswcr.2023.08.007

İmamoğlu A & Dengiz O (2017). Determination of soil erosion risk using RUSLE model and soil organic carbon loss in Alaca catchment (Central Black Sea region, Turkey). *Rendiconti Lincei* 28(1): 11-23. https://doi.org/10.1007/s12210-016-0556-0

Jackson M L (1958). Soil Chemical Analysis. Prentice Hall Inc., Englewood Cliffs, N.J.

Kamali Saraji M, Streimikiene D & Ciegis R (2022). A novel Pythagorean fuzzy-SWARA-TOPSIS framework for evaluating the EU progress towards sustainable energy development. *Environmental monitoring and assessment* 194(1): 42. https://doi.org/10.1007/s10661-021-09685-9

Keshavarzı A & Sarmadian A (2012). Mapping of spatial distribution of soil salinity and alkalinity in a semi-arid region. *Annals of Warsaw University of Life Sciences, Land Reclamation* 44(1): 3–14

Keršulienė V, Zavadskas E K & Turskis Z (2010). Selection of rational dispute resolution method by applying new step-wise weight assessment ratio analysis (Swara). *Journal of Business Economics and Management* 11(2): 243-258. https://doi.org/10.3846/jbem.2010.12

Klute A & Dirksen C (1986). Hydraulic conductivity and diffusivity: Laboratory methods. *Methods of soil analysis: Part 1 physical and mineralogical methods* 5: 687-734

Liebig M A, Varvel G & Doran J (2001). A simple performance-based ındex for assessing multiple agroecosystem functions. *Soil and Crop Management* 93(2): 313-318. https://doi.org/10.2134/agronj2001.932313x

Malczewski J & Rinner C (2015). Multicriteria decision analysis in geographic information science 1:55-77. *New York: Springer*

Miháliková M & Dengiz O (2019). Towards more effective irrigation water usage by employing land suitability assessment for various irrigation techniques. *Irrigation and Drainage* 68(4): 617-628. https://doi.org/10.1002/ird.2349

Mercan Ç (2023). Coğrafi bilgi sistemi ve ahp ile arıcılık faaliyet alanları için arazi uygunluk değerlendirmesi: Bitlis/Türkiye örneği (Land suitability assessment for Apiculture (Beekeeping) activity areas using Geographic Information System and AHP: A Case Study Bitlis/Türkiye, Extended Abstract in English). *U. Arı D. / U. Bee J* 23(1): 61-77. DOI: 10.31467/uluaricilik.1245078

Pacci S, Saflı M E & Dengiz O (2023). Fuzzy-Analitic Hierarchy Process Approach in Soil Erodibility Studies under SemiHumid Ecological Conditions. *COMU Journal of Agriculture Faculty* 11(1): 148-165. https://doi.org/10.33202/comuagri.1276110 (In Turkish)

Panagos, P., Meusburger, K., Ballabio, C., Borrelli, P., & Alewell, C. (2014). Soil erodibility in Europe: A high-resolution dataset based on LUCAS. *Science of the total environment* 479: 189-200. https://doi.org/10.1016/j.scitotenv.2014.02.010

Parlak M, Yiğini Y & Ekinci H (2014). Seasonal Change of Erodibility in the Soils of Çanakkale-Umurbey Plain. *COMU Journal of Agriculture Faculty* 2(1): 123–131(In Turkish)

Patrono A (1998). Multi-Criteria Analysis and Geographic Information Systems: Analysis of Natural Areas and Ecological Distributions. Multicriteria Analysis for Land-Use Management, Edited by Euro Beinat and Peter Nijkamp, *Kluwer Academic Publishers, Environment and Management* 9: 271-292. AA Dordrecht, TheNetherlands. https://doi.org/10.1007/978-94-015-9058-7_15

Peng X & Yang Y (2016). Pythagorean fuzzy choquet integral based MABAC method for multiple attribute group decision making. *Int. J. Intell. Syst* 31 (10): 989–1020. https://doi.org/10.1002/int.21814

Perez Rodriguez R, Marques M J & Bienes R (2007). Spatial variability of the soil erodibility parameters and their relation with the soil map at subgroup level. *Sci. Total Environ* 378 (1e2): 166e173. https://doi.org/10.1016/j.scitotenv.2007.01.044

Pimentel D & Burgess M (2013). Soil erosion threatens food production. *Agriculture* 3(3): 443-463. https://doi.org/10.3390/agriculture3030443

Pirmoradian N, Rezaei M, Davatgar N, Tajdari K & Abolpour B (2010). Comparing of interpolation methods in rice cultivation vulnerability mapping due to groundwater quality in Guilan, north of Iran. *International Conference on Environmental Engineering and Applications (ICEEA)* 147–150. Singapore

Pontes S F, Silva Y J A B D, Martins V, Boechat C L, Araújo A S F, Dantas J S & Barbosa R S (2022). Prediction of Soil Erodibility by Diffuse Reflectance Spectroscopy in a Neotropical Dry Forest Biome. *Land* 11(12): 2188. https://doi.org/10.3390/land11122188

Rani P, Mishra A R, Mardani A, Cavallaro F, Štreimikienė D & Khan S A R (2020). Pythagorean fuzzy SWARA–VIKOR framework for performance evaluation of solar panel selection. *Sustainability* 12(10): 4278. https://doi.org/10.3390/su12104278

Renard K G, Foster G R, Weesies G A, McCool D K & Yoder D C (1997). Predicting Soil Erosion by Water: A Guide to Conservation Planning with the Revised Universal Soil Loss Equation (RUSLE). *USDA Agriculture Handbook* 703. USDA, Washington, DC, USA

Saeidi P, Mardani A, Mishra A R, Cajas V E C & Carvajal M G (2022). Evaluate sustainable human resource management in the manufacturing companies using an extended Pythagorean fuzzy SWARA-TOPSIS method. *Journal of Cleaner Production* 370: 133380. https://doi.org/10.1016/j.jclepro.2022.133380

Saygın S D, Basaran M, Ozcan A U, Dolarslan M, Timur O B, Yilman F E & Erpul G (2011). Land degradation assessment by geo-spatially modeling different soil erodibility equations in a semi-arid catchment. *Environmental Monitoring and Assessment* 180: 201-215. https://doi.org/10.1007/s10661-010-1782-z

Teegavarapu R & Chandramouli V (2005). Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records. *J Hydrol* 312: 191–206. https://doi.org/10.1016/j.jhydrol.2005.02.015

Valkanou,K, Karymbalis E, Papanastassiou D, Soldati M, Chalkias C & GakiPapanastassiouK (2021). Assessment of Neotectonic Landscape Deformation in Evia Island, Greece, Using GIS-Based Multi-Criteria Analysis. *ISPRS Int. J. Geo.-Inf.* 10: 118. https://doi.org/10.3390/ijgi10030118

Wilding L P (1985). Spatial variability: its documentation, accomodation and implication to soil surveys. *In Soil spatial variability* 166-194

Wischmeier W H& Smith D D (1965). Predicting Rainfall-Erosion Losses from Cropland East of the Rocky Mountains. *USDA*, Washington, DC, USA

Wischmeier W H& Smith D D (1978). Predicting rainfall erosion losses. *USDA*, Washington, DC, USA

Wu SJ& Wei GW (2017). Pythagorean fuzzy hamacher aggregation operators and their application to multiple attribute decision making. *Int. J. Knowledge-based Intell. Eng. Syst* 21: 189–201. DOI: 10.3233/KES-170363

Yager R R (2013). Pythagorean membership grades in multicriteria decision making. *IEEE Transactions on Fuzzy Systems* 22(4): 958–965. doi: 10.1109/TFUZZ.2013.2278989

Yang D, Kanae S, Oki T, Koike T& Musiake K (2003). Global potential soil erosion with reference to land use and climate changes. *Hydrol. Process* 17: 2913–2928. https://doi.org/10.1002/hyp.1441

Zadeh L A (1965). Fuzzy sets. *Information and control* 8(3): 338-353

Zavadskas E K, Čereška A, Matijošius J, Rimkus A & Bausys R (2019). Internal Combustion engine analysis of energy ecological parameters by neutrosophic MULTIMOORA and SWARA methods. *Energies* 12(8): 1415. https://doi.org/10.3390/en12081415

Zhang H, Zhang J, Zhang S, Yu C, Sun R, Wang D, Zhu C& Zhang J (2020). Identification of Priority Areas for Soil and Water Conservation Planning Based on Multi-Criteria Decision Analysis Using Choquet Integral. *Int. J. Environ. Res. Public Health* 17: 1331

Zhang X& Xu Z (2014). Extension of TOPSIS to multiple criteria decision making with pythagorean fuzzy sets. *Int. J. Intell. Syst* 29 (12): 1061–1078. https://doi.org/ 10.1002/int.21676

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

In vivo Metabolic Investigation of Oxygen, Light, and Temperature Effects on Dormancy Alleviation of Sesame (*Sesamum indicum* L.) Seeds

Honghao Cai*[a], Xiayi Ruan[b], Yumin Wan[b], Mengting Chen[a], Xianqin Wu[c], Yingqiang Cai[b]

[a]*Department of Physics, School of Science, Jimei University, Xiamen, Fujian Province, CHINA*

[b]*School of Marine Engineering, Jimei University, Xiamen, Fujian Province, CHINA*

[c]*School of Finance and Economics, Jimei University, Xiamen, Fujian Province, CHINA*

ARTICLE INFO
Research Article
Corresponding Author: Honghao Cai, E-mail: hhcai@jmu.edu.cn
Received: 28 July 2024 / Revised: 17 September 2024 / Accepted: 26 September 2024 / Online: 14 January 2025

Cite this article
Cai H, Ruan X, Wan Y, Chen M, Wu X, Cai Y (2025). *In vivo* Metabolic Investigation of Oxygen, Light, and Temperature Effects on Dormancy Alleviation of Sesame (*Sesamum indicum* L.) Seeds. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):196-206. DOI: 10.15832/ankutbd.1523409

ABSTRACT
As an edible seed, sesame seeds require careful storage to maintain their quality. Dormancy helps seeds extend their lifespan by slowing down metabolic processes, reducing energy consumption and natural aging. However, seeds may exit dormancy and begin germination during storage due to variations in temperature, light, and oxygen conditions. This transition is not easily visible, but nutritional components within the seeds can start to deplete. In this study, non-invasive magnetic resonance spectroscopy and imaging were used to monitor sesame seeds stored under different temperature, light, and oxygen conditions for over 120 hours. Results showed that seeds remained dormant at 15 °C under oxygen deprivation and in the absence of light. When exposed to continuous light at 15 °C, under anaerobic or aerobic conditions, changes in metabolic resonances were observed through spectroscopy, indicating moisture and fatty acid transfer between seed structures. Despite these changes, magnetic resonance imaging showed that the embryo did not develop. At 24 °C with continuous light and aerobic conditions, both spectroscopy and imaging analyses revealed significant metabolic changes, and all internal seed structures developed normally, with visible signs of germination. This study highlights that although sesame seeds are non-photoblastic, light can still trigger metabolic activity within the seeds, while suitable temperature is essential for complete seed development. These findings provide valuable insights into the dynamic molecular-level metabolic changes from dormancy to early seed germination using magnetic resonance technology and offer guidance for maintaining seed dormancy during storage.

Keywords: NMR, MRI, ^{1}H NMR spectroscopy, Chemical Shift Selective Imaging, Germination, Storage

1. Introduction

Seed storage conditions, including temperature, oxygen availability, and light exposure, significantly impact their viability, germination rate, and overall quality (Gebeyehu 2020). Low temperatures help to slow down seed metabolism and reduce moisture loss, which aids in maintaining seed viability. However, temperatures that are either too low or too high may have detrimental effects on the seeds (Simon et al. 1976; Egli & Wardlaw 1980; Geneve 2003). Oxygen is a critical gas for seed metabolism, promoting water absorption and thereby promoting seed germination (Gay et al. 1991; Corbineau 2022). Certain seeds are light-sensitive, and exposure to light can trigger photosynthesis and produce excessive free radicals which may compromise seed quality (An et al. 2020; Cheng et al. 2023). As a widely consumed edible seed, sesame seeds require meticulous attention to their storage conditions to ensure their consumable quality and safety. Currently, the primary technical methods for investigating the effects of storage and cultivation conditions on seeds include: the measurement of phenotypic date (Cao et al. 2020; Chen et al. 2023; Elboghdady et al. 2023; Pasternak et al. 2024), gas chromatography mass spectrometry (Taylor et al. 1999; Yan et al. 2018), liquid chromatography-mass spectrometry (Chiwocha et al. 2003), high-performance liquid chromatography mass spectrometry (Jander et al. 2004; Kołodziejczyk et al. 2015; Wu et al. 2017), infrared spectroscopy and imaging (Tigabu et al. 2004; Song et al. 2015; Zhang et al. 2020). However, during the transition from dormancy to the early stage of germination, seeds do not exhibit observable morphological changes. As a result, phenotypic data collection is more applicable to the mid-to-late stage of germination, where visible changes occur. During germination, even in the absence of external changes, various metabolites may still be produced within the seeds. Therefore, understanding the physiological changes, as well as the composition and distribution of metabolites, in seeds during the early stage of germination is crucial. Gas chromatography mass spectrometry (Mota et al. 2021), liquid chromatography-mass spectrometry (Li et al. 2020) can obtain detailed, high-resolution metabolic profiles. However, these methods are destructive, and cannot be used for dynamic monitoring of seeds. Infrared spectroscopy and imaging, while enabling non-destructive detection of seeds (Amanah et al. 2020; Ma et al.

2020; Xiao et al. 2024), suffer from relatively low resolution and decreased sensitivity when applied to seeds with low moisture content. In addition, a variety of data preprocessing steps is required before analyzing infrared spectral data, such as the baseline correction, smoothing, and standard normal variate.

Nuclear magnetic resonance (NMR) is a suitable non-invasive technique for studying plants. Time-domain nuclear magnetic resonance relaxometry offers a reliable approach for assessing both oil seed temperature within soils and soil thermal diffusivity (Carosio et al. 2018). ^{31}P NMR spectroscopy helps to enhance the comprehension of herbicide-induced metabolic changes detected in live plants (d'Avignon et al. 2018). Quantitative and high-resolution fingerprints of marine microalgae have been established using ^{1}H High-resolution magic angle spinning (MAS) NMR (Caprara et al. 2023). The field of magnetic resonance imaging (MRI) is also rapidly expanding and offers significant advantages in the *in vivo* plant sciences (Blystone et al. 2024). Magnetic resonance imaging technology can measure the water content of leaves and study its variation throughout the day, such as detecting changes in the cellular and tissue water status and distribution in *Brassica napus* leaves during blade development and dehydration processes (Boulc'h et al. 2024). It also plays a beneficial role in detecting plant diseases. For instance, it facilitates the early detection of cotton *Verticillium wilt* (Tang et al. 2023), as well as the characterization of changes induced by *Phytophthora cactorum* infection in strawberry seedlings (Tuomainen et al. 2024).

Presently, research using NMR and MRI technologies to observe metabolic changes in seeds during the germination process remains limited in scope (Sarkar et al. 2009; Cai et al. 2018a; Xiao et al. 2024; Liao et al. 2024). Sesame, as an oil seed, primarily derives its economic and nutritional value from its rich oil content, which includes a range of fatty acids beneficial to human health, including linoleic acid, oleic acid, stearic acid, and palmitic acid (Dossa et al. 2017). However, during the germination process, sesame seeds use their stored oils as a source of nutrients and energy required for growth (Xiao et al. 2024). Therefore, the effective management of seed dormancy and prevention of the transition to the germination stage is of significant importance. As far as we know, there is limited research (Cai et al. 2018a) on the metabolism of seeds from dormancy to the early germination stages under different storage conditions.

One-dimensional ^{1}H NMR spectroscopy (Bharti & Roy 2012) is among the most prevalent techniques in NMR. It requires only a single pulse sequence to acquire spectra, with relatively fewer parameter adjustments. In contrast to certain other NMR methodologies like MAS NMR (Cai et al. 2016a; Cai et al. 2018a) and intermolecular multi-quantum coherence technique (Cai et al. 2016b; Cai et al. 2018b), ^{1}H NMR does not necessitate the use of specialized rotors, complex sequences and parameter configurations. Chemical Shift Selective (CHESS) imaging is a specific MRI technique (Haase et al. 1985). It selectively detects specific metabolites through the chemical shift effect, enabling accurate quantification of their concentration and distribution within complex biological tissues (Fujii et al. 2022). Given the distinctive features and advantages of these two techniques, this study proposed to dynamically monitor the biophysical properties of sesame seeds using one-dimensional ^{1}H NMR spectroscopy and CHESS. These approaches would elucidate how temperature, light, and oxygen influence metabolite variations and distributions. By investigating the impact of these factors on sesame seed metabolites, the aim is to identify key factors that induce sesame seeds to transition from dormancy to early germination, providing scientific insights for sesame seed storage.

2. Material and Methods

2.1. Sample preparation

Black sesame seeds (*Sesamum indicum var. radiatum*) were purchased from a local seed retailer one week after their harvest. A germination test of 50 seeds was performed to determine their viability, and the germination rate reached 96%, meeting the requirement for research purposes. For preparing acidic liquid culture medium (pH=5.4), the HNO_3 was purchased from Sigma-Aldrich Co (St. Louis, MO, USA). The procedure for preparing the medium is as follows: Add deionized water to a diluted 10 ml 1% HNO_3 solution and use a pH meter to monitor the pH until the target pH is reached. After preparation, sterilize the medium using an autoclave. Once cooled, store the medium in sterile containers to avoid contamination. The reason for preparing an acidic liquid culture medium was that this study investigated the impact of temperature, light, and oxygen on sesame seeds from dormancy to early germination. To enable rapid germination under controlled conditions, sesame seeds were cultivated in an acidic liquid culture medium. Oxygen conditions were varied by the extent to which the seeds were submerged in the acidic liquid. Complete submersion simulates hypoxic conditions, while partial coverage mimics aerobic conditions. Light for the sesame seeds was provided by a 50-watt full-spectrum LED lamp (Opple Lighting Co., China).

2.2. Experiment 1

Sesame seeds were not exposed to light, cultivated at 15 °C, and fully covered with acidic liquid culture medium. There measurements were taken at 35 time points: 1-110, 2-510, 3-835, 4-1120, 5-1560, 6-1915, 7-2160, 8-2630, 9-2880, 10-3300, 11-3615, 12-3960, 13-4320, 14-4680, 14-5080, 15-5474, 16-5820, 17-6140, 18-6505, 19-6890, 20-7260, and 21-7660 min.

2.3 Experiment 2

Sesame seeds were exposed to continuous light, cultivated at 15 °C, and fully covered with acidic liquid culture medium. The

measurements were taken at 35 time points: 1-360, 2-480, 3-720, 4-1050, 5-1410, 6-1680, 7-1800, 8-1920, 9-2160, 10-2760, 11-3240, 12-3360, 13-3600, 14-3840, 15-4380, 16-5280, 17-5490, 18-5580, 19-5740, 20-6120, 21-6720, 22-7440, 23-7600, 24-7825, 25-8390, 26-9060, 27-9760, 28-9900, 29-10080, 30-10440, 31-10800, 32-11320, 33-12360, 34-13800, and 35-15180 min.

2.4 Experiment 3

Sesame seeds were exposed to continuous light, cultivated at 15 °C, and two-thirds covered with acidic liquid culture medium. The measurements were taken at 35 time points: 1-100, 2-510, 3-540, 4-1050, 5-1260, 6-1920, 7-2310, 8-2790, 9-3300, 10-3840, 11-4170, 12-4590, 13-5070, 14-5670, 15-5970, and 16-9950 min.

2.5 Experiment 4

Sesame seeds were exposed to continuous light, cultivated at 20 °C, and two-thirds covered with acidic liquid culture medium. The measurements were taken at 35 time points: 1-90, 2-720, 3-1080, 4-1440, 5-1620, 6-2460, 7-2730, 8-3240, 9-4159.8, 10-5740.2, 11-6439.8, 12-6840, and 13-8650.2 min.

2.6 ^{1}H NMR spectroscopy

The ^{1}H NMR spectra were acquired on a Varian UNITY INOVA 500 MHz NMR spectrometer (Agilent Technologies Inc., Santa Clara, CA, USA) at 20 °C. A single 90° pulse excitation was used with the following parameters: a 90° pulse width of 2 μs, a recycle delay of 3 s, an acquisition time of 0.02 s, and 16 transients. The spectra were processed using the Vnmrj 4.0 software (Agilent Technologies Inc., Santa Clara, CA, USA), which involved Fourier transformation, phase correction, baseline correction, zero filling. To ensure reproducibility and reliable error control, the experiments were conducted on five sets of samples, and the results were averaged.

2.7 Chemical shift selective imaging

CHESS experiments were conducted using a Varian UNITY INOVA 500 MHz NMR spectrometer with a μ-MRI System, employing the CHESS sequence (Haase *et al.*, 1985) at 20 °C. The TR and TE used for CHESS were 1000 ms and 15 ms respectively, and the contrast of the CHESS images was T1-weighted. As the sesame seeds grew, some imaging parameters of CSSI were adjusted accordingly.

2.8 Experimental flowchart

The experimental flowchart is shown in Figure 1. The process can be summarized as follows: Sesame seeds were cultivated under four different conditions. ^{1}H NMR and CHESS were used to analyze seeds at various time points. The obtained NMR data were combined with phenotypic data for analysis to study and compare the metabolic and structural changes in sesame under different cultivation environments.

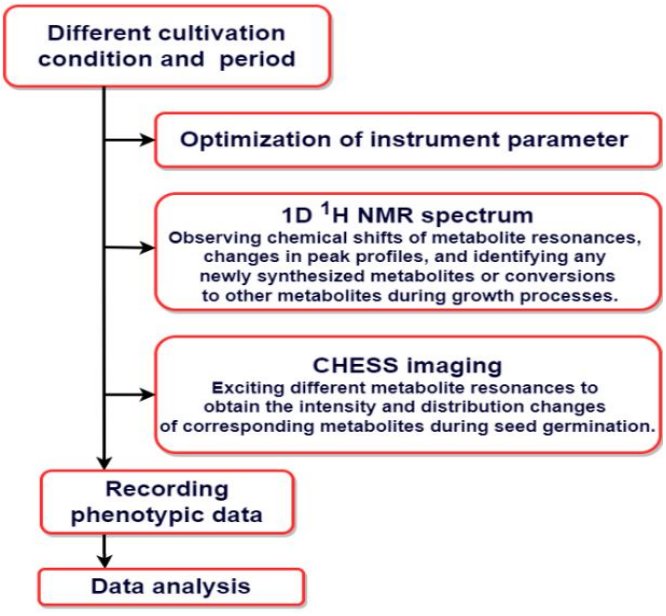


Figure 1- Experimental flowchart

3. Results and Discussion

3.1 Experiment 1

The 1D ^{1}H NMR spectra collected at various time points during Experiment 1 are shown in Figure 1. The primary identifiable resonances in the spectra are I_1, I_2, and I_3. Peaks I_1 and I_2 are associated with fatty acid features, specifically $–(CH_2)_n–$ (methylene protons) and $–CH_2–CH=CH–CH_2$ (allylic methylene protons), respectively. Peak I_3 corresponds to the water signal, H–O–H (water protons). Over time, there were no significant changes in the chemical shifts of these characteristic peaks, nor were any new or transformed metabolites detected. This suggested that the sesame seeds remained in a dormant state throughout the duration of the experiment (Terskikh et al. 2011).

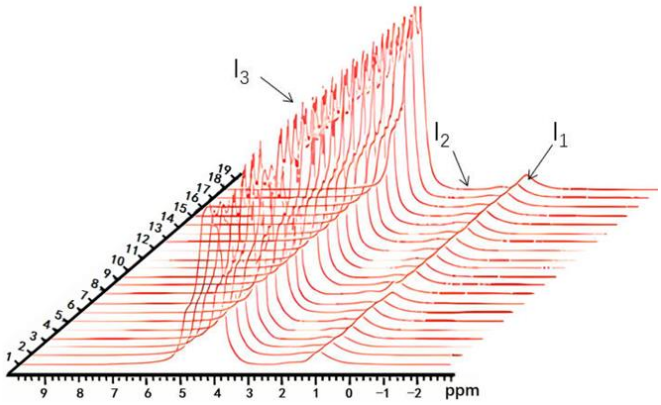


Figure 2- The 1D ^{1}H spectra of sesame seeds at different time points (Experiment 1)

Figure 3 shows the images obtained using CHESS after exciting the resonances of sesame seeds labelled as I_1, I_2, and I_3 at 7260 min. From the images, it was evident that the sesame seeds exhibited no apparent signs of germination. Lipids were evenly distributed within the sesame seeds, while moisture predominantly accumulated in the underdeveloped cotyledons, embryo shoot, axis, and root. The metabolic profiles indicated no spatial interactions or transformations, reflecting the seed dormancy in physiological structure (such as seed coat, endosperm and embryo) and metabolism, which was consistent with the findings from the ^{1}H NMR spectroscopy.

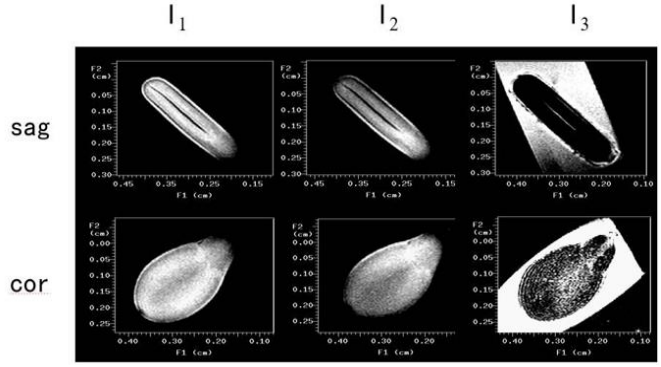


Figure 3- The CHESS sagittal (sag) and coronal (cor) images of sesame seeds at 7250 min, excited resonances at I_1, I_2, and I_3. Imaging parameters: FOV = 0.5 × 0.5 cm, matrix size = 260 × 130, slice thickness = 0.3 mm

3.2 Experiment 2

The ^{1}H NMR spectra at various time points for Experiment 2 are presented in Figure 4. From the spectra, a notable difference compared to Experiment 1 was observed: due to continuous light exposure, a marked increase in the water signal at 720 min indicated that the sesame seeds absorbed moisture through the micropyle, a key indicator of seed readiness for germination. However, from 720 to 3600 min, the spectra of Experiment 2 closely resembled those of Experiment 1, showing relative stability in water and lipid signals without apparent additions or transformations in metabolites. Starting at 3840 min, the initially sharp water signal transformed into a broader, lower peak. This change stems from two factors: a portion of the free water is oil seed d for cellular expansion and metabolic activities, while another portion binds with biomolecules within cells (such as proteins, polysaccharides), forming bound water. The signal associated with bound water appears as a broader and lower peak in the ^{1}H NMR spectra, primarily due to restricted molecular movement and shorter relaxation time (Xiao et al. 2024).

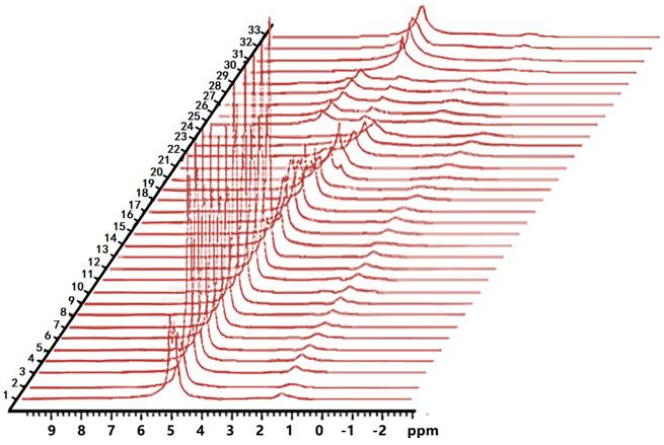


Figure 4- The 1D ^{1}H spectra of sesame seeds at different time points (Experiment 2)

To gain a clearer understanding of changes in lipid signals, we selected the spectra at key time points as shown in Figure 5. From the figure, at 8390 min, the prominent lipid characteristic peak I_1 near 1.35 ppm dramatically decreased and became less detectable, while the I_2 peak around 2.50 ppm notably increased. Moreover, both signals exhibited chemical shift variations simultaneously. During the time interval from 9760 to 10800 min, a new signal I_{new} appeared near 6.20 ppm, which subsequently disappeared after 11320 min. This suggested that various metabolic processes occurred within sesame seed cells, potentially yielding intermediate products such as metabolic transformations of fatty acids. As sesame germination progressed, these metabolites might have undergone further transformation or consumption, leading to signal disappearance.

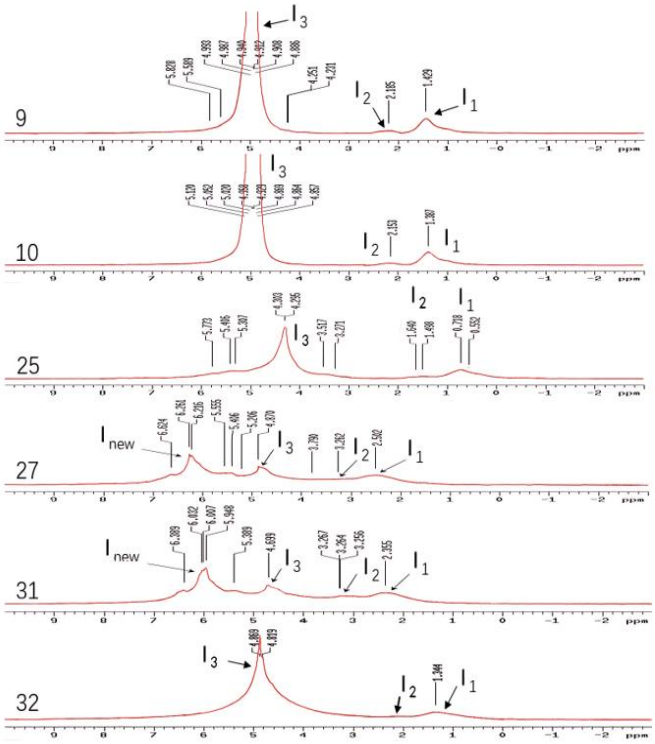


Figure 5- The 1D ^{1}H spectra of sesame seeds at key time points (Experiment 2)

Figure 6 depicts the images obtained using CHESS after exciting after exciting the resonances of sesame seeds labelled as I_1, I_2, and I_3 at 7440 min. Compared to Figure 3, the lipid distribution was no longer uniform but was concentrated around the radicle and micropylar endosperm near the micropyle, with lesser presence in the cotyledon portion of the embryo. From the sagittal image of I_3, it was observed that moisture, after permeating through the micropyle, primarily resided in the endosperm and internal cavities of the seed, gradually absorbed by the internal embryo. Moisture remained primarily concentrated in the underdeveloped cotyledons, embryo shoot, axis, and root, but compared to Figure 3, the extent of water signals had expanded. At the end of the experiment at 10800 min, the sesame seeds showed no visible signs of germination. However, observations from the spectra and CHESS imaging indicated that the sesame seeds had exited dormancy.

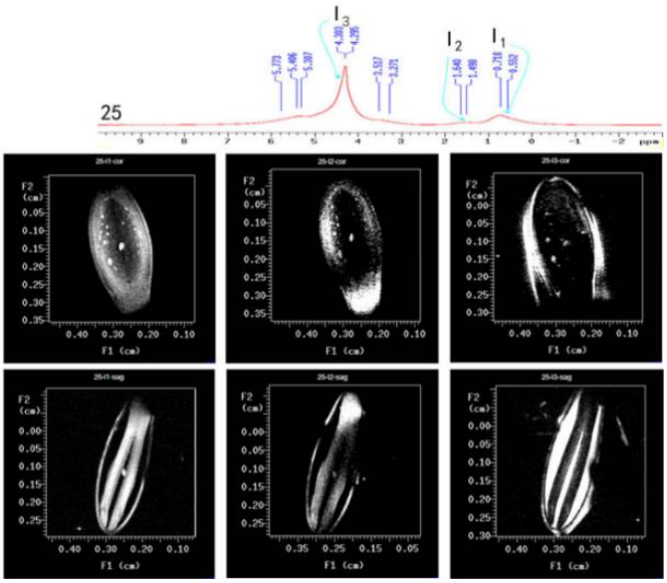


Figure 6- The CHESS sagittal (sag) and coronal (cor) images of sesame seeds at 7440 min, excited resonances at I_1, I_2, and I_3 (The images are arranged from right to left). Imaging parameters: FOV = 0.5 × 0.5 cm, matrix size = 260 × 130, nt = 16, slice thickness = 0.3 mm

3.3 Experiment 3

The ^{1}H NMR spectra at various time points for Experiment 3 are shown in Figure 7. Compared to Experiment 2, the spectra from Experiment 3 reveal continuous metabolic transformations beginning from the second (2-510) time point, leading to consistent fluctuations in chemical shifts. Notably, the signals at 5.2 ppm and 5.7 ppm were significantly enhanced, corresponding to the metabolic products of aliphatic protons and phenolic compounds. This indicated that metabolic activity in the sesame seeds in Experiment 3 had become markedly more active. At the twelfth (12-4590) time point, a new characteristic peak around 6 ppm appeared, likely corresponding to glycoside compounds. This signal subsequently diminished as these compounds were metabolized into energy and carbon sources.

In Experiment 2, with the seeds completely submerged in the culture medium and under anoxic conditions, the metabolic processes of the seeds were likely restricted. The seeds may have shifted towards anaerobic metabolic pathways, such as lactic acid fermentation or alcoholic fermentation, to generate a small amount of energy to sustain basic survival needs, thereby limiting seed growth and development. In Experiment 3, where the seeds were only partially submerged in the culture medium, the seeds were able to undergo respiration, using carbohydrates to release energy. This process generated substantial amounts of energy, which supported seed growth and germination. Moreover, aerobic conditions promote the oxidation of fatty acids, providing additional energy.

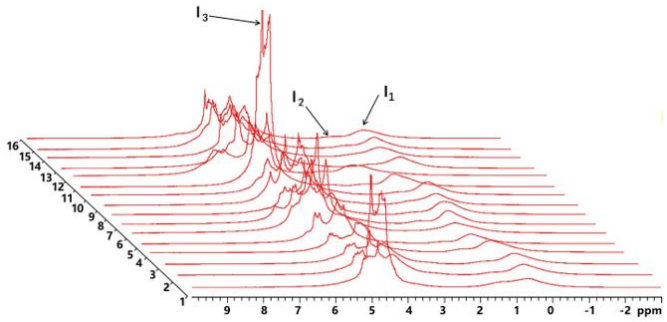


Figure 7- The 1D ^{1}H spectra of sesame seeds at different time points (Experiment 3)

Thus, Experiment 3 yielded a higher production of metabolites in sesame seeds. We excited these metabolites by using CHESS, and the results are depicted in Figure 8. From the CHESS images, we observed a broader distribution and stronger signals of metabolites, indicating enhanced seed growth and metabolism under aerobic conditions. However, similar to Experiments 1 and 2, no visible signs of germination were observed until the final time point of 9950 min.

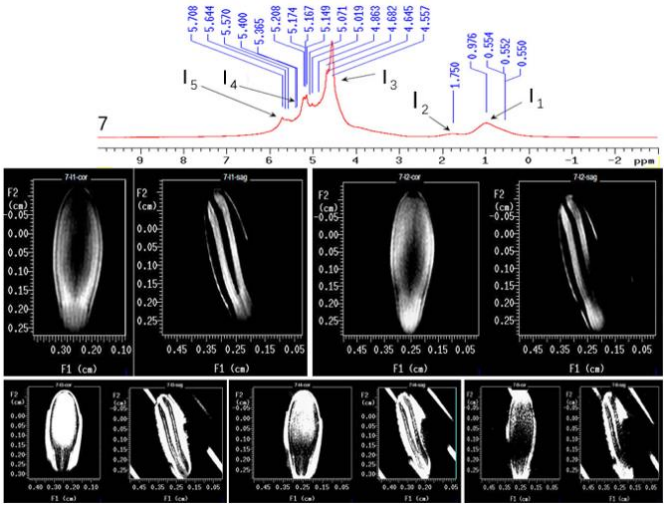


Figure 8- The CHESS sagittal (sag) and coronal (cor) images of sesame seeds at 7250 min, excited resonances at I_1, I_2, I_3, I_4, and I_5 (Clockwise arrangement of the images). Imaging parameters: FOV = 0.5 × 0.5 cm, matrix size = 160 × 80, nt = 16, slice thickness = 0.3 mm

3.4 Experiment 4

The ^{1}H NMR spectra at various time points for Experiment 4 are depicted in Figure 9, showing four distinct growth phases. From the figure, it is evident that seed growth changes rapidly. At the first (1-90) and second (2-720) time points, a conversion phenomenon was observed (characteristic peaks of sesame metabolites shift to higher frequencies), followed by a movement towards lower frequencies by the fourth (4-1440) time point. At the fifth (5-1620) time point, a new characteristic peak appeared at a high-frequency position, marking the onset of the third phase. Subsequent chemical shifts near lower frequencies indicated transformations of sesame characteristic peaks towards higher frequencies. By the eleventh (11-6439.8) time point, these peaks were completely engulfed by water peaks, yet a new peak emerged around 6.67 ppm, identified as phenolic compounds, gradually diminished thereafter. Ultimately, the overall waveform broadened and flattened, indicating that the sesame seeds were maturing, as the reduction or transformation of lipid metabolic pathways impacted the intensity of lipid components reflected in the NMR spectra.

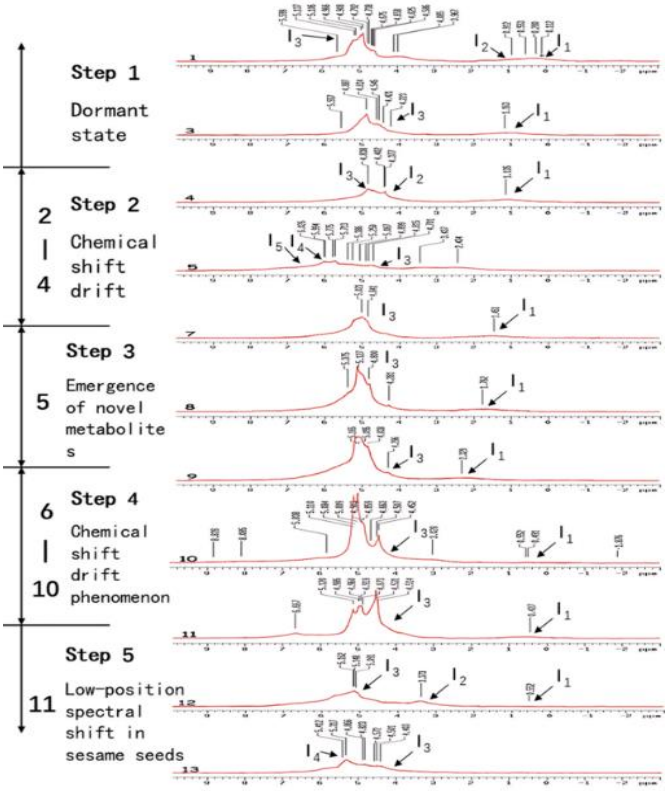


Figure 9- The 1D ^{1}H spectra of sesame seeds at different time points (Experiment 4)

A series of array images of sesame seed I_1 at the time point 4610 min obtained using CHESS are shown in Figure 9. From the figure, it was observed that the radicle of the sesame seed had completely penetrated the seed coat and endosperm, initiating germination. Metabolic signals were present in the future first pair of cotyledons. Seed metabolic activity is typically more active during this stage, including respiration, enzyme activity, and synthesis of metabolic products. Plant hormone synthesis and signalling are typically optimized, facilitating regulation of seed germination and growth processes (Castroverde & Dina 2021).

At the final time point of Experiment 4, 8650 min, sesame seeds exhibited clear signs of germination, such as visibly softened seed coats and the emergence of the embryo shoot extending from the seed coat. This phenomenon was distinct from Experiments 1, 2, and 3.

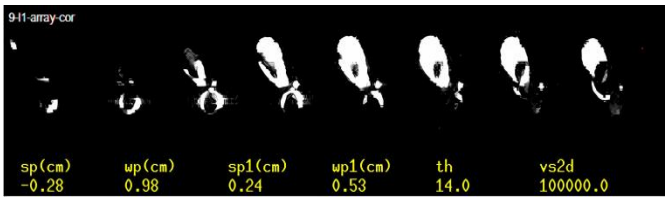


Figure10- The CHESS array images of sesame of seeds at 4160 min excited resonances at I_1. Imaging parameters were FOV = 1 × 1 cm, matrix size = 128 × 64, nt = 16, slice thickness = 0.3 mm

Based on the above experimental results, we found that sesame seeds only remain dormancy, both physiologically and metabolically, under conditions of low temperature, darkness, and low oxygen. In continuous light and aerobic environments, although metabolic reactions occur inside the seeds, they cannot germinate normally due to the lack of suitable temperature. Seeds can only develop properly when exposed to the right temperature. Yin et al. conducted a similar study on *Ottelia alismoides* seeds (Yin et al. 2013). Their study also found that an optimal temperature (25 to 30 °C) was favorable for seed germination. However, they observed that germination occurred only in light conditions, which is inconsistent with our findings. Sesame seeds are not photoblastic; they can germinate in both light and dark conditions. Additionally, our results are more comprehensive, as we found that even in light conditions, the seeds only escape metabolic dormancy but still do not germinate normally without suitable temperature. Although studies on the effects of oxygen, light, and temperature on seed germination are already well-established (Simon et al. 1976; Chahtane et al. 2017; An et al. 2020), these studies primarily rely on phenotypic data observed through naked eye or microscope. Such research aims to detect significant signs of germination in seeds. However, at the point when the seeds have recently exited dormancy and are approaching the germination stage, no apparent signs of germination are evident. Therefore, previous research on seed germination primarily focused on the later stages of observable germination. The findings indicated that temperature and oxygen were essential conditions for seed germination (Simon et al. 1976; Corbineau 2022), whereas light was not a prerequisite (Zhang et al. 2012). Our study provides molecular-level metabolic information of seeds transitioning from dormancy to early germination, a phase in which sesame growth changes are not visible to the naked eye. Through our method, a more precise analysis of sesame seed germination can be achieved. For example, we found that light exposure alone triggers sesame seeds to exit dormancy. However, previous studies (Kim 1983) indicates that sesame seeds are shown not to be photoreceptive, but this is not contradictory. In Experiment 2, although light triggered sesame seeds to exit dormancy and initiate metabolic activity, normal development did not occur.

Biological biochemical methods (Yan et al. 2020) such as enzyme activity analysis, proteomics, and gene expression profiling, are destructive and cannot provide continuous monitoring of seeds. Regarding NMR studies (Allen et al. 2009; Munz et al. 2017), these research efforts have used either NMR spectroscopy or MRI alone. In contrast, our study combines both methods. As a result, our approach provides a more comprehensive analysis and reveals both metabolic and physiological dynamics during seed germination.

This study has several limitations. First, it did not quantitatively discuss the effects of light, oxygen, and temperature on seed germination. Secondly, humidity is also an important factor influencing seed germination (Kauth et al. 2015; Li et al. 2020), Due to equipment limitations, we did not investigate the influence of humidity on seed dormancy release. Finally, applying the machine learning and computer vision models that we developed in other fields (Gao et al. 2024a; Gao et al. 2024b; Qiu et al. 2024; Xiao et al. 2024) to analyze the NMR and MRI data of seeds significantly enhances the practical applicability of our research.

4. Conclusions

This study used ^{1}H-NMR and CHESS to examine the variations in metabolite intensity and distribution during the transition of sesame seeds from dormancy to early germination under different light, oxygen, and temperature conditions. Under conditions of hypoxia (with seeds fully submerged in the culture medium), low temperature (15 °C in this study), and absence of light, NMR spectra and CHESS analysis revealed no changes in the intensity and distribution of metabolites, indicating that the seeds remained in a dormant state. Exposure to light alone or in combination with oxygen triggered dormancy release in sesame seeds, yet normal development did not occur; instead, metabolic activity occurred internally within the seeds, the embryo cannot

penetrate the seed coat. This indicates that although sesame seeds are non-photoblastic, light can still trigger metabolic activity within the seeds. Normal germination and metabolic activity were observed only when the appropriate temperature was applied (20 °C in this study). Therefore, to enhance the germination, temperature is a decisive factor. Conversely, for sesame seed storage, to prevent metabolic activity from depleting nutritional reserves, seeds should be maintained under low temperature, darkness, and hypoxic conditions. Future work will focus on analyzing and studying metabolites in sesame seeds under various temperatures, oxygen concentrations, light intensities, and humidity levels to enable precise control of sesame seed storage conditions, and establish mathematical models, thereby maximizing seed quality and germination rates.

Funding:

The Natural Science Foundation of Fujian Province of China No.2022J01821.
The Natural Science Foundation of Fujian Province of China No.2022J05163.
The National Natural Science Foundation of China No.11705068.
The National Key R&D Program of China 2023YFD2100603.

Data availability:

The data that support the findings of this study are available upon request.

Conflict of interest statement:

The authors declare no conflict of interest.

Permission to reproduce material from other sources:

All the materials in this manuscript are original.

References

Allen D K, Ohlrogge J B & Shachar-Hill Y (2009). The role of light in soybean seed filling metabolism. The Plant Journal 58 (2): 220-234. https://doi.org/10.1111/j.1365313X.2008.03771.x.

Amanah H Z, Joshi R, Masithoh R E, Choung M G, Kim K H, Kim G & Cho B K (2020). Nondestructive measurement of anthocyanin in intact soybean seed using Fourier Transform Near-Infrared (FT-NIR) and Fourier Transform Infrared (FT-IR) spectroscopy. Infrared Physics and Technology 111: 103477. https://doi.org/10.1016/j.infrared.2020.103477.

Bharti S K & Roy R (2012). Quantitative ^{1}H NMR spectroscopy. TrAC Trends in Analytical Chemistry 35: 5-26. https://doi.org/10.1016/j.trac.2012.02.007.

Blystone S, Nuixe M, Traoré A S, Cochard H, Picon-Cochard C & Pagés G (2024). Towards portable MRI in the plant sciences. Plant Methods 20 (1): 31. https://doi.org/10.1186/s13007-024-01152-z.

Boulc'h P N, Collewet G, Guillon B, Quellec S, Leport L & Musse M (2024). Quantitative MRI imaging of parenchyma and venation networks in Brassica napus leaves: effects of development and dehydration. Plant Methods 20(1): 69. https://doi.org/10.1186/s13007-024-01187-2.

Cai H, Cheng R, Jin Y, Ding S & Chen Z (2016a). Evaluation of oolong teas using ^{1}H and ^{13}C solid-state NMR, sensory analysis, and multivariate Statistics. Journal of the Chinese Chemical Society 63(9): 792-799. https://doi.org/10.1002/jccs.201600183.

Cai H, Lin L, Ding S, Cui X & Chen Z (2016b). Fast quantification of fatty acid profile of intact fish by intermolecular double-quantum coherence ^{1}H-NMR spectroscopy. European Journal of Lipid Science and Technology 118(6): 1150-1159. https://doi.org/10.1002/ejlt.201500309.

Cai H, Chuang W, Cui X, Cheng R, Chiu K, Chen Z & Ding S (2018a). High resolution ^{31}P NMR spectroscopy generates a quantitative evolution profile of phosphorous translocation in germinating sesame seed. Scientific Reports 8: 359. https://doi.org/10.1038/s41598-017-18722-y.

Cai H, Jin Y & Cui X (2018b). Feasibility of ultrafast high-resolution spectroscopy in the analysis of molecular-mobility-restricted samples in deuterium-free environments. Journal of the Chinese Chemical Society 65(6): 674-680. https://doi.org/10.1002/jccs.201700430.

Cao J, Li G, Qu D, Li X & Wang Y (2020). Into the seed: auxin controls seed development and grain yield. International Journal of Molecular Sciences 21 (5): 1662. https://doi.org/10.3390/ijms21051662.

Caprara C, Mathias T K, Santos M, D'Oca M G, D'Oca C D R, Roselet F, Abreu P C & Ramos D F (2023). Application of ^{1}H HR-MAS NMR-based metabolite fingerprinting of marine microalgae. Metabolites 13(2): 202. https://doi.org/10.3390/metabo13020202.

Carosio M G, Bernardes D F, Carvalho A & Colnago L A (2018). Non-invasive measurements of oilseed temperature in soil and soil thermal diffusivity using time-domain NMR relaxometry. Applied Magnetic Resonance 49(10): 1119-1127. https://doi.org/10.1007/s00723-018-1028-8.

Castroverde C D M & Dina D (2021). Temperature regulation of plant hormone signaling during stress and development. Journal of Experimental Botany 72(21): 7436-7458. https://doi.org/10.1093/jxb/erab257.

Chahtane H, Kim W & Lopez-Molina L (2017). Primary seed dormancy: A temporally multilayered riddle waiting to be unlocked. Journal of Experimental Botany 68(4): 857-869. https://doi.org/10.1093/jxb/erw377.

Chen D, Yuan H, Bao J, Zhao X, Fu X & Hu X (2023). Dry storage alters intraspecific variation in phenotypic traits at early life stages: evidence from a dominant alpine meadow species. Seed Science Research 33(2): 203-212. https://doi.org/10.1017/S0960258523000223.

Cheng Y, Chen H, Zhao Y, Cheng X, Wang L & Guo X (2023). Effect of light quality on polyphenol biosynthesis in three varieties of mung bean sprouts with different color seed coats. Plant Cell Reports 42(2): 253-268. https://doi.org/10.1007/s00299-022-02954-y.

Chiwocha S D, Abrams S R, Ambrose S J, Cutler A J, Loewen M, Ross A R & Kermode A R (2003). A method for profiling classes of plant hormones and their metabolites using liquid chromatography-electrospray ionization tandem mass spectrometry: an analysis of hormone regulation of thermos dormancy of lettuce (Lactuca sativa L.) seeds. The Plant Journal 35(3): 405-417. https://doi.org/10.1046/j.1365-313X.2003.01800.x.

Corbineau F (2022) Oxygen, a key signaling factor in the control of seed germination and dormancy. Seed Science Research 32(3): 126-136. https://doi.org/10.1017/S096025852200006X.

d'Avignon D A & Ge X (2018). In vivo NMR investigations of glyphosate influences on plant metabolism. Journal of Magnetic Resonance 292: 59-72. https://doi.org/10.1016/j.jmr.2018.03.008.

Dossa K, Diouf D, Wang L, Wei X, Zhang Y, Niang M, Fonceka D, Yu J, Mmadi M A & Yehouessi L W (2017). The emerging oilseed crop Sesamum indicum enters the "Omics" era. Frontiers in Plant Science 8: 1154. https://doi.org/10.3389/fpls.2017.01154.

Egli D & Wardlaw I (1980). Temperature response of seed growth characteristics of soybeans. Agronomy Journal 72(3): 560-564. https://doi.org/10.2134/agronj1980.00021962007200030036x.

Elboghdady A E A, Gomma A H, Hamed A M & Abdallatif A M (2023). Assessment of phenotypic diversity of some date palm male genotypes growing under Egyptian conditions. Revista Brasileira De Fruticultura 45: e-896. https://doi.org/10.1590/0100-29452023896.

Fujii H, Sato N, Kimura Y, Mizutani M, Kusama M, Sumitomo N, Chiba E, Shigemoto Y, Takao M & Takayama Y (2022). MR imaging detection of CNS lesions in tuberous sclerosis complex: The usefulness of T1WI with chemical shift selective images. American Journal of Neuroradiology 43(8): 1202-1209. https://doi.org/10.3174/ajnr.A7573.

Gao Z, Chen S, Huang J & Cai H (2024a). Real-time quantitative detection of hydrocolloid adulteration in meat based on Swin Transformer and smartphone. *Journal of Food Science* 89(7): 4359-4371. https://doi.org/10.1111/1750-3841.17159.

Gao Z, Huang J, Chen J, Shao T, Ni H & Cai H (2024b). Deep transfer learning-based computer vision for real-time harvest period classification and impurity detection of Porphyra haitnensis. Aquaculture International 32: 5171-5198. https://doi.org/10.1007/s10499-024-01422-6.

Gay C, Corbineau F & Côme D (1991). Effects of temperature and oxygen on seed germination and seedling growth in sunflower (Helianthus annuus L.). Environmental and Experimental Botany 31(2): 193-200. 10.1016/0098-8472(91)90070-5.

Gebeyehu B (2020). Review on: Effect of seed storage period and storage environment on seed quality. International Journal of Applied Agricultural Sciences 6(6): 185-190. https://doi.org/10.11648/J.IJAAS.20200606.14.

Geneve R L (2003). Impact of temperature on seed dormancy. HortScience 38(3): 336-340. https://doi.org/10.21273/HORTSCI.38.3.336.

Haase A, Frahm J, Hanicke W & Matthaei D (1985). ^{1}H NMR chemical shift selective (CHESS) imaging. Physics in Medicine and Biology 30(4): 341. https://doi.org/10.1088/0031-9155/30/4/008.

Jander G, Norris S R, Joshi V, Fraga M, Rugg A, Yu S, Li L & Last R L (2004). Application of a high-throughput HPLC-MS/MS assay to Arabidopsis mutant screening; evidence that threonine aldolase plays a role in seed nutritional quality. The Plant Journal 39(3): 465-475. https://doi.org/10.1111/j.1365-313X.2004.02140.x.

Kauth P J & Biber P D (2015). Moisture content, temperature, and relative humidity influence seed storage and subsequent survival and germination of Vallisneria americana seeds. Aquatic Botany 120: 297-303. https://doi.org/10.1016/j.aquabot.2014.09.009.

Kim C S (1983). Studies on the germination characteristics of sesame (Sesamum indicum L.). Korean Journal of Agricultural Science 10(1): 28-60

Kołodziejczyk I, Bałabusta M, Szewczyk R & Posmyk M M (2015). The levels of melatonin and its metabolites in conditioned corn (Zea mays L.) and cucumber (Cucumis sativus L.) seeds during storage. Acta Physiologiae Plantarum 37: 1-11. https://doi.org/10.1007/s11738-015-1850-7.

Liao W, Cai H & Ni H (2024). Sesame seed metabolism during germination under auxin: An in vivo NMR study. *Journal of Plant Growth Regulation*. https://doi.org/10.1007/s00344-024-11574-7

Li G, Chen M, Chen J, Shang Y, Lian X, Wang P, Lei H & Ma Q (2020). Chemical composition analysis of pomegranate seeds based on ultra-high-performance liquid chromatography coupled with quadrupole-Orbitrap high-resolution mass spectrometry. Journal of Pharmaceutical and Biomedical Analysis 187: 113357. https://doi.org/10.1016/j.jpba.2020.113357.

Li X, Simpson W, Song M, Bao G, Niu X, Zhang Z, Xu H, Liu X, Li Y & Li C (2020). Effects of seed moisture content and Epichloe endophyte on germination and physiology of Achnatherum inebrians. South African Journal of Botany 134: 407-414. https://doi.org/10.1016/j.sajb.2020.03.022.

Ma T, Tsuchikawa S & Inagaki T (2020). Rapid and non-destructive seed viability prediction using near-infrared hyperspectral imaging coupled with a deep learning approach. Computers and Electronics in Agriculture 177: 105683. https://doi.org/10.1016/j.compag.2020.105683.

Mota M F S, Waktola H D, Nolvachai Y & Marriott P J (2021). Gas chromatography–mass spectrometry for characterisation, assessment of quality and authentication of seed and vegetable oils. TrAC Trends in Analytical Chemistry 138: 116238. https://doi.org/10.1016/j.trac.2021.116238.

Munz E, Rolletschek H, Oeltze-Jafra S, Fuchs J, Guendel A, Neuberger T, Ortleb S, Jakob P M & Borisjuk L (2017). A functional imaging study of germinating oilseed rape seed. New Phytologist 216: 1181-1190. https://doi.org/10.1111/nph.14736.

Pasternak T, Pérez-Pérez J M, Ruperti B, Aleksandrova T & Palme K (2024). A new in vitro growth system for phenotypic characterization and seed propagation of arabidopsis thaliana. Journal of Plant Growth Regulation 43(2): 652-658. https://doi.org/10.1007/s00344-023-11093-x.

Qiu J, Lin Y, Wu J, Xiao Y, Cai H & Ni H (2024). Rapid beef quality detection using spectra pre-processing methods in electrical impedance spectroscopy and machine learning. International Journal of Food Science and Technology 59(3): 1624-1634. https://doi.org/10.1111/ijfs.16915.

Sarkar B K, Yang W, Wu Z, Tang H & Ding S (2009). Variations of water uptake, lipid consumption, and dynamics during the germination of Sesamum indicum seed: A nuclear magnetic resonance spectroscopic investigation. Journal of Agricultural and Food Chemistry 57(18): 8213-8219. https://doi.org/10.1021/jf9019129.

Simon E, Minchin A, McMenamin M M & Smith J (1976). The low temperature limit for seed germination. New Phytologist 77(2): 301-311. https://doi.org/10.1111/j.1469-8137.1976.tb01519.x.

Song L, Wang Q, Wang C, Lin Y, Yu D, Xu Z, Huang Q & Wu Y (2015). Effect of γ-irradiation on rice seed vigor assessed by near-infrared spectroscopy. Journal of Stored Products Research 62: 46-51. https://doi.org/10.1016/j.jspr.2015.03.009.

Tang W, Wu N, Xiao Q, Chen S, Gao P, He Y & Feng L (2023). Early detection of cotton verticillium wilt based on root magnetic resonance images. Frontiers in Plant Science 14: 1135718. https://doi.org/10.3389/fpls.2023.1135718.

Taylor A, Lee P & Zhang M (1999). Volatile compounds as indicators of seed quality and their influence on seed aging. Seed Technology 21

(1): 57-65
Terskikh V, Müller K, Kermode A R & Leubner-Metzger G (2011). In vivo ^{1}H-NMR microimaging during seed imbibition, germination, and early growth. Seed Dormancy: Methods and Protocols 773: 319-327. https://doi.org/10.1007/978-1-61779-231-1_18.
Tigabu M & Odén P C (2004). Rapid and non-destructive analysis of vigour of Pinus patula seeds using single seed near infrared transmittance spectra and multivariate analysis. Seed Science and Technology 32(2): 593-606. https://doi.org/10.15258/sst.2004.32.2.28.
Tuomainen T V, Toljamo A, Kokko H & Nissi M J (2024). Non-invasive assessment and visualization of Phytophthora cactorum infection in strawberry crowns using quantitative magnetic resonance imaging. Scientific Reports 14(1): 2129. https://doi.org/10.1038/s41598-024-52520-7.
Wu H, Deng J, Yang C, Zhang J, Zhang Q, Wang X, Yang F, Yang W & Liu J (2017). Metabolite profiling of isoflavones and anthocyanins in black soybean [Glycine max (L.) Merr.] seeds by HPLC-MS and geographical differentiation analysis in Southwest China. Analytical Methods 9(5): 792-802. https://doi.org/10.1039/C6AY02970A.
Xiao Y, Cai H & Ni H (2024). Identification of geographical origin and adulteration of Northeast China soybeans by mid-infrared spectroscopy and spectra augmentation. *Journal of Consumer Protection and Food Safety* 19(1): 99-111. https://doi.org/10.1007/s00003-023-01471-8.
Yan A & Chen Z (2020). The control of seed dormancy and germination by temperature, light and nitrate. The Botanical Review 86: 39-75. https://doi.org/10.1007/s12229-020-09220-4.
Xiao Y, Liao W, Cai H & Ni H (2024). Dynamic investigation of auxin on germination of sesamum indicum seed by ^{1}H spectroscopy and chemical shift selection Imaging. *Chiang Mai Journal of Science* 51: e2024033. https://doi.org/10.12982/CMJS.2024.033.
Yan S, Huang W, Gao J, Fu H & Liu J (2018). Comparative metabolomic analysis of seed metabolites associated with seed storability in rice (Oryza sativa L.) during natural aging. Plant Physiology and Biochemistry 127: 590-598. https://doi.org/10.1016/j.plaphy.2018.04.020.
Yin L Y, Zhang R F, Xie Z M, Wang C Y & Li W (2013). The effect of temperature, substrate, light, oxygen availability and burial depth on Ottelia alismoides seed germination. Aquatic Botany 111: 50-53. https://doi.org/10.1016/j.aquabot.2013.09.001.
Zhang M, Zhu J & Yan Q (2012). Review on influence mechanisms of light in seed germination. *Chinese Journal of Plant Ecology* 36(8): 899. https://doi.org/10.3724/SP.J.1258.2012.00899.
Zhang T, Fan S, Xiang Y, Zhang S, Wang J & Sun Q (2020). Non-destructive analysis of germination percentage, germination energy and simple vigour index on wheat seeds during storage by Vis/NIR and SWIR hyperspectral imaging. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 239: 118488. https://doi.org/10.1016/j.saa.2020.118488.

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Daily reference evapotranspiration prediction using empirical and data-driven approaches: A case study of Adana plain

Deniz Levent Koç[a*], Semin Topaloğlu Paksoy[b]

[a]*Department of Agricultural Structures and Irrigation, Agriculture Faculty, Çukurova University, 01250, Sarıçam, Adana, TÜRKIYE*
[b]*Department of Econometrics, Economics and Administrative Sciences Faculty, Çukurova University, 01250, Sarıçam, Adana, TÜRKIYE*

ARTICLE INFO
Research Article
Corresponding Author: Deniz Levent Koç, E-mail: leventk@cu.edu.tr
Received: 09 May 2024 / Revised: 12 September 2024 / Accepted: 30 September 2024 / Online: 14 January 2025

Cite this article
Koç D L, Topaloğlu Paksoy S (2025). Daily reference evapotranspiration prediction using empirical and data-driven approaches: A case study of Adana plain. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):207-229. DOI: 10.15832/ankutbd.1481207

ABSTRACT
Precise determination of the reference evapotranspiration (ET_0) is vital to studying the hydrological cycle. In addition, it plays a significant role in properly managing and allocating water resources in agriculture. The objective of this research was to examine the effectiveness of five different data-driven techniques, including artificial neural networks "multilayer perceptron" (ANN), gene expression programming (GEP), random forest (RF), support vector machine "radial basis function" (SVM), and multiple linear regression (MLR) to model the daily ET_0. These methods were also compared with Hargreaves-Samani (HS), Oudin, Ritchie, Makkink (MAK), and Jensen Haise (JH) empirical models and their calibrated versions. The empirical models JH and MAK performed better than the models HS and Oudin after being calibrated by linear regression. All data-driven methods with four inputs were superior to the original and calibrated empirical models. Generally, data-driven models provided increased accuracy and enhanced generalization in predicting daily reference evapotranspiration compared to empirical models. The RF and ANN methods generally demonstrated better estimation accuracy than other data-driven methods. The performance of the RF and ANN models that utilized T_{max}, T_{min}, and R_s inputs, as well as those that incorporated T_{max}, T_{min}, R_s, and U_2 inputs, proved to be superior to their corresponding MLR-based and GEP-based models for predicting ET_0 in the Adana plain, which is characterized by a Mediterranean climate. Nevertheless, the GEP and MLR methods have the advantage of utilizing explicit algebraic equations, making them more convenient to apply, especially in the context of agricultural irrigation practices.

Keywords: Reference evapotranspiration, Data-driven approaches, Empirical models, Calibration, Adana plain

1. Introduction

Evapotranspiration is vital in maintaining the land's water and energy balance, significantly influencing water resource management, irrigation, environmental studies, and hydrological systems (Sabziparvar & Tabari 2010; Izadifar 2010). Various methods, including high-cost micrometeorological techniques, remote sensing, and water budget measurements, can be utilized to measure evapotranspiration. However, these approaches have inherent limitations, such as high expenses, maintenance requirements, and complexity (Liu & Zhu 2018; Niaghi et al. 2021).

Using mathematical models based on measured meteorological parameters is a cost-effective way to estimate reference evapotranspiration (ET_0) (Negm et al. 2018). Crop evapotranspiration (ET_c) can be calculated by multiplying the crop-specific coefficient known as K_c by ET_0 The FAO56 Penman-Monteith (PM) equation is a recognized method for estimating ET_0 (Allen et al. 1998) and can be used in diverse environments and climate conditions (Landeras et al. 2008; Shiri et al. 2012). However, the most crucial drawback of the FAO56-PM equation is that it requires air temperatures (T), solar radiation (R_s), relative humidity (RH), and wind speed (U) data. Many models have been developed to estimate ET_0 using reduced climate data due to incomplete meteorological data in some areas (Tabari et al. 2013a). Empirical methods often necessitate less data and fewer meteorological parameters, making them particularly applicable in agricultural settings for farmers and water managers. However, it is essential to note that the accuracy of empirical ET_0 methods can vary significantly due to differences in data requirements and theoretical assumptions (Dong et al. 2024). These methods are often tailored to specific geographical areas and local weather patterns, necessitating local calibration for optimal performance, unlike the standardized FAO56-PM equation (Gao et al. 2015; Pereira et al. 2015; Ferreira et al. 2019). HS, Oudin, JH, Ritchie, and MAK models are among the commonly used models (Tabari et al. 2012; Ferreira et al. 2019). FAO recommends that the HS equation be applied to calculate ET_0 in case of the absence of R_s, RH, and U data, which are crucial inputs of the FAO56-PM model (Allen et al. 1998). To date, several studies have been conducted aiming to calibrate the HS, Ritchie, MAK, Oudin, and JH models (Citakoglu et al. 2014; Almorox

& Grieser 2016; Feng et al. 2016; Feng et al. 2017; Shiri 2017; Cobaner et al. 2017; Ferreira et al. 2018; Banda et al. 2018; Gomariz Castillo et al. 2018; Srivastava et al., 2018; Ferreira et al. 2019; Khodayvandie et al. 2022). These papers highly recommended a regional or local calibration of the empirical equations used in the studies.

Another option for predicting ET_0 involves utilizing data-driven or soft computing methods such as ANN, GEP, MLR, SVM, and RF. There has been a significant emphasis on using these methods to estimate ET_0 in the past few years (Shiri 2018; Dou & Yang 2018; Gavili et al. 2018; Abrishami et al. 2019; Mattar & Alazba 2019; Reis et al. 2019; Yirga 2019; Ferreira et al. 2019; Wang et al. 2019; Mohsin & Lone, 2021; Jang et al. 2021; Niaghi et al. 2021; Achite et al. 2022; Dimitriadou & Nikolakopoulos 2022; Wang et al. 2022; Bayram & Çıtakoğlu 2023).

Data-driven approaches provide several advantages, including simplified development processes compared to physically based models. These models are independent of underlying boundary conditions, other assumptions, or initial forcing and can effectively operate at localized positions (Prasad et al. 2017). However, data-driven models require large amounts of high-quality data and can become overfitted to the training data. Developing and implementing data-driven models, particularly deep learning models, may necessitate substantial computational resources and time.

However, despite all the disadvantages, data-driven approaches have become increasingly popular for predicting ET_0, leading water resource and irrigation engineering experts to employ these methods in different applications. Huo et al. (2012) compared the performances of ANN models with those of MLR, Penman, PT and HS models in an arid area of northwest China. The results showed that ANN models estimated ET_0 more accurately than other models. Yassin et al. (2016) conducted a comparative analysis between ANN and GEP models to estimate potential evapotranspiration (ET_0) in arid environments. Based on the findings, it was observed that ANN-based models exhibited a marginally higher level of precision than GEP-based models. Antonopoulos & Antonopoulos (2017) used ANN, PT, MAK, HS and mass-transfer models to predict ET_0 in West Macedonia of northern Greece. In the study, the ANN model with four input variables estimated daily ET_0 most accurately. In Guangxi, which is situated in the southwest of China's Pearl River basin, Wang et al. (2019) utilized RF-based and GEP-based models to estimate ET_0. The findings concluded that the RF-based ET_0 models outperformed the GEP-based ones, albeit by a small margin. Mattar & Alazba (2019) modelled ET_0 from various combinations of climatic variables using GEP and MLR techniques in Egypt, where climatic conditions changed from warm to temperate. The study's findings indicated that GEP models outperformed MLR, HS, and MAK models. In the semiarid region of Brazil, Reis et al. (2019) conducted a study comparing the performance of empirical equations with those of ANN, ELM (extreme learning machine), and MLR models. The results showed that ANN, ELM, and MLR models had similar accuracy and were also more accurate than HS and calibrated HS models. Üneş et al. (2020) compared various models for modeling daily ET_0 and found that data-driven techniques, such as ANFIS and radial basis function SVM, outperformed empirical equations. ANFIS showed the highest correlation coefficient while radial basis function SVM had the minimum errors. Turc empirical formula was found to be better than other empirical equations. Chen et al. (2020) evaluated the performance of deep learning methods for ET_0 estimation in the Northeast Plain, China, using incomplete meteorological data. Three deep learning models, deep neural network (DNN), temporal convolution neural network (TCN), and long short-term memory neural network (LSTM), were developed and compared with classical machine learning models (SVM, RF) and empirical equations. The results showed that the deep learning models outperformed the empirical models, especially when temperature-based features were available. Kaya et al. (2021) found that support vector regression (SVR), multi-layer perceptron (MLP), and multiple linear regression (MLR) models outperformed commonly used empirical formulas in estimating daily ET_0 in the Košice City area of eastern Slovakia. The results from the Hargreaves-Samani equation closely aligned with the FAO 56-PM equation, demonstrating superior performance compared to other empirical equations. Niaghi et al. (2021) conducted a study to evaluate several machine learning techniques, comprising GEP, SVM, MLR, and RF, using three different input combinations to estimate ET_0 in the Red River Valley in the USA. The findings indicated that the RF model was the most effective approach for all input combinations among the four models considered. In Algeria's semiarid region, Achite et al. (2022) conducted a study to explore the potential of ANN and GEP models to estimate ET_0 by utilizing various combinations of climatic variables. The results showed that modelling ET_0 utilizing the ANN technique gave better estimates than the GEP models.

Sarıgöl & Katipoğlu (2024) evaluated the effectiveness of hybrid machine learning models for predicting monthly evaporation in the Southeast Anatolia Project Area. The research found that combining the gradient boosting machines (GBM) technique with signal decomposition methods generally provided more accurate evaporation estimations than using the GBM model alone.

As the reviewed literature shows, the studies of combined application of data-driven techniques to estimate ET_0 and applications of GEP, RF, MLR, and SVM approaches for modelling ET_0 are minimal, especially the RF method was rarely applied in hydrological and irrigation research. Also, in Türkiye, very few studies applied data-driven methods to estimate ET_0 and used the calibration of empirical models to estimate ET_0. As far as we know, only three studies have been carried out to adjust empirical models (Citakoglu et al. 2014; Çobaner et al. 2017; Gharehbaghi & Kaya 2022) and no studies have also addressed evaluating the GEP models to estimate daily ET_0 and developing the models' mathematical expressions in Türkiye. Other researchers preferred to compare the outcomes of the original empirical models solely. One of the unique features of this paper is the calibration of five widely used empirical models for estimating ET_0 and comparing them with data-driven models.

Forecasting reference evapotranspiration accurately is crucial for optimizing agricultural production and effective water resource management, which highlights the importance of addressing gaps mentioned above in the literature through this study.

2. Material and Methods

2.1. Study area and data used

The Adana Plain, situated in the Eastern Mediterranean region of Türkiye, was the focus of the study. Adana's climatic conditions are categorized under the Csa per the Köppen-Geiger classification system. This type of climate is known for its mild winters and extremely hot, dry summers. It is considered a typical Mediterranean climate (TSMS 2024). Adana Plain covers 27% of the province's territory. The basin's southern part is called Çukurova, and the northern part is called Anavarza (CSB 2023). Turkey's leading producer of citrus fruits, watermelons, soybeans, and peanuts is the Adana Plain (Kades 2019). Irrigation is necessary for crop production in Adana Plain due to inadequate precipitation levels and distribution during the growing season. The study's timeframe is from April through October, encompassing the growth season of the main crops in Adana Plain. This study used long-term daily climate data (2000-2021). The data was gathered from the Adana weather station in the Adana Plain (TSMS 2022) (Figure 1). According to the long-term climate data (1929-2022), the region experiences an annual precipitation of 668.8 mm, and approximately 50% of this precipitation falls during the winter months of December, January, and February. Temperatures are highest in July and August, with an average daily temperature of 28.2 and 28.7 °C. (TSMS 2023).

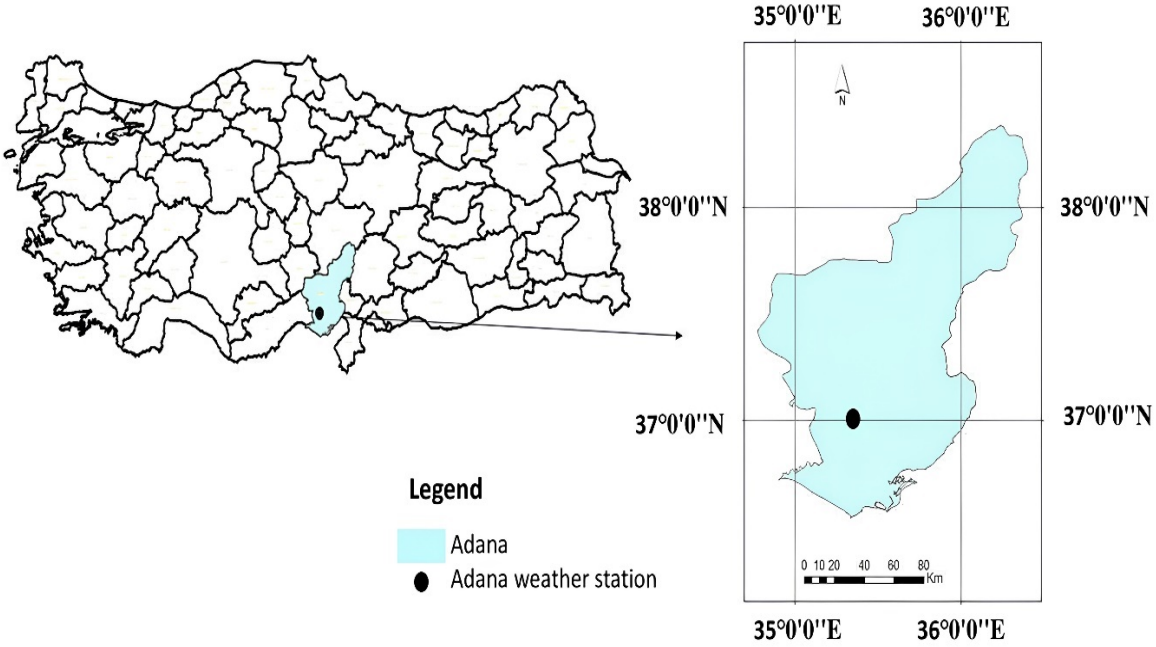


Figure 1- Position of the Adana climate station in Adana Plain in Türkiye (37 00' 14" N; 35 20' 39" E; altitude: 24 m)

The statistics regarding the data used can be found in Table 1. The average highest temperature (T_{max}) and lowest temperature (T_{min}) during the training period are 30.8 and 19.5 ºC, respectively. The mean relative humidity (RH) is 64.7%, solar radiation (R_s) is 19.6 MJ m^{-2} d^{-1} and wind speed at a two-meter height (U_2) is 0.92 m s^{-1} during the training period. The RH has a skewness coefficient of -0.94, while the U_2 has a skewness coefficient 0.41. The U_2 has a kurtosis coefficient 1.34, and T_{min} kurtosis coefficient of -0.56. In the validation period, the average values for T_{max}, T_{min}, RH, and R_s are 31.5 ºC, 20.6 ºC, 63.8 %, and 20.8 MJ m^{-2} d^{-1}, respectively. However, the average U_2 during the validation period is notably higher at 1.33 m s^{-1}. The coefficient of variation for U_2 is higher in the validation period, indicating more variability in wind speed during this period. The coefficients of variation of other climate parameters are fairly similar. The skewness coefficients of U_2 differ, while the other parameters are similar in both the training and validation periods. The kurtosis coefficients of T_{max} and U_2 differ, while the other parameters are similar in both the training and validation periods. The mean ET_0 is slightly higher in the validation period. However, the coefficient of variation is the same in both periods, indicating similar variability in ET_0 in both periods.

Table 1- Key statistical information of the dataset utilized in the study during the training and validation periods

Period	*Statistic*	T_{max} (°C)	T_{min} (°C)	*RH* (%)	R_s (MJ m^{-2} d^{-1})	U_2 (m s^{-1})	ET_0 (mm d^{-1})
Training	Maximum	42.1	29.8	94.5	30.4	3.07	8.9
	Minimum	15.0	3.5	16.0	0.8	0.10	0.9
	Mean	30.8	19.5	64.7	19.6	0.92	4.2
	Standard Deviation	4.64	4.80	11.23	5.11	0.39	1.17
	Skewness	-0.67	-0.42	-0.94	-0.60	0.41	-0.19
	Kurtosis	-0.09	-0.56	1.02	0.23	1.34	-0.38
	Coefficient of Variation	0.15	0.25	0.17	0.26	0.43	0.28
Validation	Maximum	45.1	29.8	87.0	30.5	3.96	10.5
	Minimum	16.5	6.6	24.0	0.3	0.45	1.1
	Mean	31.5	20.6	63.8	20.8	1.33	4.7
	Standard Deviation	4.57	4.65	10.21	5.58	0.44	1.31
	Skewness	-0.72	-0.52	-1.08	-0.84	1.58	-0.21
	Kurtosis	0.43	-0.53	0.94	0.44	4.99	-0.31
	Coefficient of Variation	0.15	0.23	0.16	0.27	0.33	0.28

Notes: ET_0 represents the ET_0 values estimated by the FAO56-PM equation

2.2. Empirical models for ET_0 prediction

This study employed the FAO56-PM ET_0 model as the benchmark for assessing the empirical and data-driven models, a common process in literature (Equation 1).

$$ET_0 = \frac{0.408\text{SVPC}(R_n - G) + \gamma \frac{900}{T + 273} U_2(e_{sat} - e_{act})}{\text{SVPC} + {}_{\varphi}(1 + 0.34U_2)} \quad (1)$$

ET_0 refers to the reference evapotranspiration in mm d^{-1}, R_n represents the net radiation in MJ m^{-2} d^{-1}, G denotes the soil heat flux density in MJ m^{-2} d^{-1}, T signifies the mean daily air temperature at the height of 2 meters in °C, U_2 represents the wind speed at a height of 2 meters in m s^{-1}, e_{sat} refers to the saturation vapour pressure in kPa, e_{act} denotes the actual vapour pressure in kPa, SVPC represents the slope vapour pressure curve in kPa $°C^{-1}$, and φ signifies the psychrometric constant in kPa $°C^{-1}$.

All parameters' daily values were calculated using the equations outlined in the FAO Irrigation and Drainage Paper 56 written by Allen et al. (1998).

Five empirical models were employed in this study, and concise descriptions of the methodologies utilized are presented below. The equations corresponding to each model can be found in Table 2.

2.2.1. Hargreaves-Samani model

Hargreaves and Samani (1985) introduced an equation for estimating ET_0 based on daily or mean values of maximum and minimum temperature. The equation was formulated using eight years of daily lysimeter data collected in Davis, California, and subsequently applied to compute ET_0 values for various locations.

2.2.2. Oudin model

In 2005, Oudin et al. suggested that the Penman-Monteith equation may not be the best choice for use in rainfall-runoff models. Instead, they put forward the Oudin method as an alternative.

2.2.3. Ritchie model

Due to a lack of meteorological data, Jones & Ritchie (1990) devised a simpler formula that correlates ET_0 with only air temperature, known as the Ritchie equation. The Ritchie equation finds extensive application in plant growth models and in research on managing agricultural water.

2.2.4. Makkink model

Makkink (1957) developed the MAK method based on the Penman-Monteith equation, and its validation was conducted using data from a lysimeter study on short grass in the Netherlands.

2.2.5. Jensen-Haise model

Jensen & Haise (1963) introduced the JH method, which utilizes global solar radiation and air temperature to develop irrigation plans based on comprehensive field data in arid and semi-arid conditions in the USA.

Allen et al. (1998) recommended calibrating empirical models locally by determining regression coefficients (a, b) using the FAO56-PM model as follows (Equation 2).

$$ET_0(PM) = a + bET_{0_{model}} \quad (2)$$

ET_0 (PM): the reference evapotranspiration computed by the FAO56-PM, ET_{0model}: the reference evapotranspiration estimated by other applied models, and a and b: regression coefficients.

Table 2- Empirical models to estimate ET_0

Model	***Reference***	***Meteorological inputs***	***Formula***
Hargreaves-Samani (HS)	Hargreaves & Samani (1985)	T_{max}, T_{min}	$ET_0 = 0.0023Ra(T_{max} - T_{min})^{0.5}(T + 17.8)$
Oudin	Oudin et al. (2005)	T_{max}, T_{min}	$ET_0 = \frac{R_a}{\lambda}(\frac{(T+5)}{100}\ If\ T + 5 > 0;\ ETo = 0\ otherwise$
Ritchie	Jones & Ritchie (1990)	T_{max}, T_{min}, R_s	$ET_0 = \alpha_1[3.87 \times 10^{-3} \times R_s(0.6T_{max} + 0.4T_{min} + 29)]$ $\alpha_1 = 1.1\ when\ 5°C < T_{max} < 35°C$ $\alpha_1 = 1.1 + 0.05(T_{max} - 35)\ when\ T_{max} > 35°C$ $\alpha_1 = 0.1exp[0.18(T_{max} + 20)]\ when\ T_{max} < 5°C$
Makkink (MAK)	Makkink (1957)	T_{max}, T_{min}, R_s	$ET_0 = 0.61\frac{SVPC}{SVPC + \gamma}\frac{R_s}{\lambda} - 0.12$
Jensen-Haise (JH)	Jensen & Haise (1963)	T_{max}, T_{min}, R_s	$ET_0 = 0.408R_s(0.0252T + 0.078)$

Notes: T= mean daily air temperature (°C), R_a = extraterrestrial radiation (MJ m^{-2} d^{-1}), λ = the latent heat flux (MJ kg^{-1}), α_1= coefficient, SVPC =slope vapour pressure curve (kPa $°C^{-1}$), γ = psychrometric constant (0.0672 kPa K^{-1})

2.3. Data-driven methods for ET_0 prediction

As per the NFLT (No Free Lunch Theorem), all optimization methods perform equally well on average. Hence, there is no agreement on whether a machine learning method can always provide better results compared to the others. Several aspects, like the size and structure of the dataset employed in the study, play a crucial role. Therefore, multiple methods are utilized and compared to obtain the most valuable or best prediction (Sterkenburg & Grünwald; 2021; Goldblum et al. 2023). The current study examines GEP, RF, ANN, MLR, and SVM models to predict the daily ET_0 using observed daily meteorological data set (T_{max}, T_{min}, U_2, and R_s) and FAO56-PM-targeted ET_0 values. For this purpose, as common in the literature (Irmak et al. 2003; Noi et al. 2017), this study employed 70% of the dataset (from April 1, 2000, to October 31, 2014) for training and calibration and the remaining 30% (from April 1, 2015, to October 31, 2021) for validating the data-driven and empirical models.

A correlation matrix was constructed to establish the connection between FAO56-PM ET_0 and several climate parameters (Figure 2). RH was found to have a very weak negative correlation (R = -0.165) with FAO56-PM ET_0. Therefore, RH was excluded when deciding on input combinations for the models. In this study, a total of four input combinations were devised to assess the effectiveness of data-driven models. Two of these combinations were the same as the input combination of the empirical equations utilized in the study: 1) T_{max} and T_{min} and 2) T_{max}, T_{min}, and R_s. The remaining combinations were defined as 3) T_{max}, T_{min}, and U_2 and 4) T_{max}, T_{min}, R_s, and U_2. Figure 3 provides a summary of the study's workflow.

Figure 2- Correlation between FAO56-PM ET_0 and various climate parameters

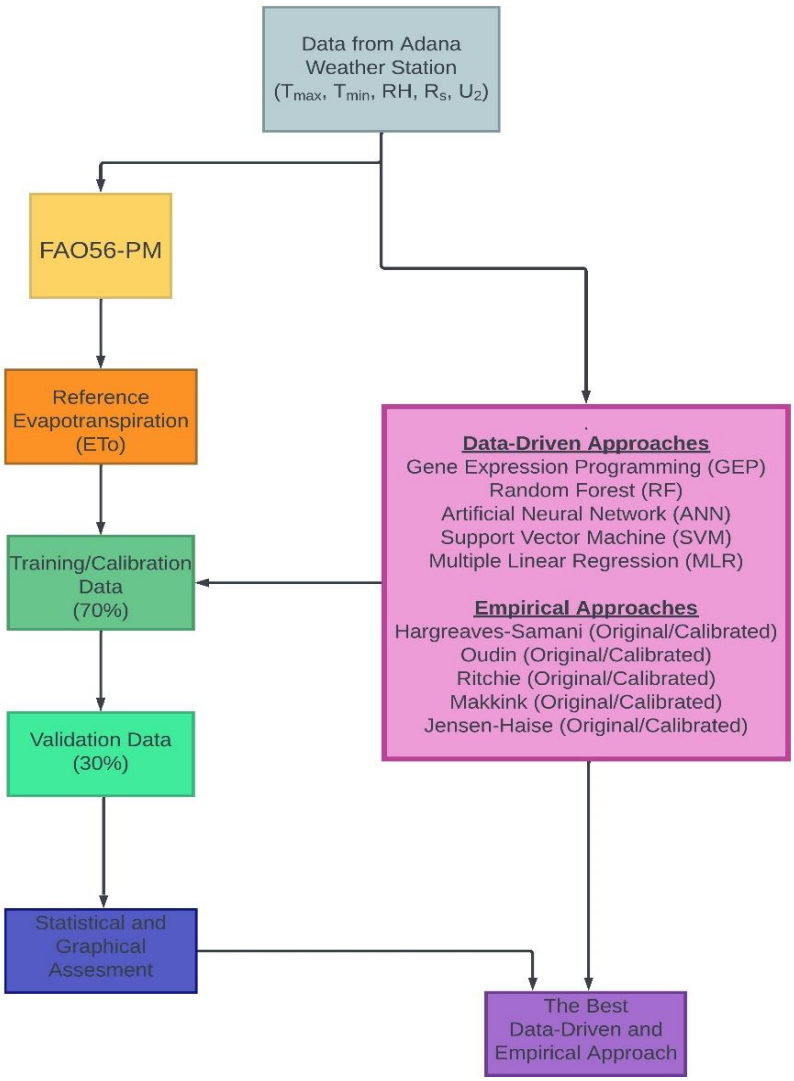


Figure 3- Flowchart outlining the research methodology used in the study

2.3.1. Gene expression programming (GEP)

GEP, invented by Ferreira in 2001, is a branch of evolutionary algorithms that models dynamic and non-linear processes. It is an advanced form of genetic algorithms and programming (Mehdizadeh et al. 2017). The sensitivity of GEP to the number of inputs is lower than the impact of the information content present in the data (Shiri, 2017). The GEP algorithm defines the solution to a problem using chromosomes that contain one or multiple genes. These genes can be combined to create chromosomes, which are then linked through a linking function (Traore & Guven 2013). Algebraic expressions for non-linear problems can be automatically generated by GEP. Moreover, GEP models have the ability to explicitly establish the connection between dependent and independent variables.

The first step in designing a GEP algorithm is to define the fitness function. The next step is to define the terminals and functions to be used. The third step is to determine the structure of chromosomes, including the number of generations, length, and number of genes. The fourth step is to determine the linking function. The fifth step is to determine the characteristics of the operators. Finally, the algorithm can be implemented. After running the program with different input combinations, the GEP model was developed once the model's accuracy stopped significantly improving (Wang et al. 2019).

2.3.2. Artificial neural networks (ANN)

ANNs can be described as mathematical models resembling biological neural networks. ANNs can learn from examples, identify patterns within the data, and adapt their solutions over time, making them efficient at processing large amounts of information (Jain et al. 2008). An ANN's internal architecture resembles a natural brain's structure, with several layers of fully interconnected nodes or neurons. A neural network is defined by its architecture, training algorithm, and activation function. The most common neural network architecture consists of three layers: the input layer for introducing data, one or more hidden layers for processing information, and the output layer for obtaining results. The type of ANN that we are referring to is known as a multilayer perceptron (MLP) (Fausett et al. 2006). The Backpropagation (BP) algorithm is used by the multilayer perceptron (MLP), which is the most widely used, persuasive, and effective neural network architecture (Choi et al. 2018). Kumar et al. (2011) provide further information about ANN. MLP neural networks with various combinations of climatic inputs and hidden layers were evaluated in this research to determine their usefulness in estimating ET_0.

2.3.3. Multiple linear regression (MLR)

Despite significant developments in data-driven modelling, using MLR for various modelling and model comparison purposes remains in demand (Izadifar 2010). MLR, a statistical modelling technique, uses multiple explanatory variables to anticipate the result of a response variable. MLR employs multiple explanatory variables, typically two or more, to approximate the result of a response variable by using a linear equation to establish a fitting linear model. The MLR model has a general form that can be expressed as follows (Equation 3):

$$Y = a_0 + a_1X_1 + a_2X_2 + \ldots + a_nX_n \quad (3)$$

Where Y: the predicted value of the dependent variable; $X_1,\ldots X_n$: independent variables; a_0: unknown intercept; $a_1,\ldots a_n$: estimated regression coefficients of the function.

2.3.4. Support vector machine (SVM)

The support vector machine (SVM) is a relatively new soft learning algorithm adopted for multiple applications in fields such as soft computing, hydrology, and environmental studies (Gocić et al. 2015). SVM is a well-known method in the field of machine learning that is based on the principles of classification and regression analysis theory. Its origin can be traced back to its developer, Vapnik (1995). Support vector regression (SVR) is commonly employed to elucidate the concept of regression in the SVM algorithm.

SVM uses the equation f(x) in regression analysis to establish the connection between a set of independent variables (x) and a dependent variable (y). This equation defines how the independent variables relate to the dependent variable (Equation 4).

$$f(x) = \hat{w}\cdot\phi(x) + þ \quad (4)$$

The weight vector is represented as 'ŵ'. '(x)' is a non-linear function that transforms the input space vector' x' into a high-dimensional feature space. The bias term is denoted as 'þ'.

SVM employs kernel functions such as linear (Lin), polynomial (Poly), and radial basis function (RBF) to transform input data into a feature space of high dimensionality. This study used the radial basis (RBF) as the kernel function since it is highly recommended in the literature (Tabari et al. 2012; Seifi & Riahi 2020). To achieve satisfactory performance, it is crucial to

appropriately set parameters C, ε, and γ when using the RBF kernel to train an SVM model. For this study, the parameters were chosen through a process of trial and error.

2.3.5. Random forest (RF)

The Random Forest (RF) model has recently become increasingly popular as an ensemble learning technique for classification and regression tasks. Its various advantages make it a preferred choice in classification and regression problems. These benefits include preventing overfitting, providing satisfactory performance, and enabling personalized parameter selection (Feng et al. 2017). Breiman (2001) first introduced random forest (RF), a machine-learning method that utilizes several decision trees to form an ensemble model. Ensemble learning is a technique that seeks to enhance the overall ability of machine-learning models to generalize by creating multiple base learners or combining multiple trees in their structure (Samadianfard et al. 2022). The RF method is one of the most effective and practical approaches to generating rules. This method is based on decision tree algorithms such as Classification and Regression Tree (CART). Compared to other decision tree ensembles, RF is known to be a more robust approach (Cutler et al. 2012). The Random Forest (RF) method can define the appropriate predictor without requiring the data to be re-scaled like other techniques. Conversely, regression trees that follow the traditional approach tend to over-fit on the training data set, which results in poor performance. Nevertheless, the RF method utilizes the characteristics of randomness to conquer this issue (Shirzad & Safari 2019).

Breiaman (2001) and Cutler et al. (2012) contain further information on RF. The current study utilized RF as the regression model for estimating ET_0. Parameter optimization was carried out with a focus on the number of trees to achieve the study's lowest possible error level.

2.4. Software and pseudo codes

The models in this study were developed using GeneXProTools 5.0 and the Orange 3.35 software suite. Due to limitations in the existing programs, the visual components were developed using the Python programming language (Phyton 3.10.12). The study utilized the following libraries: Matplotlib 3.7.1, Numpy 1.26.4, Pandas 2.1.4, and Seaborn 0.13.1. Applied model parameters and pseudo codes are provided in Appendix 1. Regarding the benefits of softwares, the following points can be articulated. Orange software is a user-friendly platform for developing data analysis and machine learning models. It has a drag-and-drop interface, making model building intuitive and not reliant on extensive technical expertise (Orange Data Mining 2024). GeneXProTools is a robust software application for modeling and analyzing data using genetic programming. It excels in automated exploration and optimization of complex mathematical models and is well-suited for data mining, modeling, optimization, and classification (Ferreira 2001).

2.5. Assessment of model performance

Table 3 demonstrates the use of multiple statistical indicators to evaluate the precision of the model's predictions (Willmott 1981; Karunanithi et al. 1994; Jacovides & Kontoyiannis 1995; Landeras et al. 2008). The current investigation assessed the performance of the models by employing various metrics such as RMSE, MAE, RaRMSE, RaMAE, and R, and applying criteria recommended by Corzo & Solomatine (2007) based on the RE values. As per criteria, an RE value of 15% or less indicates a small error, whereas an RE value between 15% and 35% indicates a moderate error, and finally, an RE value greater than or equal to 35% is considered a large error.

In the study, the conformance between the ET_0 derived from the FAO56-PM reference method and the modeled ET_0 values was assessed using a Taylor diagram, employing the Pearson correlation coefficient (R), RMSE, and standard deviation statistics. This method facilitated the identification of the most realistic model. Moreover, the research employed box plots to visually represent the distribution of numerical data. These box plots illustrate the central tendency (median), variability (interquartile range), and potential skewness through quartiles. Additionally, they help in identifying outliers and comparing the overall spread of different data sets (Mcgill et al. 1978; Taylor 2001).

Table 3- The statistical parameters employed to assess the performances of the models

Statistical indices	*Symbol*	*Equation*
Root mean square error	RMSE	$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2}$
Root mean square error ratio	RaRMSE	$\mathrm{RaRMSE} = 1 - \frac{\mathrm{RMSE}_{\text{calibrated model}}}{\mathrm{RMSE}_{\text{non-calibrated model}}}$
Relative error (%)	RE	$\mathrm{RE} = \frac{\mathrm{RMSE}}{\bar{O}} \times 100$
Mean absolute error	MAE	$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\lvert P_i - O_i \rvert$
Mean absolute error ratio	RaMAE	$\mathrm{RaMAE} = 1 - \frac{\mathrm{MAE}_{\text{calibrated model}}}{\mathrm{MAE}_{\text{non-calibrated model}}}$

Notes: P_i represents the estimated ET_0 values by the models, measured in mm d^{-1}; O_i represents the FAO56-PM ET_0 values, also measured in mm d^{-1}; n represents the total number of data; and $\bar{O}$ represents the mean of FAO56-PM ET_0 values, measured in mm d^{-1}.

3. Results and Discussion

3.1 Empirical models

Table 4 displays the statistical outcomes of both the original and calibrated empirical models that have been taken into account for this study. The linear regression parameters used for calibrating empirical models are presented in Table 5. The JH model (non-calibrated) demonstrated the poorest performance with the values of RMSE of 1.845 mm d^{-1}, MAE of 1.658 mm d^{-1}, and RE of 39.1% in the validation period. According to criteria suggested by Corzo & Solomatine (2007), the JH model showed high errors due to RE ≥ 35%. In contrast, during the validation period, the HS model (non-calibrated) demonstrated the best performance with an RMSE value of 0.742 mm d^{-1}, an MAE value of 0.548 mm d^{-1}, and a RE value of 15.7%, and this model showed medium due to RE values. Similarly, Sabziparvar & Tabari (2010) also showed that HS and Turc models estimated ET_0 more accurately than MAK and Priestley-Taylor (PT) models in semiarid climates. It can be argued that the varying efficacy of empirical approaches in estimating ET_0 stems from the fact that each method is tailored to distinct geographic areas and specific local climatic conditions. Initially developed in the United States for implementation in arid and semi-arid conditions, the JH equation based on T_{max}, T_{min}, and R_s may not provide accurate estimates for regions under different climatic influences. The HS equation based on T_{max} and T_{min}, endorsed globally by the Food and Agriculture Organization (FAO), yielded superior results compared to other empirical equations in our study. The Makkink equation based on T_{max}, T_{min}, and R_s, originally developed under the Netherlands' mild and humid climate conditions as an adaptation of the Penman-Monteith equation, exhibited limited efficacy in estimating ET_0 within our research area, characterized by a Mediterranean climate. The Oudin equation, tested in French basins, demonstrated comparable predictive performance to the HS model in estimating ET_0 within this research. The Ritchie equation, which incorporates coefficients adjusted for specific temperature ranges, did not yield satisfactory results in estimating ET_0.

This disparity suggests the need for region-specific adjustments to improve predictive accuracy. In this instance, it is necessary to ascertain calibration coefficients tailored to the specific local climatic conditions or to calibrate equations using linear regression as recommended by Allen et al. in 1998.

In the study, after calibration by linear regression, the ET_0 estimation performances of Jensen-Haise (JH) and Makkink (MAK) models significantly increased with a sharp decline in RMSE, MAE and RE values, as shown in Table 4. These models strongly correlated with the FAO56-PM method (R = 0.94 for JH; R = 0.93 for MAK, as shown in Table 5) and gave low errors due to RE ≤ 15%. This study calculated RaRMSE and RaMAE values to assess the calibrated models' accuracy. Figure 4 presents a graphical representation of the RaRMSE and RaMAE values for the linear regression-based calibrated models, compared to the non-calibrated (original) models during the validation period. Figure 4 shows that positive values represent progress, while negative values indicate a decline in the models. The Cal_HS model exhibited a 25.6% increase in RMSE and MAE values, as shown in Figure 4. In comparison, Cal_Oudin demonstrated a 13.7% increase in RMSE values and a 21.9% increase in MAE values. It's worth mentioning that the HS and Oudin models calibrated locally had worse statistics than their non-calibrated counterparts. Similarly, a study by Khodayvandie et al. (2022) reported that the locally calibrated HS model gave more inadequate statistics than the original HS model at two locations in Iran. Some studies have also stated that there was no improvement in the prediction performances of some empirical models after they were calibrated (Landeras et al. 2008; Valipour 2015; Djaman et al. 2016; Shiri, 2017; Farias et al. 2020). The Cal_HS and Cal_Oudin showed medium errors like their original versions, with 19.8 and 19.0 % RE values, respectively. Therefore, HS and Oudin models do not need to be calibrated by linear regression for the best performance in this study region. The Jensen-Haise (JH) model showed the most significant improvement

(decreases of 73 and 80% of RMSE and MAE in the validation period, respectively), ranked first among all models after calibration, and was followed by calibrated Makkink (MAK). Ritchie model's performance slightly improved after calibration, and the model's RMSE and MAE values have decreased by 16.3% and 13.7%, respectively. Similar to our research, studies by Djaman et al. (2016), Shiri (2017), Banda et al. (2018), Ferreira et al. (2019), Khodayvandie et al. (2022), and Gharehbaghi & Kaya (2022) confirmed that there was generally an improvement in the performance of empirical models after local calibration. As in our study, research conducted by Farzanpour et al. (2019) in semiarid regions of Iran found that the JH model had the highest performance improvement among the radiation-based models after the local calibration procedure. Similarly, Gharehbaghi & Kaya (2022) found that the calibrated JH model best estimated the ET_0 among the empirical models in the Kutahya province of Türkiye, which has dry, hot summers and snowy and cold winters in their study. It can be said that the calibration coefficients obtained by linear regression, as shown in Table 4 in the present study, are valid for the Adana Plain, which has a hot summer Mediterranean climate. Suppose the available data is insufficient or unreliable to solve the FAO-56 PM equation. In that case, it is possible to use calibration coefficients for JH and MAK models when dealing with irrigation practices in the studied area.

Table 4- The statistical summary of the empirical models

	Calibration period			*Validation period*		
Model	*RMSE* $mm\ d^{-1}$	*MAE* $mm\ d^{-1}$	*RE* %	*RMSE* $mm\ d^{-1}$	*MAE* $mm\ d^{-1}$	*RE* %
JH	1.846	1.654	44.3	1.845	1.658	39.1
MAK	0.793	0.687	19.0	1.089	0.970	23.1
Ritchie	1.296	1.014	31.1	1.379	1.054	29.3
HS	0.953	0.762	22.9	0.742	0.548	15.7
Oudin	0.743	0.566	17.8	0.789	0.575	16.7
Cal_ JH	0.390	0.288	9.4	0.491	0.340	10.4
Cal_ MAK	0.419	0.308	10.1	0.564	0.408	12.0
Cal_ Ritchie	0.956	0.741	22.9	1.154	0.910	24.5
Cal_HS	0.703	0.543	16.9	0.932	0.745	19.8
Cal_ Oudin	0.684	0.524	16.4	0.897	0.702	19.0

Notes: Calibrated models denoted as Cal

Table 5- Empirical and correlation coefficients used for calibrating empirical models

		$ET_0(PM) = a + bET_{0_{model}}$	
Model	***a***	***b***	***R***
JH	0.8334	0.5766	0.94
MAK	0.4183	1.0718	0.93
Ritchie	1.8943	0.4855	0.57
HS	0.2288	0.8238	0.80
Oudin	0.2685	0.8805	0.81

Notes: a represents intercept, b represents slope, and R represents correlation coefficient

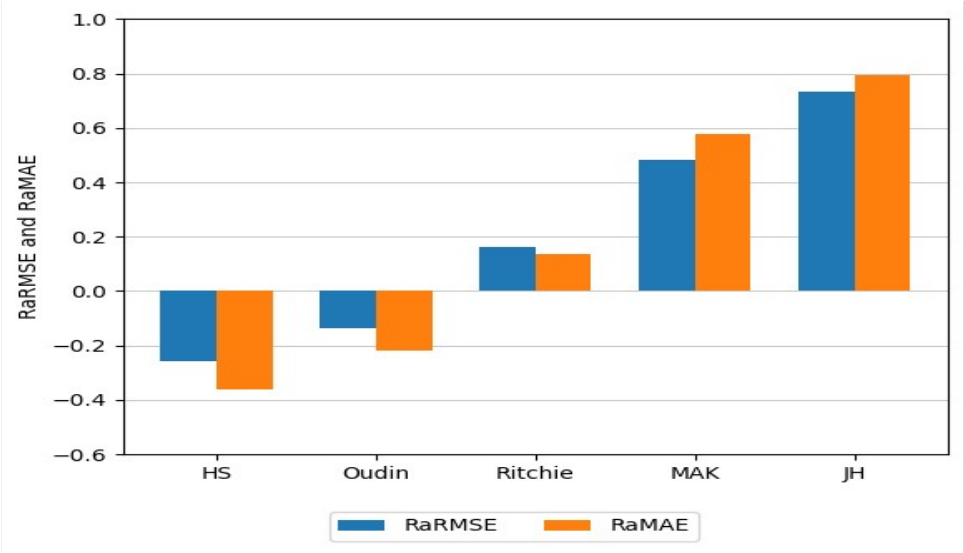


Figure 4- Root mean square error ratio (RaRMSE) and mean absolute error ratio (RaMAE) values of the linear regression-based calibrated empirical models for the validation period

3.2. Data-driven approaches

The statistics of the data-driven models created for the four combinations are shown in Table 6. Table 7 presents the GEP and MLR models' mathematical expressions, respectively. Table 7 shows that, generally, there is a minimal difference between the RMSE, RE, and R values of all models during the training and validation periods. However, the training period's statistical results were slightly better than the validation period for all models. The models with T_{max}, T_{min}, and R_s inputs (3rd combination) and T_{max}, T_{min}, R_s, and U_2 inputs (4th combination) of all data-driven methods successfully estimated ET_0, as evidenced by their high R and low RMSE, MAE and RE values (RE ≤ 15 %). 4th combination of all data-driven methods provided the most accurate predictions of ET_0 during both the training and validation periods. The RMSEs were found to be less than 0.414 mm per day, while the REs were not more than 9.1% in value. So, the findings from the present research coincide with the existing literature suggesting that incorporating more climate variables usually results in improved accuracy of model estimation (Shiri 2017; Mattar & Alazba 2019; Niaghi et al. 2021; Yamaç 2021; Yıldırım et al. 2023; Bayram & Çıtakoğlu 2023).

The RF models showed the lowest RMSE, MAE, and RE in the training and validation periods compared to the GEP, ANN, MLR, and SVM models (during the validation period, the RF1 model was the only exception). The RF4 model achieved the best results with the values of RMSE of 0.116 mm d^{-1}, MAE of 0.076 mm d^{-1}, RE of 2.0 %, and R of 0.995 in the training period, and values of RMSE of 0.224 mm d^{-1}, MAE of 0.151 mm d^{-1}, RE of 3.9 %, and R of 0.983 in the validation period. Overall, the RF and ANN models displayed the best performance among all data-driven models. GEP, MLR, and SVM models showed nearly similar performance considering T_{max}, T_{min}, U_2 and T_{max}, T_{min}, R_s, and U_2 combinations. However, SVM1 and SVM2 models displayed worse performance than their analogous among all data-driven models (RMSE = 1.505-1.454 mm d^{-1}, MAE =1.273-1.244 mm d^{-1}, RE = 30.4-29.6 %, respectively in the validation period). The hyperparameters for the SVM models were determined through trial and error during the study. Due to the limited options for hyperparameter selection provided by the Orange package, SVM1 and SVM2 models are considered not to reach the desired performance level. According to the data presented in Table 6, SVM3 and SVM4 demonstrated slightly better performance levels than those of GEP3 and GEP4, although somewhat inferior to other analogous data-driven models. The results obtained were consistent with previous studies conducted by Sayyadi et al. (2009); Rahimikhoob (2010); Traore et al. (2010), and Yurtseven & Serengil (2021). The studies mentioned utilized the ANN (MLP) method and found it more accurate than other methods when estimating ET_0 in different climates worldwide. Similar to our research, according to Wang et al. (2022), RF-based ET_0 models outperformed GEP-based ET_0 models in their study performed in southwest China's Pearl River basin, which has a tropical and subtropical humid climate. Also, Niaghi et al. (2021) found that in the Red River Valley in the USA, the RF model had a superior ET_0 performance compared to the GEP, SVM, and MLR approaches. In contrast, Yurtseven & Serengil (2021) found that ANN and SVM methods performed better than the RF method in estimating ET_0 in semiarid highland environments.

Studies in the existing literature have explored the impact of different data preprocessing techniques on estimation accuracy in ANNs. For instance, Katipoğlu et al. (2023) investigated the influence of various data preprocessing methods on the accuracy of evaporation estimation using ANN in Adana, which has a Mediterranean climate. Their findings revealed that the standard scalar optimization algorithm presented the highest level of accuracy, and the power transformer showed second-degree promising results. Recently, a novel method employed in evaporation modeling involves implementing hybrid models. Katipoğlu (2023) predicted evaporation with wavelet-based hyperparameter optimized k-nearest neighbors (KNN) and extreme gradient boosting (XGBoost) algorithms in a semi-arid environment. The study specifically investigated the effectiveness of different "mother wavelet" types in enhancing the accuracy of evaporation prediction models. Combining the biorthogonal 2.2 (rbio2.2) mother wavelet with the KNN algorithm yielded the most accurate evaporation predictions. Decomposing the data into sub-signals using wavelet transform generally improved the performance of both KNN and XGBoost models.

In our study, the results presented in Table 6 and Figure 5 indicate that T_{max}, T_{min}, U_2 input (2nd combination) models had higher accuracy than T_{max} and T_{min} input (1st combination) models, and T_{max}, T_{min}, and R_s input (3rd combination) models had higher accuracy than T_{max}, T_{min}, U_2 input (2nd combination) models in two studied periods for all data-driven methods. Additionally, adding U_2 and R_s into the first combination significantly improved the accuracy of the models. During the calibration and validation periods, the first and second model combinations had medium errors, while the third and fourth combinations had low errors based on RE values. The data presented in Figure 5 shows that the RE (relative error) values for the first and second combinations of GEP, ANN, MLR, and RF methods are less than 25%. On the other hand, for the first and second combinations of SVM, the RE values are around 30%. The RE values for the 3rd and 4th combinations of RF and ANN are below 10%, while for the 3rd and 4th combinations of GEP, MLR, and SVM, they are below 15%.

Including U_2 data in the RF1 model improved the accuracy of ET_0 estimation for the new model (RF2), reducing RMSE and MAE values by 20.0 and 19.5%, respectively, during the validation period. On the other hand, when R_s data were incorporated into the RF1 model, the accuracy of ET_0 estimation for the new model (RF3) significantly improved, with a decrease in RMSE and MAE values by 60.9 and 65.7%, respectively, during the validation period. The addition of U_2 and R_s data to the RF1 model led to significant improvement in the accuracy of the ET_0 estimation of the new model (RF4), as evidenced by the decrease in RMSE and MAE values by 78.7% and 81.9%, respectively, during the validation period. Other data-driven methods also demonstrated behaviour similar to that of the RF method. The research conducted by Traore and Guven (2012) and Citakoglu et al. (2014) revealed similar outcomes. It was observed that incorporating U_2 with other meteorological variables enhanced the

efficiency of the models. It is clear from here that the R_s variable has more effect on estimating ET_0 than the T_{max}, T_{min}, and U_2 variables. In Section 2.3, Figure 2 illustrates the strong positive correlation (R = 0.875) between R_s and ET_0, while indicating a notably weak negative correlation (R = -0.165) between RH and ET_0. It is noteworthy that the addition of RH to the GEP4 model (T_{max}, T_{min}, R_s, U_2) during the formulation of the GEP models did not prompt consideration of the RH parameter by the GEP software when establishing the equations. This observation suggests that RH holds minimal influence within the study area. Notably, all data-driven models constructed using the 4-input combination effectively predicted ET_0 without the inclusion of RH, exemplified by the RF model yielding RE = 2.0% and RMSE = 0.116 mm d^{-1} during the training period (Table 5). Consequently, RH was omitted in this study's formulation of input combinations. Pereira et al. (2015) state that R_s can be the dominant factor in predicting ET_0 during summer in humid and sub-humid climates. This is because, in these conditions, the impact of the radiation term is greater than that of the pressure deficit of water vapor and wind factor of the PM method. Studies conducted by Yamaç (2021), Dimitriadou and Nikolakopoulos (2022), Yıldırım et al. (2023), Bayram and Çıtakoğlu (2023) found similar results that R_s has more effect on estimating ET_0 than other variables. A study was conducted in the Peloponnese region of Greece, which has a Mediterranean climate (Csa) similar to our study area. Dimitriadou and Nikolakopoulos (2022) found that net radiation (R_n) and sunshine hours (n) had a more significant impact on ET_0 than other variables.

Table 6- Performance statistics of data-driven models in training and validation periods

		Training period				***Validation period***			
Model	***Input Parameters***	*RMSE mm d^{-1}*	*MAE mm d^{-1}*	*RE %*	*R*	*RMSE mm d^{-1}*	*MAE mm d^{-1}*	*RE %*	*R*
GEP1	T_{max}, T_{min}	0.921	0.759	22.3	0.613	1.094	0.921	23.3	0.405
GEP2	T_{max}, T_{min}, U_2	0.820	0.680	19.7	0.713	0.850	0.708	18.1	0.582
GEP3	T_{max}, T_{min}, R_s	0.438	0.335	10.7	0.928	0.591	0.411	13.1	0.858
GEP4	T_{max}, T_{min}, R_s, U_2	0.337	0.257	8.3	0.958	0.413	0.300	9.0	0.901
ANN1	T_{max}, T_{min}	0.968	0.794	21.8	0.631	0.983	0.807	21.7	0.595
ANN2	T_{max}, T_{min}, U_2	0.825	0.684	18.8	0.751	0.831	0.694	18.7	0.734
ANN3	T_{max}, T_{min}, R_s	0.420	0.303	8.1	0.942	0.422	0.296	7.7	0.939
ANN4	T_{max}, T_{min}, R_s, U_2	0.264	0.197	5.6	0.977	0.265	0.194	5.4	0.976
MLR1	T_{max}, T_{min}	0.920	0.754	22.1	0.615	1.107	0.937	23.5	0.633
MLR2	T_{max}, T_{min}, U_2	0.818	0.679	19.6	0.713	0.847	0.707	18.0	0.767
MLR3	T_{max}, T_{min}, R_s	0.383	0.280	9.2	0.944	0.518	0.355	11.0	0.941
MLR4	T_{max}, T_{min}, R_s, U_2	0.249	0.183	6.0	0.977	0.299	0.227	6.3	0.974
RF1	T_{max}, T_{min}	0.633	0.488	12.9	0.862	1.054	0.832	22.0	0.507
RF2	T_{max}, T_{min}, U_2	0.450	0.341	9.1	0.933	0.843	0.670	17.7	0.725
RF3	T_{max}, T_{min}, R_s	0.206	0.138	3.5	0.986	0.412	0.285	7.2	0.941
RF4	T_{max}, T_{min}, R_s, U_2	0.116	0.076	2.0	0.995	0.224	0.151	3.9	0.983
SVM1	T_{max}, T_{min}	1.503	1.264	30.9	-0.670	1.505	1.273	30.4	-0.718
SVM2	T_{max}, T_{min}, U_2	1.445	1.235	29.9	-0.583	1.454	1.244	29.6	-0.644
SVM3	T_{max}, T_{min}, R_s	0.565	0.455	11.8	0.944	0.573	0.456	11.5	0.883
SVM4	T_{max}, T_{min}, R_s, U_2	0.364	0.272	7.6	0.957	0.369	0.275	7.6	0.953

Table 7- The GEP and MLR models' mathematical expressions

Model	*Formula*
GEP1	$ET_0 = \sqrt{(3.41) + \frac{(3.26)\text{x}(-3.63) + T_{max}}{\sqrt{T_{max}}}} + \sqrt{\frac{\frac{T_{max}\ \text{x}\ T_{max}}{6.52\ \text{x}\ T_{min}}}{\sqrt{T_{max}^{\,6.52}}}} + \left(\frac{\frac{T_{max}\ \text{x}\ T_{max}}{(-9.58) - T_{min}}}{\sqrt{3.75^{4.28}}}\right)^2$
MLR1	$ET_0 = -0.391 + 0.129 \times T_{max} + 0.03 \times T_{min}$
GEP2	$ET_0 = U2 + \log(T_{max} + U2 + T_{min}\ \text{x}\ (-9.07) + T_{max}^{\,2} + T_{max}) + (-8.45) + \sqrt{T_{max} + U_2 + 1.57} + (-7.90)$
MLR2	$ET_0 = -1.908 + 0.170 \times T_{max} - 0.009\ \text{x}\ T_{min} + 1.127\ \text{x}\ U_2$
GEP3	$ET_0 = \max\left(\frac{\frac{R_s + 8.45}{2} + \frac{R_s^{\,2}}{2}}{2},\ \frac{1}{8.45}\ \text{x}\ (R_s + T_{min})\right) + \tanh(-0.14) + \frac{(-0.14)\ \text{x}\ R_s}{2} + \frac{(-1.72)}{R_s} + \frac{2}{\frac{T_{max} + T_{min}}{2} - \frac{(-4.34) + T_{max}}{2} + \frac{(-4.34) - T_{max}}{T_{max}}}$ while $\max(x,y) = \begin{cases} y, & x \le y \\ x, & x > y \end{cases}$
MLR3	$ET_0 = -1.336 + 0.020\ \text{x}\ T_{max} + 0.070\ \text{x}\ T_{min} + 0.179\ \text{x}\ R_s$
GEP4	$ET_0 = max\left(\left(tanh(R_s) + \frac{(-2.09 + U2)}{2}\right)^2, \left(\frac{R_s + 5.97}{2} + \frac{U2 + (-9.02)}{2}\right)\right) + log\left(min\left((T_{max} - U2 + T_{max}),\ \left((U2)^2 + \frac{(-1.72) + R_s}{2}\right)\right)\right) + arctan\left(\frac{U2 + arctan\left(\frac{R_s + T_{min}}{2}\right) - U2 - R_s + 0.78}{2}\right)$ $while\ max(x,y) = \begin{cases} y, & x \le y \\ x, & x > y \end{cases}\ and\ \ min(x,y) = \begin{cases} x, & x \le y \\ y, & x > y \end{cases}$
MLR4	$ET_0 = -2.350 + 0.054\ \text{x}\ T_{max} + 0.040\ \text{x}\ T_{min} + 0.792\ \text{x}\ U_2 + 0.169\ \text{x}\ R_s$

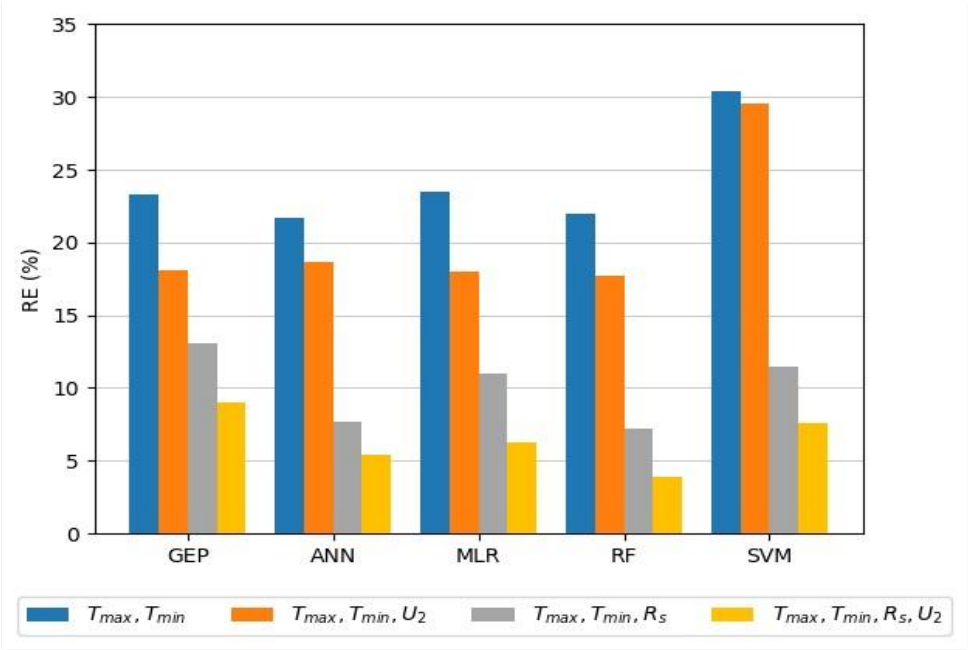


Figure 5- Performance of the data-driven models during the validation period

3.3. Comparison of empirical and data-driven approaches

This section of the study compares the performance of original (non-calibrated), locally calibrated empirical models and data-driven methods in estimating ET_0. The HS and Oudin models in this study correspond to the GEP1, ANN1, MLR1, RF1, and SVM1 models, while the JH, MAK, and Ritchie models align with the GEP3, ANN3, MLR3, RF3, and SVM3 models based on

their respective input combinations. As shown in Table 4 and Table 6, the models that used three or four inputs were more effective than the models that used only two inputs when it came to estimating ET_0, whereas the four-input data-driven models (T_{max}, T_{min}, R_s, U_2 input) outperformed all original and calibrated empirical models. ANN1 and RF1 models (T_{max} and T_{min} inputs) provided higher accuracy than their analogous data-driven models in training and validation periods. On the other hand, original (non-calibrated) HS and Oudin models (T_{max} and T_{min} input) performed better than their equivalent data-driven models in the validation period. The calibrated JH (Cal_JH) model predicted ET_0 slightly better than the GEP3, MLR3 and SVM3 models, while the ANN3 and RF3 models outperformed the Cal_JH model in the validation period. Cal_MAK model estimated somewhat better than the GEP3 model in the validation period, whereas the models ANN3, MLR3, RF3, and SVM3 were found to be more accurate than the Cal_MAK model when estimating ET_0 in the validation period. Similar to our research, Mehdizadeh et al. (2017) have found that some calibrated radiation-based empirical models performed similarly to soft computing approaches in their study.

On the other hand, HS and ANN1 models with T_{max} and T_{min} inputs presented more accurate results than their equivalent models, while among all the models created with T_{max} and T_{min} inputs, the HS model was the most accurate. Cal_JH and RF3 models created using T_{max}, T_{min}, and R_s inputs performed more accurately than their analogous models, whereas out of all the models that were tested with T_{max}, T_{min} and R_s input, the RF3 model displayed the highest level of accuracy. The RF4 model proved to be the most accurate in predicting the ET_0 in all models.

Figures 6, 7, and 8 show diagrams depicting the Taylor, the box plot, and the prediction-errors box plot, and Figure 9 shows time series graphics for the best empirical and data-driven model during the validation period, respectively. Table 8 enlists the descriptive statistics of prediction errors for the top input combinations. In Figure 6, it is evident that the RF4 model outperformed the other models. The RF4 model performed best with low RMSE and high correlation coefficient. The RF4 model provided estimates of ET_0 similar to those of the FAO56-PM method. Also, it is apparent from the Figure 6 that RF3 and Cal_JH models produce similar outcomes. The boxplot charts of the developed data-driven models are similar, as shown in Figure 7. Average ET_0 values estimated by data-driven methods range from 4.35 to 4.38 mm d^{-1}, while that of FAO56-PM is 4.37 mm d^{-1}. However, both HS and Cal_JH had a higher ET_0 compared to FAO56-PM. The average ET_0 of the HS was 4.80 mm d^{-1}, while the average ET_0 of Cal_JH was 4.47 mm d^{-1}. According to Table 8, the RF4 model followed the corresponding values with the lowest lower quartile (0.046) and upper quartile (0.1966), and the standard deviation value is less than that of other data-driven and empirical models. According to Figure 8, it can be inferred that RF4, characterized by the lowest median value and narrow interquartile range, emerges as the most effective model. According to the information depicted in Figure 9, it is apparent that data-driven models were more effective than empirical equations in forecasting peak ET_0 values. Among the models assessed, RF4 exhibited the highest level of accuracy in predicting peak ET_0 values. In contrast, the HS and Cal_JH equations demonstrated deficiencies in precisely forecasting peak ET_0 values. Based on Figure 9, it can be deduced that the Hargreaves-Samani equation generally overestimates ET_0, while the Cal_JH equation tends to underestimate ET_0.

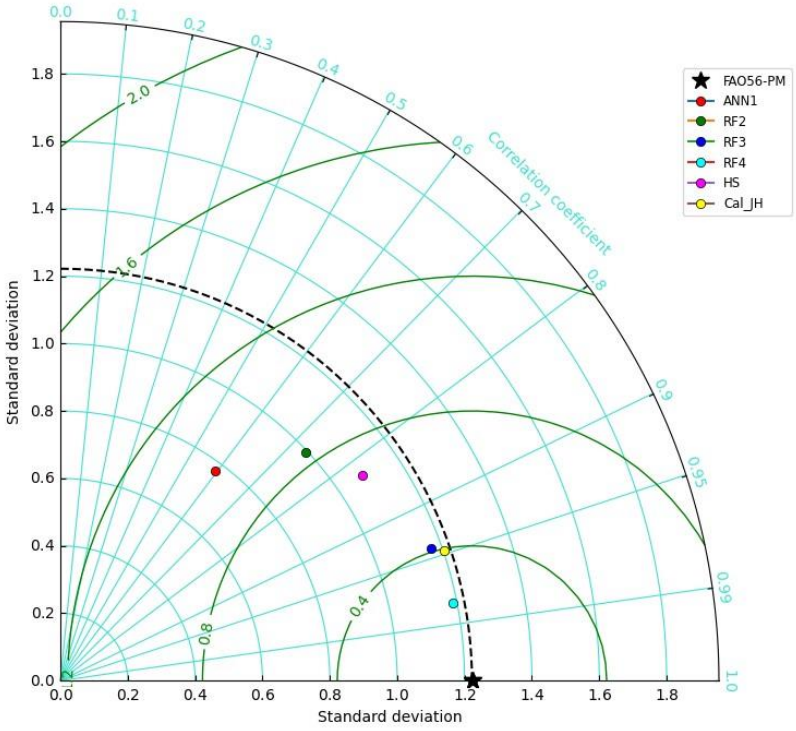


Figure 6- Taylor diagram for the best empirical and data-driven model according to each input combination

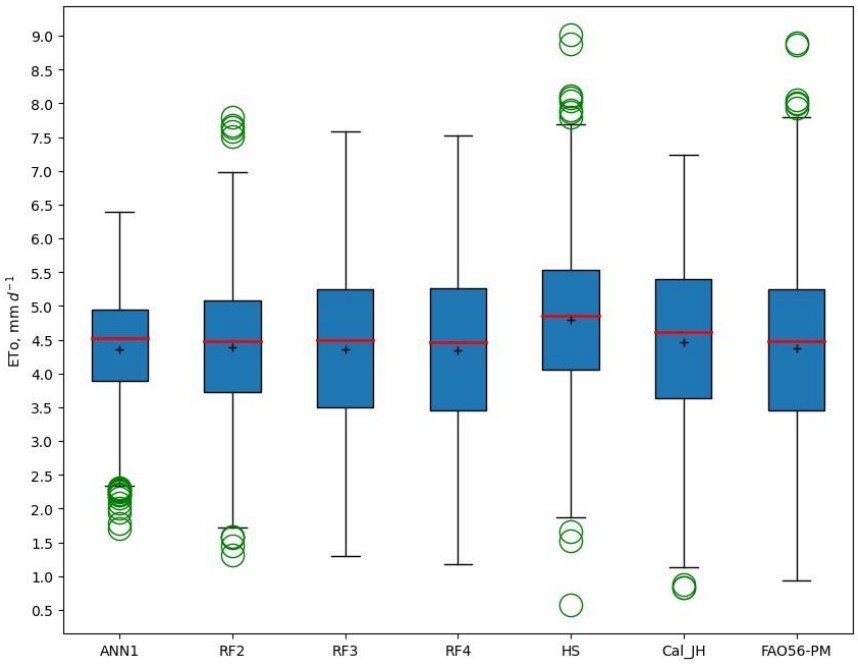


Figure 7- Box plot diagram for the best empirical and data-driven model according to each input combination

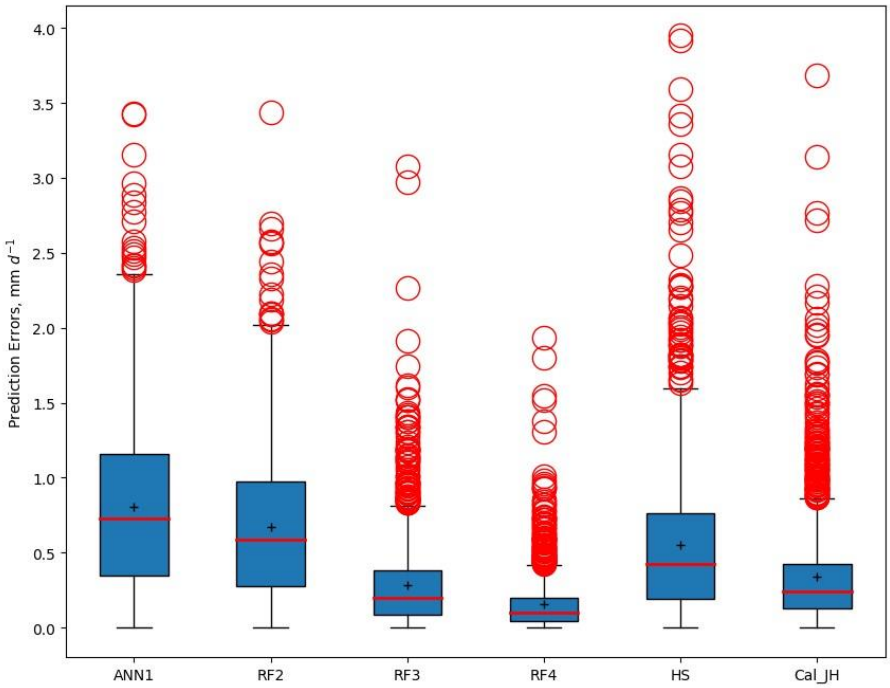
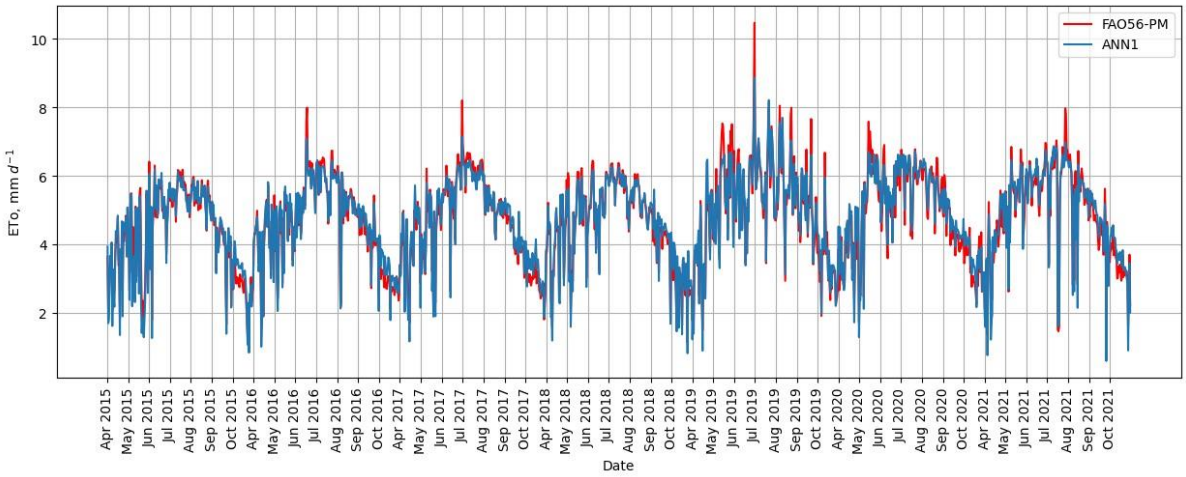
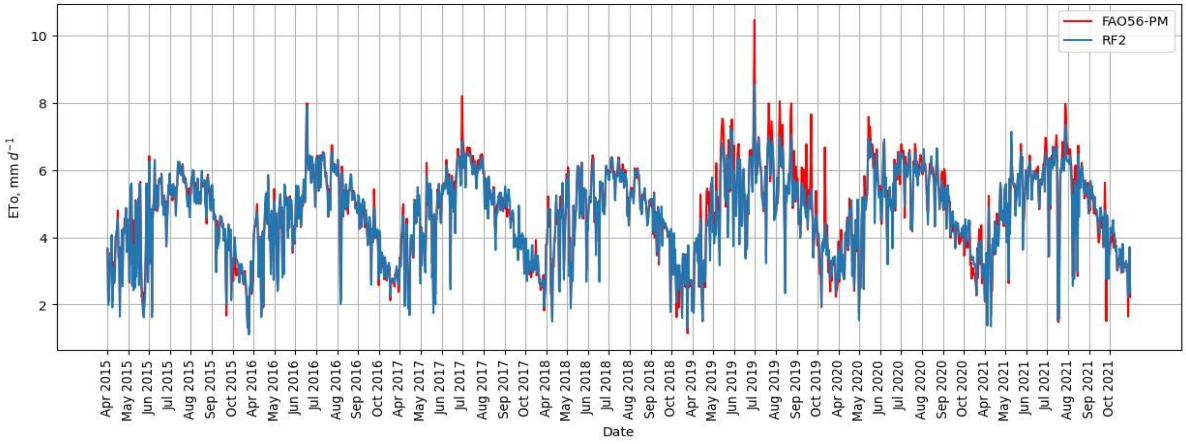
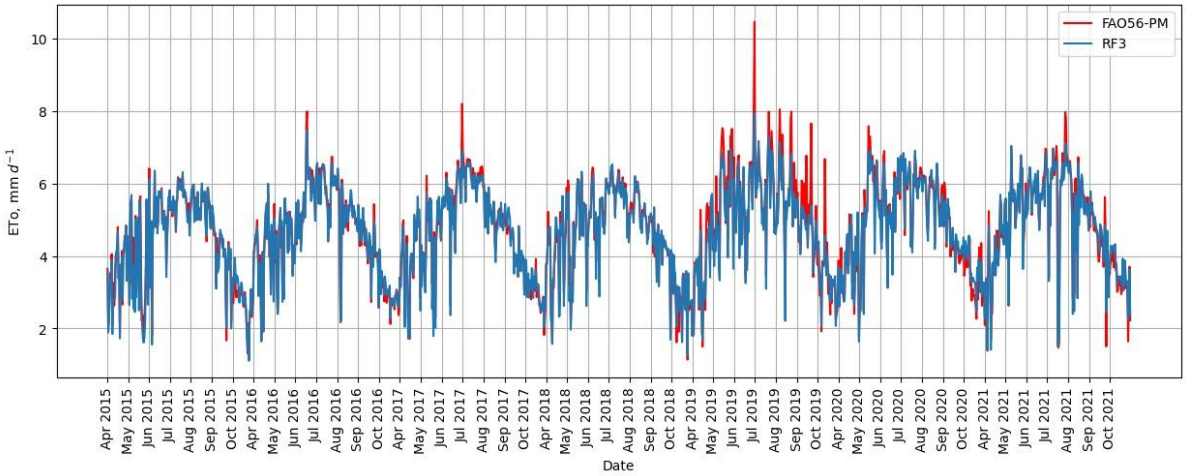
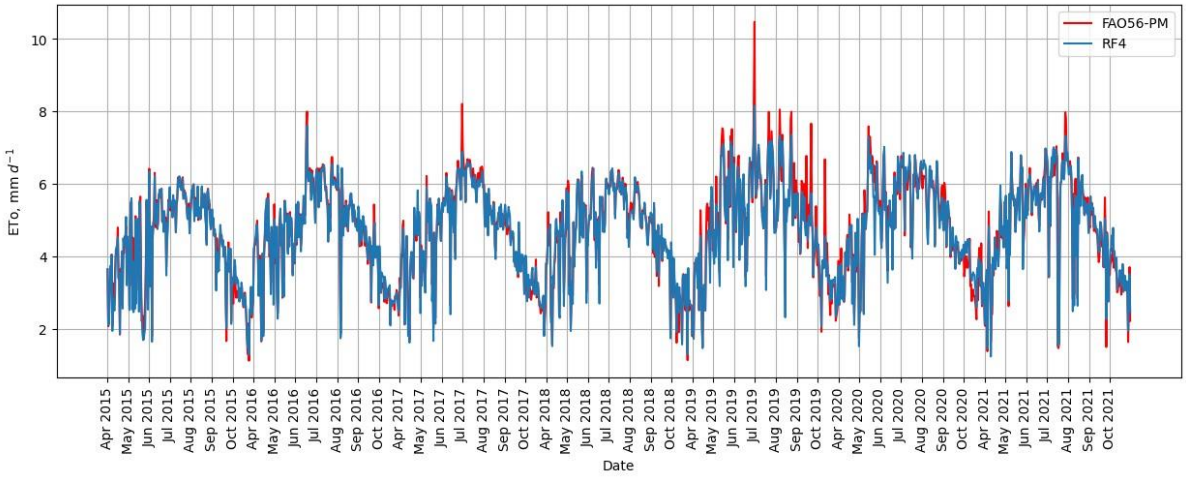


Figure 8- Prediction-errors box plot diagram for the best empirical and data-driven model according to each input combination


FAO56-PM
ANN1
ETo, mm d^{-1}
Date


FAO56-PM
RF2
ETo, mm d^{-1}
Date


FAO56-PM
RF3
ETo, mm d^{-1}
Date


FAO56-PM
RF4
ETo, mm d^{-1}
Date

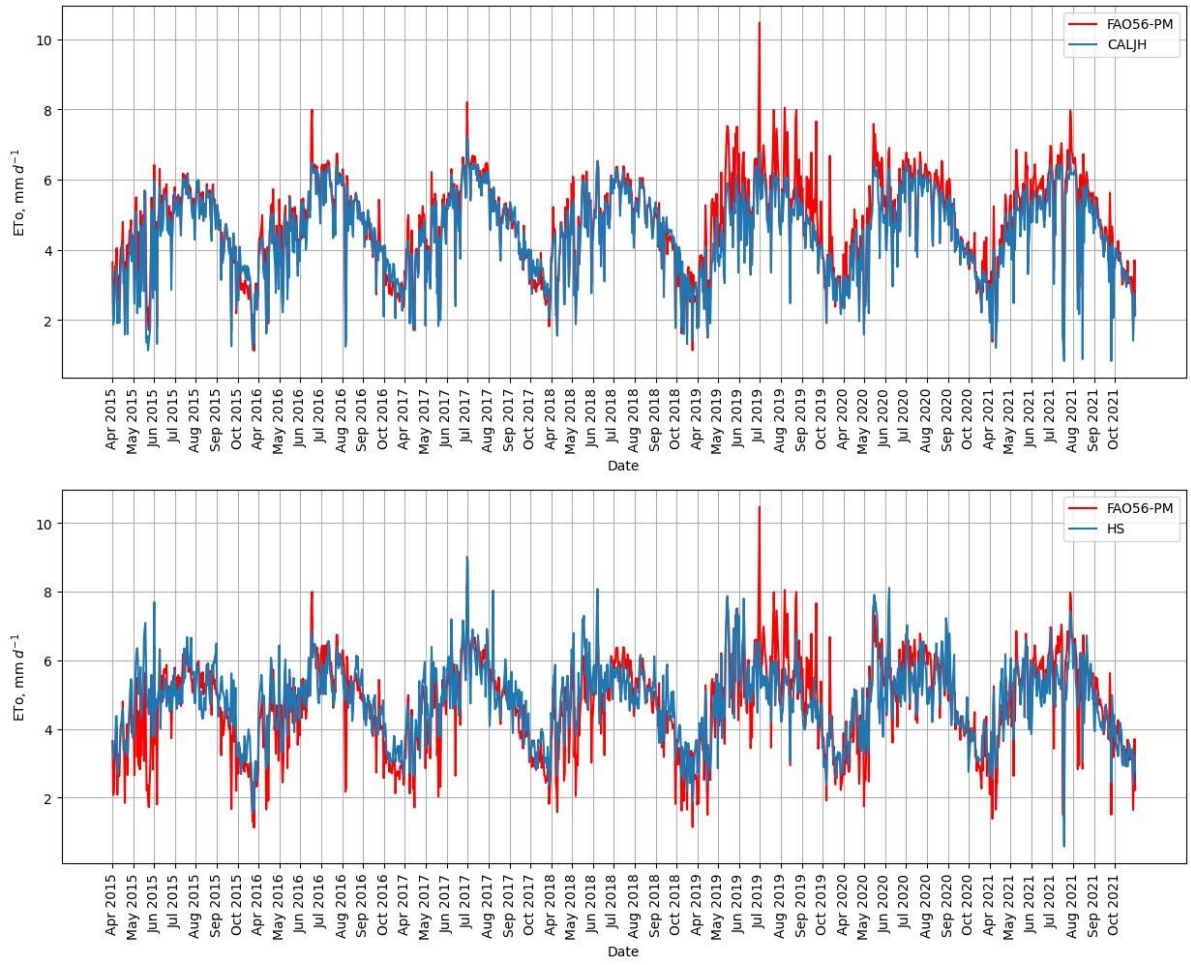


Figure 9- Time series graphics of daily FAO56-PM ETo and daily model-estimated ETo for the optimal data-driven and empirical model based on each input combination

Table 8- Box plot statistics for the optimal models based on each input combination

Model	Lower Quartile	Upper Quartile	Mean	Median	Standard Deviation
ANN1	0.3496	1.1593	0.8064	0.7282	0.5619
RF2	0.2762	0.9745	0.6730	0.5860	0.5005
RF3	0.0877	0.3829	0.2844	0.2006	0.2965
RF4	0.0466	0.1966	0.1544	0.0973	0.1799
HS	0.1902	0.7615	0.5482	0.4250	0.4999
Cal_JH	0.1259	0.4215	0.3398	0.2382	0.3542

Many studies show that data-driven models with all input parameters are superior to empirical and calibrated empirical models, as in our study (Benzaghta et al. 2012; Karimaldini et al. 2012; Tabari et al. 2013b). Benzaghta et al. (2012) found that using all input parameters in ANN leads to better evaporation estimation performance. Jain et al. (2008) highlighted that the accurate estimation of ET_0 requires temperature and radiation data as crucial inputs in Southwestern Idaho, USA, which has a semi-arid climate. In our study, the performance of equation methods was significantly impacted by the presence or absence of critical input. However, the performance values in data-driven models may vary despite having the same input set due to model dynamics.

In this study, the hyperparameter values used to create data-driven models in the Orange software were limited to the software's provided constraints. Consequently, precise determination of optimal model performance values through trial and error was not feasible in the SVM models. Nonetheless, utilizing software such as python enables the performance of machine learning model tuning and hyperparameter optimization using techniques like “Grid Search” and “Random Search”. We recommend employing these methods in future studies to enhance the accuracy of ET_0 predictions. In recent research, hybrid models have shown promise in enhancing the accuracy of estimating ET_0. These models have demonstrated improved estimation accuracy by leveraging the strengths of diverse algorithms. Furthermore, it is suggested that future studies incorporate the use of hybrid models and integrate deep learning models alongside traditional data-driven methods to optimize ET_0 estimation further.

4. Conclusions

The estimation of reference evapotranspiration (ET_0) is crucial for water resource management, irrigation practices, and agricultural and hydrometeorological research. Thus, the precise prediction of ET_0 is of the utmost importance.

Based this study, the key findings derived from this study are outlined as follows: The study conclusively establishes that the application of data-driven techniques yields satisfactory outcomes in estimating daily ET_0, even in cases where not all the climate parameters required for the reference method FAO56-PM ET_0 are available. The MLR and GEP equations acquired can have applications in agricultural irrigation within the researched area. The calibration process significantly improved the accuracy of the JH and MAK empirical equations in forecasting ET_0. We recommend utilizing the calibration coefficients derived from these equations to make precise ET_0 projections within the research area. The Orange software's limitations in selecting hyperparameters, particularly in this study, resulted in reduced ET_0 estimation performance from the SVM1 and SVM2 models compared to other equivalent model combinations. It is advisable to consider employing methodologies such as "grid search" and "random search" for the purpose of hyperparameter optimization in forthcoming research endeavors within the domain of machine learning.

In future research, the predictive accuracy of reference ET_0 could be investigated utilizing deep learning models, hybrid models, and methodologies incorporating lag values.

Funding

The authors did not receive support from any organization for the submitted work.

Data Availability

The data supporting this study's findings are available on request from the corresponding author.

Conflict of Interest

The authors declare that there is no conflict of interest.

References

Abrishami N, Sepaskhah A R & Shahrokhnia M H (2019). Estimating wheat and maize daily evapotranspiration using artificial neural network. Theoretical and Applied Climatology 135: 945-958. https://doi.org/10.1007/s00704-018-2418-4

Achite M, Jehanzaib M, Sattari M T, Toubal A K, Elshaboury N, Wałęga A, Krakauer N Y, Yoo J & Kim T (2022). Modern techniques to modeling reference evapotranspiration in a semiarid area based on ANN and GEP models. Water 14(8): 1210. https://doi.org/10.3390/w14081210

Allen R G, Pereira L S, Raes D & Smith M (1998). Crop evapotranspiration. Guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper (56) Rome: FAO

Almorox J & Grieser J (2016). Calibration of the Hargreaves–Samani method for the calculation of reference evapotranspiration in different Köppen climate classes. Hydrology Research 47(2): 521-531. https://doi.org/10.2166/nh.2015.091

Antonopoulos V Z & Antonopoulos A V (2017). Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate variables. Computers and Electronics in Agriculture 132: 86-96. https://doi.org/10.1016/j.compag.2016.11.011

Banda P, Cemek B & Küçüktopcu E (2018). Estimation of daily reference evapotranspiration by neuro computing techniques using limited data in a semiarid environment. Archives of Agronomy and Soil Science 64(7): 916-929. https://doi.org/10.1080/03650340.2017.1414196

Bayram S & Çıtakoğlu H (2023). Modeling monthly reference evapotranspiration process in Turkey: application of machine learning methods. Environmental Monitoring and Assessment 195(1): 67. https://doi.org/10.1007/s10661-022-10662-z

Benzaghta M A, Mohammed T A, Ghazali A H, Amin M S M & Soom M A B M (2012). Prediction of evaporation in tropical climate using artificial neural network and climate based models. Scientific Research and Essays 7(36): 3133-3148. https://doi.org/10.5897/SRE11.1311

Breiman L (2001). Random forests. Machine Learning 45: 5-32

Chen Z, Zhu Z, Jiang H and Sun S (2020). Estimating daily reference evapotranspiration based on limited meteorological data using deep learning and classical machine learning methods. *Journal of Hydrology* 591(2020): 125286. https://doi.org/10.1016/j.jhydrol.2020.125286

Choi Y, Kim M, O'Shaughnessy S, Jean J, Kim Y & Song W J (2018). Comparison of artificial neural network and empirical models to determine daily reference evapotranspiration. *Journal of Korean Society of Agricultural Engineers* 60(6): 43-54. https://doi.org/10.5389/KSAE.2018.60.6.043

Citakoglu H, Cobaner M, Haktanir T & Kisi O (2014). Estimation of monthly mean reference evapotranspiration in Turkey. Water Resources Management 28: 99-113. https://doi.org/10.1007/s11269-013-0474-1

Cobaner M, Citakoglu H, Haktanir T & Kisi O (2017). Modifying Hargreaves–Samani equation with meteorological variables for estimation of reference evapotranspiration in Turkey. Hydrology Research 48: 480-497. https://doi.org/10.2166/nh.2016.217

Corzo G & Solomatine D (2007). Baseflow separation techniques for modular artificial neural network modelling in flow forecasting. *Hydrological Sciences Journal* 52(3): 91-507. https://doi.org/10.1623/hysj.52.3.491

CSB (2023). Urbanization and climate change. The governor of Adana, Ministry of Environment. https://adana.csb.gov.tr/ilim izi-taniyalim-i-1222, Accessed 16 February 2023.

Cutler A, Cutler D R & Stevens J R (2012). Random forests. Ensemble machine learning: Methods and applications, pp. 157-175. https://doi.org/10.1007/978-1-4419-9326-7_5.

Dimitriadou S & Nikolakopoulos K G (2022). Multiple linear regression models with limited data for the prediction of reference evapotranspiration of the Peloponnese, Greece. Hydrology 9(7): 124. https://doi.org/10.3390/hydrology9070124

Djaman K, Tabari H, Balde A B, Diop L, Futakuchi K & Irmak S (2016). Analyses, calibration and validation of evapotranspiration models to predict grass-reference evapotranspiration in the Senegal river delta. *Journal of Hydrology*: Regional Studies 8: 82-94. https://doi.org/10.1016/j.ejrh.2016.06.003

Dong J, Xing L, Cui N, Guo L, Liang C, Zhao L, Wang Z & Gong D (2024) Estimating reference crop evapotranspiration using optimized empirical methods with a novel improved Grey Wolf Algorithm in four climatic regions of China. Agricultural Water Management 291: 2024. https://doi.org/10.1016/j.agwat.2023.108620

Dou X & Yang Y (2018). Evapotranspiration estimation using four different machine learning approaches in different terrestrial ecosystems. Computers and Electronics in Agriculture 148: 95-106. https://doi.org/10.1016/j.compag.2018.03.010

Farias V D, Costa D L, Pinto J V, Souza P J, Souza E B & Ortega-Farias S (2020). Calibration of reference evapotranspiration models in Pará. Acta Scientiarum Agronomy 42. https://doi.org/10.4025/actasciagron.v42i1.42475

Farzanpour H, Shiri J, Sadraddini A A & Trajkovic S (2019). Global comparison of 20 reference evapotranspiration equations in a semiarid region of Iran. Hydrology Research 50(1): 282-300. https://doi.org/10.2166/nh.2018.174

Fauset L V (2006). Fundamentals of neural networks: architectures, algorithms and applications. Pearson Education India.

Feng Y, Cui N, Zhao L, Hu X & Gong D (2016). Comparison of ELM, GANN, WNN and empirical models for estimating reference evapotranspiration in humid region of Southwest China. *Journal of Hydrology* 536: 376-383. https://doi.org/10.1016/j.jhydrol.2016.02.053

Feng Y, Peng Y, Cui, N, Gong D & Zhang K (2017). Modeling reference evapotranspiration using extreme learning machine and generalized regression neural network only with temperature data. Computers and Electronics in Agriculture 136: 71-78. https://doi.org/10.1016/j.compag.2017.01.027

Ferreira C (2001). Gene expression programming: A new adaptive algorithm for solving problems. Complex Systems 13(2): 87-129

Ferreira L B, Cunha F F, Duarte A B, Sediyama G C & Cecon P R (2018). Calibration methods for the Hargreaves-Samani equation. Ciencia E Agrotecnologia 42: 104-114. http://dx.doi.org/10.1590/1413-70542018421017517

Ferreira L B, Cunha F F, de Oliveira R A & Filho E I (2019). Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM - A new approach. *Journal of Hydrology* 572: 556-570. https://doi.org/10.1016/j.jhydrol.2019.03.028

Gao X, Peng S, Xu J, Yang S & Wang W (2015). Proper methods and its calibration for estimating reference evapotranspiration using limited climatic data in Southwestern China. Archives of Agronomy and Soil Science 61(3): 415-426. https://doi.org/10.1080/03650340.2014.933810

Gavili S, Sanikhani H, Kisi O & Mahmoudi M H (2018). Evaluation of several soft computing methods in monthly evapotranspiration modelling. Meteorological Applications 25(1): 128-138. https://doi.org/10.1002/met.1676

Gharehbaghi A & Kaya B (2022). Calibration and evaluation of six popular evapotranspiration formula based on the Penman-Monteith model for continental climate in Turkey. Physics and Chemistry of the Earth, Parts A/B/C, 127, 103190. https://doi.org/10.1016/j.pce.2022.103190

Gocić M, Motamedi S, Shamshirband S, Petković D, Sudheer C, Roslan H & Muhammad A (2015). Soft computing approaches for forecasting reference evapotranspiration. Computers and Electronics in Agriculture 113: 164-173. https://doi.org/10.1016/j.compag.2015.02.010

Goldblum M, Finzi M, Rowan K & Wilson A G (2023). The no free lunch theorem, Kolmogorov complexity, and the role of inductive biases in machine learning. arXiv preprint arXiv:2304.05366. https://doi.org/10.48550/arXiv.2304.05366

Gomariz-Castillo F, Alonso-Sarria F & Cabezas-Calvo-Rubio F (2018). Calibration and spatial modelling of daily ET_0 in semiarid areas using Hargreaves equation. Earth Science Informatics 11(3): 325-340. https://doi.org/10.1007/s12145-017-0327-1

Hargreaves G H & Samani Z A (1985). Reference crop evapotranspiration from temperature. Applied Engineering in Agriculture 1: 96-99.

Huo Z, Feng S, Kang S & Dai X (2012). Artificial neural network models for reference evapotranspiration in an arid area of northwest China. *Journal of Arid Environments* 82: 81-90. https://doi.org/10.1016/j.jaridenv.2012.01.016

Irmak S, Irmak A, Allen R G & Jones J W (2003). Solar and net radiation-based equations to estimate reference evapotranspiration in humid climates. *Journal of Irrigation and Drainage Engineering* 129(5): 336-347. https://doi.org/10.1061/(ASCE)0733-9437(2003)129:5(336)

Izadifar Z (2010). Modeling and analysis of actual evapotranspiration using data driven and wavelet techniques. Undergraduate Dissertation, University of Saskatchewan.

Jacovides C P & Kontoyiannis H (1995). Statistical procedures for the evaluation of evapotranspiration computing models. Agricultural Water Management 27: 365-371. https://doi.org/10.1016/0378-3774(95)01152-9

Jain S K, Nayak P C & Sudheer K P (2008). Models for estimating evapotranspiration using artificial neural networks, and their physical interpretation. Hydrological Processes 22(13): 2225-2234. https://doi.org/10.1002/hyp.6819

Jang J C, Sohn E H, Park K H & Lee S (2021). Estimation of daily potential evapotranspiration in real-time from GK2A/AMI data using artificial neural network for the Korean Peninsula. Hydrology 8(3): 129 https://doi.org/10.3390/hydrology8030129

Jensen M E & Haise H R (1963). Estimating evapotranspiration from solar radiation. *Journal of the Irrigation and Drainage Division* 89(4): 15-41. https://doi.org/10.1061/JRCEA4.000028

Jones J W & Ritchie J T (1990). Crop growth models. In: G. J. Hoffman, T. A. Howel, & K. H. Solomon (eds.) Management of farm irrigation systems. ASAE Monograph 9: 63–69

Kades C (2019). Economy of Adana in 2019 [in Turkish] https://www.adanato.org.tr/ WebDosyalar/V2/Dosyalar/2020/8/13/adana-ekonomisi-2019-16-53-18.pdf, Accessed 14 January 2023

Karimaldini F, Shui L T, Mohamed T A, Abdollahi M & Khalili N (2012). Daily evapotranspiration modeling from limited weather data by using neuro-fuzzy computing technique *Journal of Irrigation and Drainage Engineering* 138(1): 21-34. https://doi.org/10.1061/(ASCE)IR.1943-4774.000034

Karunanithi N, Grenney W J, Whitley L D & Bovee K (1994). Neural networks for river flow prediction. *Journal of Computing in Civil Engineering* 8(2): 201-220. https://doi.org/10.1061/(ASCE)0887-3801(1994)8:2(201)

Katipoğlu O M (2023). Evaporation prediction with wavelet-based hyperparameter optimized K-Nearest Neighbors and Extreme Gradient Boosting Algorithms in a semi-arid environment. Environmental Processes 10(4): 50-51. https://doi.org/10.1007/s40710-023-00669-0

Katipoğlu O M, Pekin M A & Akil S (2023). The Impact of preprocessing approaches on neural network performance: A case sudy on evaporation in Adana, a Mediterranean climate." *Indonesian Journal of Earth Sciences* 3(2): A821-A821 https://doi.org/10.52562/injoes.2023.821

Kaya Y Z, Zeleňáková M, Üneş F, Demirci M, Hlavatá H & Mésároš P (2021). Estimation of daily evapotranspiration in Košice City (Slovakia) using several soft computing techniques. Theoretical and Applied Climatology 144(2021): 287-298. https://doi.org/10.1007/s00704-021-03525-z

Khodayvandie B, Nazemi A H, Shiri J & Trajkovic S (2022). Adopting regional calibration scenarios for ensuring reliable ETo estimates in semiarid regions: assessing the ancillary data supply method. *ISH Journal of Hydraulic Engineering* 28(sup1): 299-309. https://doi.org/10.1080/09715010.2020.1784803

Kisi O (2016). Modeling reference evapotranspiration using three different heuristic regression approaches. Agricultural Water Management 169: 162-172. https://doi.org/10.1016/j.agwat.2016.02.026

Kumar M, Raghuwanshi N S & Singh R (2011). Artificial neural networks approach in evapotranspiration modeling: A review. Irrigation Science 29: 11-25. https://doi.org/10.1007/s00271-010-0230-8

Landeras G, Ortiz-Barredo A & López J J (2008). Comparison of artificial neural network models and empirical and semi-empirical equations for daily reference evapotranspiration estimation in the Basque Country (Northern Spain). Agricultural Water Management 95(5): 553-565. https://doi.org/10.1016/j.agwat.2007.12.011

Liu S & Xu Z (2018). Micrometeorological methods to determine evapotranspiration. In: Li, X., Vereecken, H. (eds) Observation and measurement. Ecohydrol. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-47871-4_7-2

Makkink G F (1957). Ekzameno de la formulo de Penman. Netherlands Journal of Agricultural Science 5: 290-305.

Mattar M A & Alazba A A (2019). GEP and MLR approaches for the prediction of reference evapotranspiration. Neural Computing and Applications 31: 5843-5855. https://doi.org/10.1007/s00521-018-3410-8

Mcgill R, Tukey J W & Larsen W A (1978). Variations of box plots. The American Statistician 32(1): 12–16. https://doi.org/10.1080/00031305.1978.10479236

Mehdizadeh S, Behmanes J & Khalili K (2017). Using MARS, SVM, GEP and empirical equations for estimation of monthly mean reference evapotranspiration. Computers and Electronics in Agriculture 139: 103-114. https://doi.org/10.1016/j.compag.2017.05.002

Mohsin S & Lone M A (2021). Modeling of reference evapotranspiration for temperate Kashmir Valley using linear regression. Modeling Earth Systems and Environment 7: 495-502. https://doi.org/10.1007/s40808-020-00921-8

Negm A, Minacapilli M & Provenzano G (2018). Downscaling of American National Aeronautics and Space Administration (NASA) daily air temperature in Sicily, Italy, and effects on crop reference evapotranspiration. Agricultural Water Management 209: 151-162. https://doi.org/10.1016/j.agwat.2018.07.016

Noi P T, Degener J & Kappas M (2017). Comparison of multiple linear regression, cubist regression, and random forest algorithms to estimate daily air surface temperature from dynamic combinations of MODIS LST Data. Remote Sensing 9(5): 398. https://doi.org/10.3390/rs9050398

Orange Data Mining (2024). Orange (Version 3.34) [Software]. https://orangedatamining.com/ Accessed 15 January 2024

Oudin L, Hervieu F, Michel C, Perrin C, Andréassian V, Anctil F & Loumagne C (2005). Which potential evapotranspiration input for a lumped rainfall–runoff model? Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling. Journal of Hydrology 303(1-4): 290-306. https://doi.org/10.1016/j.jhydrol.2004.08.026

Pereira L S, Allen R G, Smith M & Raes D (2015). Crop evapotranspiration estimation with FAO56: past and future. Agricultural Water Management 147: 4-20. https://doi.org/10.1016/j.agwat.2014.07.031

Prasad R, Deo R C, Li Y & Maraseni T (2017). Input selection and performance optimization of ANN-based streamflow forecasts in the drought-prone Murray Darling Basin region using IIS and MODWT algorithm. Atmospheric Research 197 (2017): 42-63. https://doi.org/10.1016/j.atmosres.2017.06.014, 2017

Rahimikhoob A (2010). Estimation of evapotranspiration based on only air temperature data using artificial neural networks for a subtropical climate in Iran. Theoretical and Applied Climatology 101: 83-91. https://doi.org/10.1007/s00704-009-0204-z

Niaghi A R, Hassanijalilian O & Shiri J (2021). Estimation of reference evapotranspiration using spatial and temporal machine learning approaches. Hydrology 8(1): 25. https://doi.org/10.3390/hydrology8010025

Reis M M, Silva A J, Junior J Z, Santos L D, Azevedo A M & Lopes E M (2019). Empirical and learning machine approaches to estimating reference evapotranspiration based on temperature data. Computers and Electronics in Agriculture 165: 104937. https://doi.org/10.1016/j.compag.2019.104937

Sabziparvar A A & Tabari H (2010). Regional estimation of reference evapotranspiration in arid and semiarid regions. Journal of Irrigation and Drainage Engineering 136: 724-731. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000242

Samadianfard S, Kargar K, Shadkani S, Hashemi S, Abbaspour A & Safari M J (2022). Hybrid models for suspended sediment prediction: Optimized random forest and multilayer perceptron through genetic algorithm and stochastic gradient descent methods. Neural Computing and Applications 34: 3033–3051. https://doi.org/10.1007/s00521-021-06550-1

Sarıgöl M & Katipoğlu O M (2024). Estimation of monthly evaporation values using gradient boosting machines and mode decomposition techniques in the Southeast Anatolia Project (GAP) area in Turkey. Acta Geophysica 2023: 1–18. https://doi.org/10.1007/s11600-023-01067-8

Sayyadi H, Oladghaffari A, Faalian A & Sadraddini AA (2009). Comparison of RBF and MLP neural networks performance for estimation of reference crop evapotranspiration. Water and Soil Science 19(1): 1-12

Seifi A & Riahi H (2020). Estimating daily reference evapotranspiration using hybrid gamma test-least square support vector machine, gamma test-ANN, and gamma test-ANFIS models in an arid area of Iran. *Journal of Water and Climate Change* 11(1): 217-240. https://doi.org/10.2166/wcc.2018.003

Shiri J, Kisi O, Landeras G, Lopez J J, Nazemi A H & Stuyt L (2012). Daily reference evapotranspiration modeling by using genetic programming approach in the Basque Country (Northern Spain). *Journal of Hydrology* 414: 302-316. https://doi.org/10.1016/j.jhydrol.2011.11.004

Shiri J (2017). Evaluation of FAO56-PM, empirical, semi-empirical and gene expression programming approaches for estimating daily reference evapotranspiration in hyper-arid regions of Iran. Agricultural Water Management 188: 101-114. https://doi.org/10.1016/j.agwat.2017.04.009

Shiri J (2018). Improving the performance of the mass transfer-based reference evapotranspiration estimation approaches through a coupled wavelet-random forest methodology. *Journal of Hydrology* 561: 737-750. https://doi.org/10.1016/j.jhydrol.2018.04.042

Shirzad A & Safari M J S (2019). Pipe failure rate prediction in water distribution networks using multivariate adaptive regression splines and random forest techniques. *Urban Water Journal* 16(9): 653-661. https://doi.org/10.1080/1573062X.2020.1713384

Srivastava A, Sahoo B, Raghuwanshi N S & Chatterjee C (2018). Modelling the dynamics of evapotranspiration using Variable Infiltration Capacity model and regionally calibrated Hargreaves approach. Irrigation Science 36(4-5): 289-300. https://doi.org/10.1007/s00271-018-0583-y

Sterkenburg T F & Grünwald P D (2021). The no-free-lunch theorems of supervised learning. Synthese 199: 9979-10015. https://doi.org/10.1007/s11229-021-03233-1

Tabari H, Kisi O, Ezani A & Talaee P H (2012). SVM, ANFIS, regression and climate based models for reference evapotranspiration modeling using limited climatic data in a semiarid highland environment. *Journal of Hydrology* 444: 78-89. https://doi.org/10.1016/j.jhydrol.2012.04.007

Tabari H, Grismer M E & Trajkovic S (2013a). Comparative analysis of 31 reference evapotranspiration methods under humid conditions. Irrigation Science 31: 107–117. https://doi.org/10.1007/s00271-011-0295-z

Tabari H, Martinez C, Ezani A & Talaee P H (2013b). Applicability of support vector machines and adaptive neurofuzzy inference system for modeling potato crop evapotranspiration. Irrigation Science 31(4): 575-588. https://doi.org/10.1007/s00271-012-0332-6

Taylor K E (2001), Summarizing multiple aspects of model performance in a single diagram, *Journal of Geophysical Research* 106(D7): 7183–7192, https://doi:10.1029/2000JD900719

Traore S, Wang Y M & Kerh T (2010). Artificial neural network for modeling reference evapotranspiration complex process in Sudano-Sahelian zone. Agricultural Water Management 97(5): 707-714. https://doi.org/10.1016/j.agwat.2010.01.002

Traore S & Guven A (2012). Regional-specific numerical models of evapotranspiration using gene-expression programming interface in Sahel. Water Resource. Management 26: 4367-4380. https://doi.org/10.1007/s11269-012-0149-3

Traore S & Guven A (2013). New algebraic formulations of evapotranspiration extracted from gene-expression programming in the tropical seasonally dry regions of West Africa. Irrigation Science 31: 1-10. https://doi.org/10.1007/s00271-011-0288-y

TSMS (2022). General directorate of state meteorology affairs. https://mevbis.mgm.gov.tr/ mevbis/ui/index.html#Workspace, accessed 15 March 2022

TSMS (2023). General directorate of state meteorology affairs. https://www.mgm.gov.tr/ veridegerlendirme/il-ve-ilceler-istatistik. aspx?k=unde- fined&m=ADANA, accessed 15 January 2024

TSMS (2024). General directorate of state meteorology affairs. https://www.mgm.gov.tr/FILES/iklim/iklim_siniflandirmalari/koppen.pdf, Accessed 12 January 2024.

Üneş F, Kaya Y Z & Mamak M (2020. Daily reference evapotranspiration prediction based on climatic conditions applying different data mining techniques and empirical equations. Theoretical and Applied Climatology 141(2020): 763-773. https://doi.org/10.1007/s00704-020-03225-0

Valipour M (2015). Investigation of valiantzas' evapotranspiration equation in Iran. Theoretical and Applied Climatology 121: 267-278. https://doi.org/10.1007/s00704-014-1240-x

Vapnik V (1995). Support-vector networks. Machine Learning 20: 273-297. https://doi.org/10.1007/BF00994018

Wang S, Lian J, Peng Y, Hu B & Chen H (2019). Generalized reference evapotranspiration models with limited climatic data based on random forest and gene expression programming in Guangxi, China. Agricultural Water Management 221: 220-230. https://doi.org/10.1016/j.agwat.2019.03.027

Wang J, Raza A, Hu Y, Buttar N A, Shoaib M, Saber K… & Ray R L (2022). Development of monthly reference evapotranspiration machine learning models and mapping of Pakistan-A comparative study. Water 14(10): 1666. https://doi.org/10.3390/w14101666

Willmott C J (1981). On the validation of models. Physical Geography 2: 184-194. https://doi.org/10.1080/02723646.1981.10642213

Yamaç S S (2021). Reference evapotranspiration estimation with kNN and ANN models using different climate input combinations in the semiarid environment. *Journal of Agricultural Sciences* 27 (2): 129-137. https://doi.org/10.15832/ankutbd.630303

Yassin M A, Alazba A A & Mattar M A (2016). Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate. Agricultural Water Management 163: 110-124. https://doi.org/10.1016/j.agwat.2015.09.009

Yildirim D, Küçüktopcu E, Cemek B & Simsek H (2023). Comparison of machine learning techniques and spatial distribution of daily reference evapotranspiration in Türkiye. Applied Water Science 13(4): 107. https://doi.org/10.1007/s13201-023-01912-7

Yirga S A (2019). Modelling reference evapotranspiration for Megecha catchment by multiple linear regression. Modeling Earth Systems and Environment 5: 471-477. https://doi.org/10.1007/s40808-019-00574-2

Yurtseven I & Serengil Y (2021). Comparison of different empirical methods and data-driven models for estimating reference evapotranspiration in semi-arid Central Anatolian Region of Turkey. *Arabian Journal of Geosciences* 14: 1-28. https://doi.org/10.1007/s12517-021-08150-8

Appendix 1- Applied model parameters and pseudo codes

Model Parameters for GEP	Model Parameters for RF	Model Parameters for ANN
Number of chromosomes: 30	Number of trees: 10	Hidden layers: 100
Head size: 7-8	Maximal number of considered features: unlimited	Activation: ReLu
Genes: 3	Maximal tree depth: unlimited	Solver: L-BFGS-B
Function set: +, -, *, /, Exp, Ln, X2, 3Rt, Inv	Stop splitting nodes with maximum instances: 5	Alpha: 0.0001
Linking function: addition	Replicable training: yes	Max iterations: 200
Fitness function: RMSE		Replicable training: yes
Mutation: 0.00138		
Inversion: 0.00546		
IS transposition: 0.00546	**Model Parameters for SVM**	**Model Parameters for MLR**
RIS transposition: 0.00546	SVM type: SVM, c=1.3, Ɛ = 0.5	Parameters: Fit intercept (unchecking it fixes it to zero)
Gene transposition: 0.00277	Kernel: RBF, exp(-auto|x-y|^2)	Regularization: No regularization
One-point recombination: 0.00277	Numerical tolerance: 0.001	
Two-point recombination: 0.00277	Iteration limit: 200	
Gene recombination: 0.00277		

Box Plot

1. Initialize necessary variables and parameters: - Define a list of column identifiers to read from the Excel file (pandas). - Define a list of labels corresponding to each dataset. - Initialize dictionary (data structure) to store the data and calculated statistics.

2. Read the Excel data: - Loop through the list of column identifiers: - For each identifier, read the corresponding column from the Excel file into a variable (pandas). - Store this data in a dictionary with the corresponding label as the key.

3. Convert the data to NumPy arrays: - Loop through the labels in the dictionary: - Convert each dataset to a NumPy array and store it back in the dictionary (NumPy).

4. Define a function to calculate statistics: - The function should take a dataset as input and return the following (NumPy): - Lower Quartile - Upper Quartile - Mean - Median - Standard Deviation

5. Calculate statistics for each dataset: - Loop through the labels in the dictionary: - Apply the statistics function to each dataset. - Store the results in a dictionary with the corresponding label as the key.

6. Print the calculated statistics: - Loop through the labels in the statistics dictionary: - For each label, print the corresponding statistics.

7. Plot the boxplots: - Initialize the plotting parameters (e.g., median line style, mean marker, etc.) (matplotlib). - Prepare the data for plotting by extracting each dataset from the dictionary. - Use the prepared data to create boxplots with labels for each dataset (matplotlib). - Set the y-axis label and range (matplotlib).
- Display the plot (matplotlib).

Error Box Plot

1. Read data from the Excel file (pandas):- Load data from columns into variables: ANN1, RF2, RF3, RF4, HS, CALJH

2. Convert the loaded data into NumPy arrays (NumPy): - For each dataset (ANN1, RF2, RF3, RF4, HS, CALJH), convert the DataFrame to a flattened NumPy array.

3. Define the properties for the boxplot (standard Python dictionary): - Set the median line properties with red color and a specific line style. - Define the properties for mean markers (black edge, blue face). - Set the properties for outlier markers (red edge, specific marker size).

4. Calculate statistics for each dataset (NumPy): - For each dataset (ANN1, RF2, RF3, RF4, HS, CALJH), calculate the lower quartile, upper quartile, mean, median, and standard deviation using the `calculate_statistics` function.

5. Plot the boxplots (matplotlib): - Set the figure size for the plot. - Define the y-axis ticks from 0 to 4.1 with intervals of 0.5. - Create a boxplot for each dataset (ANN1, RF2, RF3, RF4, HS, CALJH) with appropriate labels and the defined properties (whiskers, median, mean). - Label the y-axis and range (matlplotlib).

Taylor Diagram

1. Calculate correlation coefficients between datasets (NumPy): - Compute the correlation between ANN1 and FAO56-PM. - Compute the correlation between RF2 and FAO56-PM. - Compute the correlation between RF3 and FAO56-PM. - Compute the correlation between RF4 and FAO56-PM. - Compute the correlation between HS and FAO56-PM. - Compute the correlation between CALJH and FAO56-PM.

2. Extract the relevant correlation values from the correlation matrices (NumPy): - Extract the correlation value from the ANN1-FAO56-PM matrix. - Extract the correlation value from the RF2-FAO56-PM matrix. - Extract the correlation value from the RF3-FAO56-PM matrix. - Extract the correlation value from the RF4-FAO56-PM matrix. - Extract the correlation value from the HS- FAO56-PM matrix. - Extract the correlation value from the CALJH- FAO56-PM matrix.

3. Store the extracted correlation values in a list (standard Python list): - Create a list to hold the correlation values for ANN1, RF2, RF3, RF4, HS, and CALJH.

4. Define the `TaylorDiagram` class: - Initialize the class with standard deviation (STD), figure, rectangular plot area, and label parameters. - Set up the polar transform for correlation angles (NumPy and matplotlib).
- Define correlation labels and convert them to polar angles (NumPy). - Set up the grid for the diagram (matplotlib). - Create a subplot within the figure (matplotlib). - Configure the axis labels and directions:
- Set up the top axis for the correlation coefficient with turquoise color. - Set up the left and right axes for standard deviation. - Hide the bottom axis. - Draw grid lines on the subplot. - Add reference points and standard deviation contours to the plot (matplotlib). - Initialize a list to collect sample points.

5. Define `add_sample` method in `TaylorDiagram` class: - Plot a sample point on the Taylor diagram using its standard deviation and correlation coefficient. - Append the plotted sample point to the list of sample points.
6. Define `add_contours` method in `TaylorDiagram` class: - Create a meshgrid of standard deviation and polar angles (NumPy). - Calculate RMSE (Root Mean Square Error) values (NumPy). - Draw contour lines on the Taylor diagram to represent RMSE levels (matplotlib).
7. Define `add_samples` function: - For each dataset, plot a sample point on the Taylor diagram using its standard deviation, correlation coefficient, and specified color. - Customize the marker appearance (matplotlib).
8. Define `srl` function: - Create a new figure for the Taylor diagram (matplotlib). - Initialize the `TaylorDiagram` class with observed standard deviation and other parameters. - Add contour lines to the diagram with specified colors and linewidths. - Add sample points to the diagram using the `add_samples` function. - Create a legend for the plotted points. - Display the Taylor diagram (matplotlib).
9. Define the data for the diagram: - Set the observed standard deviation using statistics from the FAO56-PM dataset. - Define lists for the standard deviations, correlation coefficients, and labels for each dataset (ANN1, RF2, RF3, RF4, HS, Cal_JH). - Specify colors for each data point.
10. Call the `srl` function: - Pass the observed standard deviation, standard deviation list, correlation coefficient list, label list, color list, and file name to the function.

Correlation Matrix

1. Read data from the Excel file (pandas): - Load data from columns into variables Tmax, Tmin, RH, Rs, U2, FAO56-PM.
2. Convert the loaded data into NumPy arrays and flatten them (NumPy): - For each dataset (Tmax, Tmin, RH, Rs, U2, FAO56-PM), convert the DataFrame to a flattened NumPy array.
3. Combine the flattened arrays into a single 2D array (NumPy): - Stack the NumPy arrays column-wise into a single 2D array.
4. Calculate the correlation matrix for the combined data (NumPy): - Compute the correlation matrix using the combined 2D array, treating each column as a variable.
5. Loop through the correlation matrix to process each pair of correlations (standard Python loop): - For each pair of variables (i, j) in the correlation matrix
6. Create a figure for the heatmap (matplotlib): Set the figure size to 8 by 6 inches size.
7. Generate a heatmap to visualize the correlation matrix (seaborn): - Plot the correlation matrix as a heatmap using the following parameters: - Annotate each cell with the correlation coefficient value, formatted to 3 decimal places. - Use the "Pink-Green" color map for the heatmap - Set annotation font size to 15.
- Define x-axis and y-axis tick labels with specific variables. - Add a label to the color bar with the text 'Correlation Coefficient'.
8. Adjust the font size and rotation of the tick labels on the axes (matplotlib):
- Set the x-axis tick labels with no rotation and a font size of 10. - Set the y-axis tick labels with a 90-degree rotation and a font size of 10.
9. Customize the color bar (matplotlib): - Access the color bar from the heatmap and adjust its label font size to 12. - Adjust the tick label font size on the color bar to 12.
10. Display the heatmap (matplotlib): - Show the generated heatmap.

Time Series

1. Set the file path for the Excel file.
2. Read data from specified columns in the Excel file (pandas): - Columns into FAO56-PM, ANN1, RF2, RF3, RF4, HS, CALJH. - Columns into DAY, MONTH, YEAR.
3. Convert the data to flattened NumPy arrays (NumPy): - Flatten DAY, MONTH, YEAR into nDAY, nMONTH, nYEAR. - Flatten ANN1, RF2, RF3, RF4 into nANN1_num, nRF2_num, nRF3_num, nRF4_num. - Flatten HS, CALJH, FAO56-PM into nHS_num, nCALJH_num, nFAO56_PM_num.
4. Define the `DataRecord` class: - Initialize the class with attributes: value, day, month, and year. - Implement a `__repr__` method to return a string representation of the object, displaying the value, day, month, and year.
5. Create lists of `DataRecord` objects for each dataset: - For each dataset (ANN1, RF2, RF3, RF4, FAO56_PM, HS, CALJH): - Use list comprehension to create `DataRecord` objects by zipping together the corresponding values, days, months, and years. - Store the resulting list of `DataRecord` objects in the appropriate variable (e.g., `ANN1_records`, `RF2_records`, etc.).
6. Define the `plot_data` function: - Accept parameters: a list of `DataRecord` objects (`records`), a plot title (`title`), and an optional step for x-axis labels.
7. Sort the records by date: - Convert `day`, `month`, and `year` attributes of each record into a `datetime` object. - Sort the records by the `datetime` object. - Extract the formatted dates as strings and the corresponding values into separate lists.
8. Set up the plot (matplotlib): - Create a figure with specified dimensions. - Plot the sorted dates against the values with a blue line. - Set the plot title, x-axis label, and y-axis label.
9. Customize x-axis labels: - Identify dates that are the 1st of each month. - Format these dates as "Month Year" for x-axis labels. - Set the x-axis labels to the formatted dates with a 90-degree rotation.
10. Display the grid and the plot (matplotlib).
11. Use `plot_data` function to visualize different datasets: - Call `plot_data` for each dataset (`ANN1_records`, `RF2_records`, `RF3_records`, `RF4_records`, `FAO56_PM_records`, `HS_records`, `CALJH_records`) with appropriate titles and step values.

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Nitrogen Fertilization Effects on Oil Content, Sucrose, α- Tocopherol, Fatty Acid and Aminoacid Compositions of Confectionary Sunflower Seed

Öner Canavar[a], Hatice Kübra Gören[a*]

[a] *Aydin Adnan Menderes University, Faculty of Agriculture, Department of Field Crops, 09970, Aydin, TURKEY*

ARTICLE INFO
Research Article
Corresponding Author: Hatice Kübra Gören, E-mail: hkubra.goren@adu.edu.tr
Received: 29 April 2024 / Revised: 13 August 2024 / Accepted: 03 October 2024 / Online: 14 January 2025

Cite this article
Canavar Ö, Gören H K (2025). Nitrogen Fertilization Effects on Oil Content, Sucrose, α- Tocopherol, Fatty Acid and Aminoacid Compositions of Confectionary Sunflower Seed. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):230-241. DOI: 10.15832/ankutbd.1475152

ABSTRACT
This study investigates the effects of nitrogen fertilization on the composition of confectionary sunflower seeds, including parameters such as oil content, alpha-tocopherol, sucrose, amino acids and fatty acid composition. Nitrogen fertilization, surprisingly, had no significant effect on sunflower oil content, with remarkable differences observed between the genotypes. The α-tocopherol content, an important antioxidant, displayed responses that were dependent on the genotype upon application of nitrogen. The Somon Beyazı genotype consistently demonstrated higher oil content compared to the Ahmet Bey genotype. In addition, both genotypes showed a decrease in α-tocopherol levels as the application of nitrogen increased. Sucrose content was higher in the Somon Beyazı genotype and decreased significantly with increasing nitrogen doses. Significant variations were observed in fatty acid compositions, further emphasizing the impact of genotype and nitrogen application. The study also revealed diverse amino acid profiles, with notable concentrations of glutamine and asparagine. This comprehensive study highlights the complex interaction between genotype and nitrogen fertilization and provides valuable insights for optimising sunflower seed production and quality. The results emphasise the significance of integrating both genetic factors and nutrient management practices into crop cultivation for improved agricultural outcomes.

Keywords: Oil content, α-tocopherol, Sucrose, Amino acid profiles, Fatty acid composition

1. Introduction

Sunflower (*Helianthus annuus* L.), which has substantially wide range of adaptation with regard to farming production to varied latitudes, longitudes and photoperiods of the World, is one of the most a good candidate of oilseed crops for supply to the gap and demand of edible oil. Sunflower is used mainly in the edible oil industry and many other industries such as paint, cake, bread and in animal feeding meal and etc. Two types of sunflower are grown: oil seed purpose and non-oil seed sunflower for commercial market (Phanindra et al. 2018; Tan & Kaya 2019). Confectionary sunflowers are highly variable for morphological characters (Tan et al. 2016 and Tan et al. 2017). Non-oil seed sunflower is white or black striped, white or black color hull and comes from large-seeded varieties with low oil content and it is used in baking and snack applications. The improvement of nutrition facts and the quality of the oil is associated with its fatty acid composition and particularly the proportions of oleic, linoleic and linolenic acids (Jalilian et al. 2012). Here's a clearer version of the sentence:

Sunflower seeds are highly important due to their essential nutrients (Lie et al. 2017). As people increasingly demand healthier food options and are also suffering from cardiovascular diseases (Ryan et al. 2015), the nutritional value of sunflower seeds becomes even more significant. Consequently, cultivating promising genotypes with high yielding ability and applying favorable agricultural practices as planting date, irrigation, plant row spacing and nitrogen fertilization offer a great opportunity to improve seed yield and oil quality of sunflower seeds (Lie et al. 2017).

Although confectionary sunflower is a significant income earner in the World, it is generally considered together with oil sunflower in the world literature. Although it is evaluated separately for confectionary type and oil type on a country basis, confectionary sunflower statistics are not included in organizations in international agricultural organizations ISA, OECD, FAO and so on (Kaya et al.). Farmers have insufficient knowledge about breeding or some farm production techniques. This situation may cause low productivity and not reaching the desired level for high seed yield and quality in production. Poor quality and non-standard sunflower seed causes difficulties in processing and quality problems in the final product to be offered to the consumer. Appropriate cultivation techniques will be effective in terms of yield and quality in confectionary sunflower farming. One of the methods used in agriculture is nitrogen fertilization. It is possible to observe the effect of nitrogen fertilization in

sunflower physiologically and morphologically (Ghani et al. 2000; Valchovski 2002; Osman & Awed 2010; El Satar et al. 2017). For example, Scheiner et al. (2002) investigated nitrogen fertilization and the need for nitrogen in sunflower in their study in Pampas, Argentina. As a result of the research, the nitrogen fertilization increased the yield by 17% and nitrogen fertilization decreased the oil concentration in the seed. In this concern, Osman & Awed (2010) showed that 100 kg N ha^{-1} fertilization was suitable for sunflower seed and the higher rate 150 kg N ha^{-1} indicates a negative effect on the oil contents and seed yield.

Even though there is a classification rule in oil type sunflower in terms of sales and marketing, no any officially classification rules for confectionary sunflower type in many countries in the World. The most important problem in the production and marketing of confectionary sunflower is the lack of a certain quality standardization. When farmers market their products, they sell according to completely only physical and relative methods without any nutrition facts of seeds such as oil content, protein, amino acids, tocopherols, sugar content.

In this study, we conducted a research on the nutritional facts content of confectionary sunflower seed, which has two different shell colors, such as oil content, α-tocopherol, sugar ingredients, fatty acids and amino acid compositions under different nitrogen fertilization treatments in the field. In recent years, although there are a lot of studies, which explain the effects of nitrogen fertilization or non-fertilization on the nutrition facts of sunflower seeds, especially lacking of studies with regards to the changing of sucrose content and α-tocopherol under different nitrogen doses applications.

α-Tocopherol exhibits the maximum vitamin E activity (in vivo). But its in vitro activity is relatively low (Bramley et al. 2000). On the other hand, α-Tocopherol is a more effective free radical scavenger than α-Tocopherol in vitro (Duthie et al. 1991). These astonishing properties have made the tocopherols an incredibly essential nutrient and which must be present in edible oil in a significant amount. (Naz et al. 2011). The reports of Zhang et al. 2012 suggested that when the vitamin E (Toc) intake is lacked either from the diet or from supplements may reduce the risk of liver cancer. There may be associated between nitrogen and tocopherols (Hussein et al. 2014), because tocopherols and nitrogen are found mostly localizially in the same plant tissues, specifically in the plastids, a cell organelle containing chlorophyll and involved in photosynthesis (Gzyl-Malcher et al. 2010) and N is a major component of chlorophyll, amino acids, enhancing photosynthetic activity and other important biomolecules such as adenosine triphosphate (ATP) and nucleic acids (Wagner 2011).

The objectives of present study were to investigate how the effects of different nitrogen fertilization doses on oil content, tocopherol, sucrose, amino acids and fatty acid compositions of two different hull color confectionary sunflower seeds. In addation, we considered that the results of our study can provide a good data set in terms of chemical. nutrition facts of seeds for more comprehensive confectionary sunflower standardization studies that will be carried out in the future.

2. Material and Methods

2.1. Plant materials and the properties of experiment site

Somon Beyazı belongs to white hull color and Ahmetbey as a black hull color; confectionary sunflower genotypes were used in the experiment as a plant material. Two confectionary sunflower varieties were provided by the Trakya Agricultural Research Institute in Turkey. This research was conducted in the Aegean Region, with an altitude of 820m and between 38° 5 'north latitude 29° 36' east longitude. The location of the experimental area showed the transition between the Mediterranean climate dominant in the Aegean and the continental climate of Central Anatolia. The soil properties of field experiment showed Table 1. Average annual precipitation is 409.56 mm. The highest rainfall is in April and the lowest in August during the summer growing season. The values of climate conditions in field experiment were given in Figure 1.

Table 1- The soil properties of field experiment

Properties	***Results***	***Rating***
The texture of soil	*Silty Clay Loam*	
pH	8.23	Alkali
% Total Salt	0.020	Low
% Lime	38.67	Silty Clay Loam
% Sand	16.83	
% Silty	44.50	
% Organic matter	1.93	Low
% Marl	46.44	Extensive
Phosphorus (P) ppm	5.16	Low
Potassium (K) ppm	291.1	Very low
Calcium (Ca) ppm	4685	High
Magnesium (Mg) ppm	528.8	Very high
Iron (Fe) ppm	2.65	Critical level
Manganese (Mn) ppm	3.68	Enough suitable
Zinc (Zn) ppm	1.19	Enough suitable
Cuper (Cu) ppm	3.68	Enough suitable
Boron (B) ppm	0.41	Very low

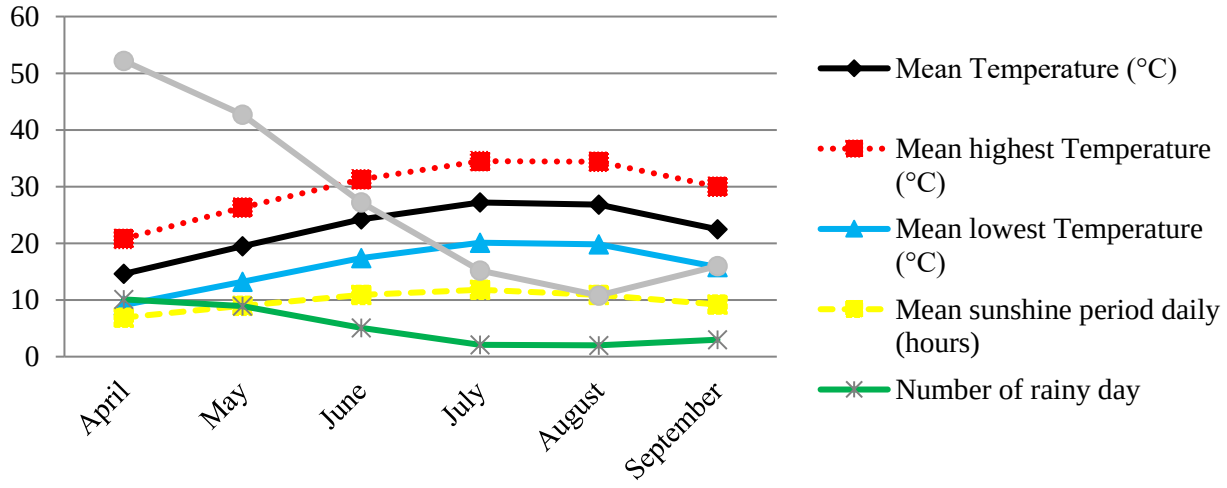


Figure 1- The climate conditions of study (long term means 1957 – 2019)

2.2. The application of fertilizer and design of field experiment

In the experiment, the main plots were set up with different nitrogen fertilizer doses of 0, 60, 120, 180 kg N ha^{-1}. *First fertilization during soil tillage before planting;* the application of different nitrogenous fertilizer dosage, which is one of the factors (main-plot) in the experiment, was carried out with the split-plot methods. Nitrogen fertilization process was applied in two different periods. During the planting preparation of the tillage soil, only for 60, 120 and 180 kg ha^{-1} nitrogen fertilization plots were fertilized by calculating Ammonium Sulphate (AS 21% N) fertilizer to provide 60 kg ha^{-1} Nitrogen (N) doses. 640 gr AS fertilizer was calculated and applied to each parcels (22.4 m^2). No nitrogenous fertilization was applied to the parcels to be treated with 0 kg ha^{-1} of nitrogen.

In addition, before planting in the experimental area, it was calculated also and applied from Triple Superphosphate (TSP 43-44% P_2O_5) fertilizer to provide 60 kg ha^{-1} pure Phosphate (P) fertilization for all parcels. 315 gr TSP fertilization was calculated and applied on all parcels of study. Potassium Sulphate (50% K_2SO_4) was also calculated and applied as 60 kg ha^{-1} pure Potassium (K) fertilization. 270 gr Potassium Sulphate fertilization was applied on all parcels.

Second fertilization before flowering of plants; Second fertilization was not applied to parcels with 0 and 60 kg N ha^{-1}. 295 gr UREA (46 % N) fertilization was fertilized on the parcels manured with 60 kg N ha to obtain 120 kg ha^{-1} Nitrogen doses application. 584 gr UREA was applied on the parcels manured with 120 kg ha^{-1} N to obtain 180 kg ha^{-1} Nitrogen doses.

The sunflower seeds were planted with a four-row pneumatic seed drill, with a distance between rows of 0.7 metres and a plant distance of 0.2 metres. 25 days after planting, the intra-rows were thinned with an inter-row as 40 cm (plant density was 3.5 plant m^{-1}). Plots consisted of four rows 8 m long × 2.8 m wide. The harvest area was 19.6 m^2 in each parcel. Furrow irrigation was applied and hand weeding was performed throughout the growing period when necessary.

2.3. Collecting plant samples

The seeds with shelled were randomly selected from sunflower plants in the middle each plot without any size classification during harvesting. The rate of sampling was approximately represented by 15 % of the seed yield in each parcel. The shell of sunflower seeds were immediately crushed by hand nippers to avoid contamination without waiting after harvested. Unshelled seeds were immediately stored at – 20 °C until analysis.

2.4. The analysis of oil content

Soxhlet method was used to determine the oil ration of unshelled confectionary sunflower seeds (Tesan et al. 2022)). After 250 mL balloons were tarred. 10g sample grounded was taken and surrounded with filter paper. Then 150 mL hexane was added and put into Soxhlet tubes. The Soxhlet tubes were immersed in the Soxhlet device for 4 hours. To remove left hexan in the tubes were put into oven for 30 min at 105 °C.

2.5. Determination of fatty acids composition using gas chromatography (GC)

First, the extracted crude sunflower oils underwent gas chromatography (GC) analysis following esterification. For sample preparation, 10 mg of extracted sunflower oil was mixed with 0.1 mL of 2 M methanolic potassium hydroxide (KOH) and blended. Then, the solution was mixed with 2 mL of isooctane. The final mixture was spun at 300 mL/minute for 5 minutes. One hundred µl was extracted from the upper layer using a micro-syringe and subsequently injected into the Gas Chromatography with Flame Ionization Detection (GC-FID) (Model X, Manufacturer Y).

To determine the peaks in the chromatogram, a standard mixture comprising each fatty acid was injected into the GC-FID, and their corresponding retention times (RT) were identified. Technical term abbreviations were always explained when used for the first time. The SP TM-256 Column, Macherey-Nagel, Duren, Germany, with a film thickness of 0.2 µm on a 100 m × 0.25 mm ID column was used. Helium served as the carrier gas and had a linear velocity of 1 mL m–1, while the split ratio was set at 40:1. The starting temperature of the column was set at 140 °C, with an air flow rate of 400 mL s–1 and an H_2 flow rate of 40 mL s–1, sustained for 10 minutes. Subsequently, the column temperature was raised by 4 °C per minute until it reached 240 °C. The fatty acid composition was analysed by computing the area percentage of each peak (Reed et al. 2004).

2.6. The analysis of sugar components (Glucose. Fructose. Maltose. Sucrose)

A 5 g portion of finely ground sunflower seeds was weighed and mixed with 40 mL of deionised water to dissolve in a 150 mL tube. Following this, 25 mL of methanol was added to the solution. Ultra-pure deionised water was then added to fill the remaining 100 mL of the solution. The mixture was passed through a 0.45 µm PTFE filter. Finally, 20 µL samples of the extract were injected into HPLC for analysis. A calibration curve was prepared by injecting a mixture of sugar components, comprised of Glucose, Fructose, Sucrose, and Sorbitol, at specific concentrations into the device.

The sugar components were analyzed using HPLC (with an RID detector) and the results were determined by substituting the values for the sugars in the sample onto the calibration curve. Technical abbreviation explanations were included where first used. The analysis was carried out using a mobile phase of Asetontril:dH2O (80:20) (V: V) with a column of NH2 Colon 5 µm, 4.6x250 mm operating at a flow rate of 1.5 mL/min at 30 °C.

2.7. The analysis of α-tocopherol using high performance liquid chromatography

α-Tocopherol content was determined using HPLC-FLD. Approximately 0.5 g of the sample was weighed and diluted to 10 mL with hexane, then vortexed. After appropriate dilution, 50 µL of the solution was injected into a Shimadzu HPLC system (mod. LC-10ADVP) with an FLD detector (mod. SPDM10AVP) and a reversed-phase column (Si 60, 5 µm, 4.0 x 250 mm). The detector was set to an excitation wavelength of 290 nm and an emission wavelength of 330 nm. The mobile phase, consisting of hexane: 2-propanol (99.5:0.5), flowed at 0.8 mL/min with a column temperature of 30 °C.

Along with the sample, the standards of α-tocopherol prepared at different concentrations were injected into the device and the calibration curve was obtained. Quantification was performed by substituting the peak area value of α-tocopherol from the sample into the calibration curve.

2.8. Amino acid profiles of seeds

The amino acid composition of the prepared sunflower seed was determined using a High Performed Liquid Chromatograph (Shimadzu Nexara XR, HPLC system). 0.1 g ground samples were hydrolysed by adding 5 mL 6 N HCL, 250 µL 2 mM Phenol, 0.1 g Na_2SO_3. The solution was waited for 24 hours at 110 °C. The pH of the sample was adjusted to the range of 6.7 – 7.3 with 5 N NaOH and then centrifuged at 4000 rpm for 5 min at room temperature.

The solvent used (for derivatization) were: Borat tampon (Diluent); [Boric acid + KOH, concentration (0.4 M 100 mL, pH: 10.2], OPA solution; [OPA + Diluent (10 mg OPA, 100 µL MeOH, 900 µL Diluent, 10 µL 3-MPA) 10 mg mL^{-1} (1mL)], FMOC solution; [ACN (2.5 mg FMOC, 1mL ACN) 2.5 mg mL^{-1})].

2.9. Statistical analysis

The statistical analysis of the study investigated the effect of different nitrogen doses on two sunflower hybrid varieties using a split-plot design with three replications. In addition, the differences between sunflowers of different hull colours and the interaction of nitrogen x genotype were assessed using a two-way split-plot ANOVA for all parameters. This was carried out using the statistical software JMP Pro16 software (SAS Institute, Cary, NC, USA) and TARIST (Açıkgöz et al., 1994). Significant differences in the means of replications were evaluated by means of Fisher's least significant difference (LSD) approach. Regression analyses were performed using Microsoft Excel to assess the correlation between the oil content, α-tocopherol, sucrose of two genotypes, and nitrogen fertilisation doses (see Figure 1). All differences reported were statistically significant at 0.05.

3. Results and Dıscussıon

3.1. Effects of nitrogen fertilization on oil content

The effect of nitrogen fertilization was not statistically found on the oil content of sunflower seed (Table 2). It was found that there was a significantly difference in terms of oil content of two different hull color confectionary sunflower (Table 2). Generally, the oil content of Somon Beyazı (white hull color) confectionary sunflower seed was higher than that of Ahmet Bey genotype under all nitrogen fertilization doses (Figure 2). Especially, two different confectionary sunflower showed differences with regards to the changing of oil content of seeds under 60 kg N ha^{-1} fertilization (Figure 2). Somon Beyazı genotype showed an increase in oil content at a dose of 60 kg nitrogen fertilization, whereas Ahmet Bey genotype decreased. But the oil content of Somon Beyazı was decreased with subsequent increasing doses of nitrogen fertilization (Figure 2). Figure 2 showed that the oil content of two confectionary genotypes were decreased with the increasing nitrogen doses fertilization, even though there was not found statistically any differences the effect of nitrogen fertilization in field experiment. These results are consistent with those of Abdel-Sabour and AboEl-Seoud (1996), Nanjundappa et al. (2001), Munir et al. (2007), and Nasim et al. (2012), who observed decreases in oil percentages with increased N application. The oil content of the Somon Beyazı genotype decreased non-linearly by 1.1%, from 38.9% to 37.8%, as nitrogen fertilization increased from 0 kg N ha^{-1} to 180 kg N ha^{-1}. In contrast, the oil content of the Ahmet Bey genotype decreased linearly by 2.0%, from 36.8% to 34.8%, with increasing nitrogen fertilization doses. This might be due to the varying genetic potential of the hybrids. The effect of the year was also not significant. These results are consistent with the findings of Roche et al. (2010) and Bukhsh et al. (2011), who found that different sunflower variety exhibit distinct responses to oil content due to differences in their genetic makeup. In addition to Steer et al. (1984) reported that excessive amounts of nitrogen can decrease the concentration of sunflower oil as a result of increased protein concentration, without any parallel increase in biomass as evidenced.

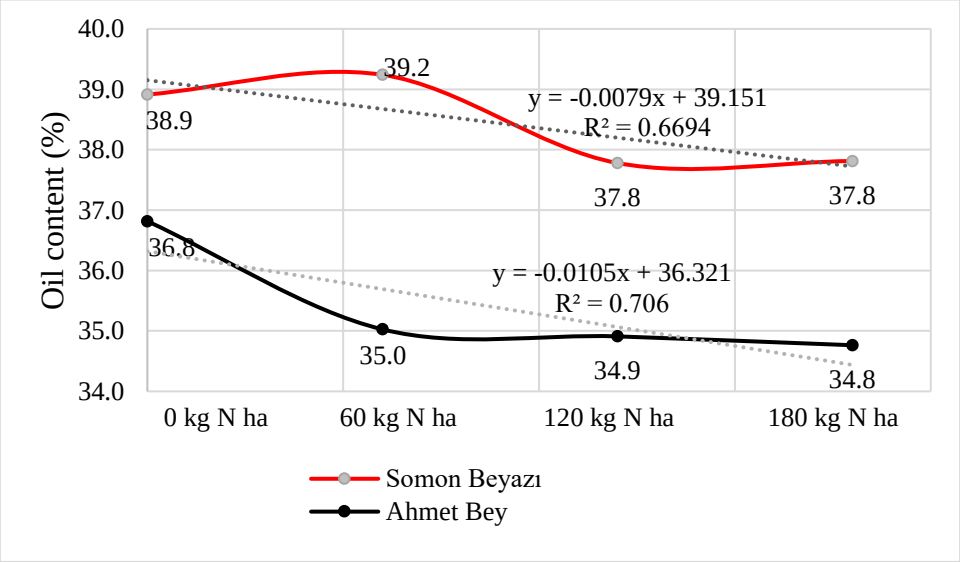


Figure 2- The effect of nitrogen on the oil content of two different hull colored confentionary sunflower seed

Table 2- The ANOVA results of oil content, α-Tocopherol and sucrose content of two confectionary sunflower under different nitrogen fertilization conditions

Variance Sources	*df*	*The Values of Mean Square*		
		Oil content (%)	*α-Tocopherol*	*Sucrose*
Replicate	2	2.0094	2783.51	0.02275
N Fert.	3	3.16594ns	43688.3ns	0.0023ns
Error 1	6	1.1609	15023.3	0.02852
Genotype	1	56.1204**	223494**	1.6485ns
N x G	3	1.32176ns	47802.6*	0.01029***
Error 2	8	3.60058	9092.9	0.028921
$LSD_{N <0.05}$		ns	ns	ns
$LSD_{G <0.05}$		2.727	2.908	ns
$LSD_{NxG <0.05}$		ns	5.257	0.335

*: P<0.05; **: P<0.01; ***: P<0.001, ns:non-significant

3.2. Effects of nitrogen fertilization on α-tocopherol

There were significant effects of genotypes and nitrogen fertilization × genotype for α-tocopherol content (Table 2). There were significantly differences between the α-tocopherol content of two genotype in all nitrogen fertilization doses (Table 2). Considering mean values, Ahmet Bey significantly higher content of α-tocopherol by 256.9 µg g^{-1} in 0 kg N ha^{-1}, 280.8 µg g^{-1} in 60 kg N ha^{-1}, 207,5 µg/g in 120 kg N ha^{-1} and 27.4 µg g^{-1} in 180 kg N ha^{-1} fertilization treatments than that of Somon Beyazı genotype. It was observed important considerable decrease of α-tocopherol content in both sunflower genotype with the increasing nitrogen fertilization doses (Figure 3). The reduction in α-tocopherol content was markedly 22.04% and 20.25% in the 120 kg N ha^{-1} nitrogen fertilization in Ahmet Bey and Somon Beyazı sunflower genotypes respectively. The study showed that application of different concentrations of nitrogen fertiliser can affect α-tocopherol levels in both sunflower genotypes to different rates.

Tocopherols play an important role in the quality of oil due to their antioxidant activity, which is more effective at relatively low concentrations (Beltrán et al. 2005). On the other hand, when α-tocopherol is present at relatively higher concentrations, a pro-oxidative effect can occur (Belizit & Grosch, 1986). In our study, there was no significant effect of the dose of nitrogen on the ratio of tocopherols. The difference observed was due solely to the genotypes, and the responses to the different doses of N were not uniform.

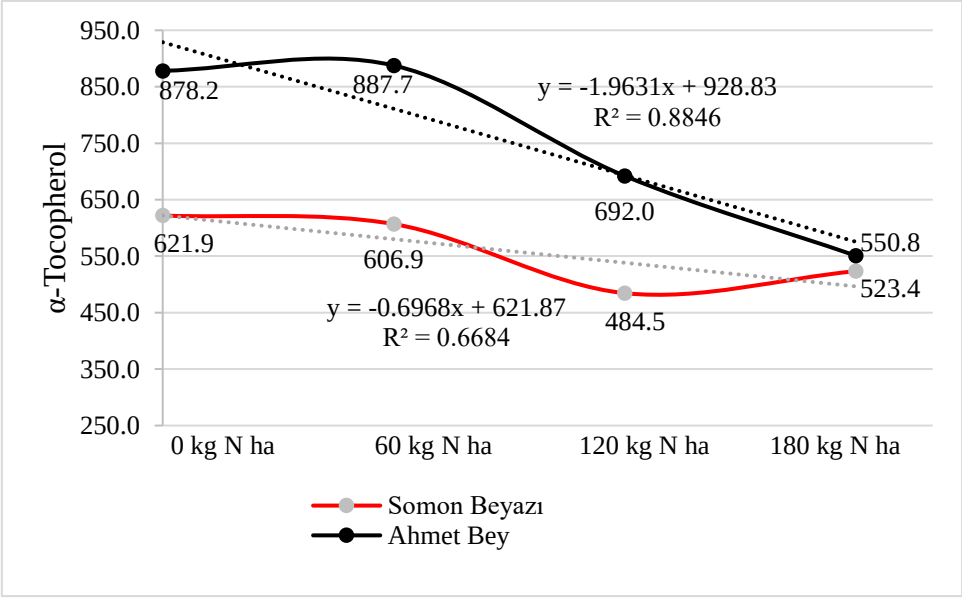


Figure 3- The effect of nitrogen on the α-Tocopherol of two different hull colored confentionary sunflower seed

3.3. Effects of nitrogen fertilization on sucrose

Figure 4 displayed that the change in sugar content of both genotypes in the application of different nitrogen dose fertilizers. The sugar content of the Somon Beyazı genotype showed higher values than the sugar content of Ahmet Bey sunflower genotype in

all nitrogen dose fertilizations (Figure 4). Furthermore, sucrose content of the two sunflower cultivars under different nitrogen doses fertilization had decreased by 19.77% at 15 DAF (Days After Flowering), 32.44% at 30 DAF, and 48.37% at 45 DAF, respectively, an average decrease of 33.53%.

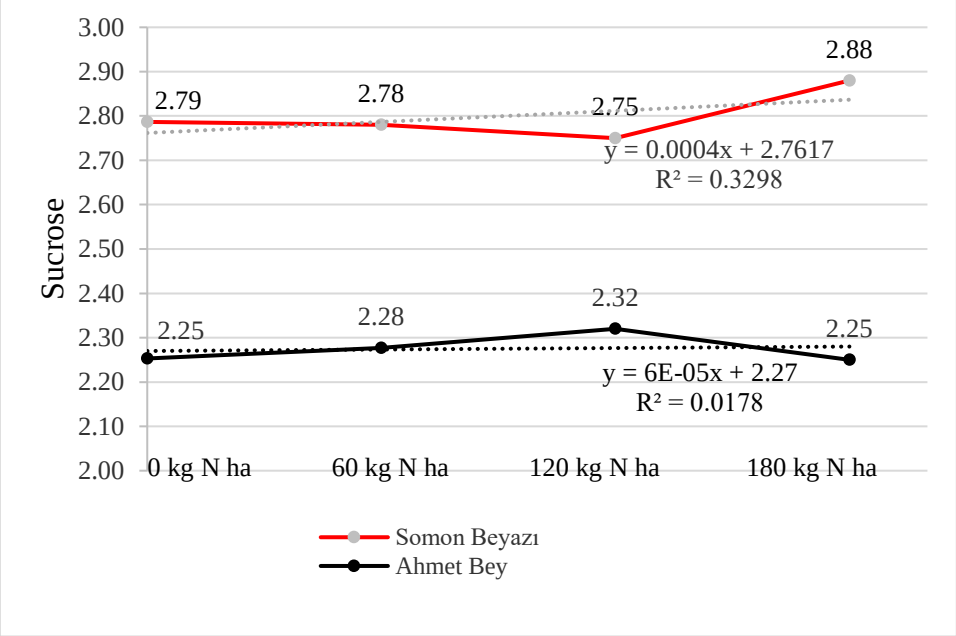


Figure 4- The effect of nitrogen on the sucrose content of two different hull colored confentionary sunflower seed

3.4. Effects of nitrogen fertilization on fatty acid compositions

Results of fatty acid composition for two confectionery sunflower genotypes under different nitrogen fertilization are shown in Table 5. Myristic acid, stearic acid, oleic acid, linoleic acid, gondoic acid, linolenic acid, behenic acid and lignoceric acid showed significant differences between genotypes. Moreover, there were significant differences in arachidic acid and triconoic acid for both N dose and genotypes, and significant effects of nitrogen fertilization × genotype interaction were observed for palmitic acid and gondoic acid. However, no statistical difference was found in palmitoleic acid and its parameters.

Table 5- ANOVA results of fatty acids composition of two confectionary sunflower genotype under different nitrogen fertilization

Variance Sources	*df*	*The Values of Mean Square*						
		Miristik Acid	*Palmitic Acid*	*Palmitoleic Acid*	*Heptadeconoic Acid*	*Stearic Acid*	*Oleic Acid*	*Linoleic Acid*
Replicate	2	9.04e-6	0.00031	0.0272	0.01395	0.00692	0.40744	0.37079
N Fert.	3	0.00002	0.01299**	0.02861	0.01421	0.05729	1.30025	1.54579
Error 1	6	5.87e-6	0.00092	0.0282	0.01431	0.02869	0.50338	0.35549
Genotype	1	0.00056***	0.70761***	0.03096	0.0187	0.9165***	35.9611***	71.3633***
N x G	3	0.00003	0.00934*	0.02861	0.01477	0.03923	1.0603	1.07065
Error 2	8	0.000010	0.001645	0.028573	0.014361	0.015964	0.52275	0.45873
$LSD_{N<0.05}$		ns	0.0141	ns	ns	ns	ns	ns
$LSD_{G<0.05}$		0.0058	0.0432	ns	ns	0.0574	0.6879	1.5556
$LSD_{NxG<0.05}$		ns	0.0561	ns	ns	ns	ns	ns

*: P<0.05; **: P<0.01; ***: P<0.001

Table 5 (Continue)- ANOVA results of fatty acids composition of two confectionary sunflower genotype under different nitrogen fertilization

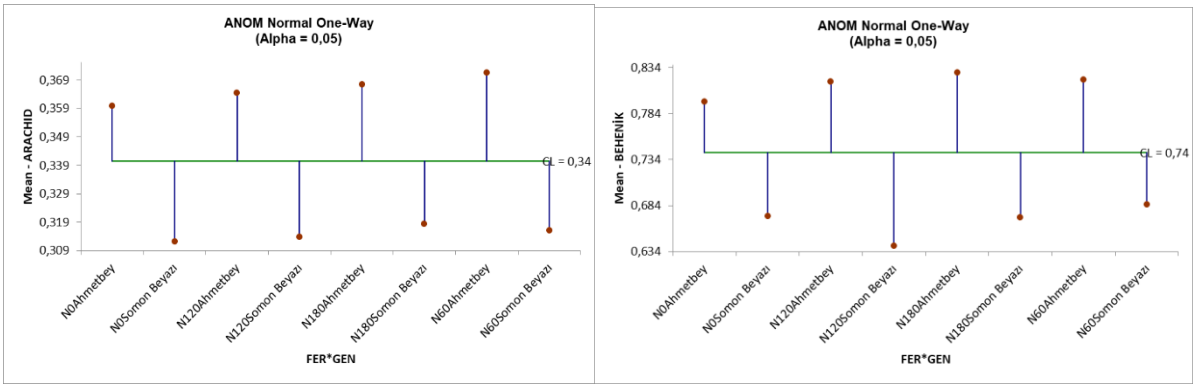
Variance Sources	*df*	*The Values of Mean Square*					
		Arachidic Acid	*Gondoic Acid*	*Linolenic Acid*	*Behenic Acid*	*Triconoic Acid*	*Lignoseric Acid*
Replicate	2	3.88e-6	3.79e-6	1.79e-6	0.00017	6.67e-7	0.00009
N Fert.	3	0.00008**	7.28e-6	8.04e-6	0.00078	2.94e-6*	0.00007
Error 1	6	6.82e-6	6.9e-6	3.29e-6	0.0011	4.44e-7	0.00002
Genotype	1	0.0154***	0.0035 ***	0.00081***	0.13336***	0.00001*	0.0373***
N x G	3	0.00002	0.00008 **	0.00001	0.00085	1.17e-6	0.00002
Error 2	8	0.000027	0.000007	0.000003	0.001053	0.0000012	0.000043
$LSD_{N<0.05}$		0.1155	ns	ns	ns	ns	ns
$LSD_{G<0.05}$		0,0387	0.0005	0.0028	0.0126	0.0010	0.0867
$LSD_{NxG<0.05}$		ns	0.0001	ns	ns	ns	ns

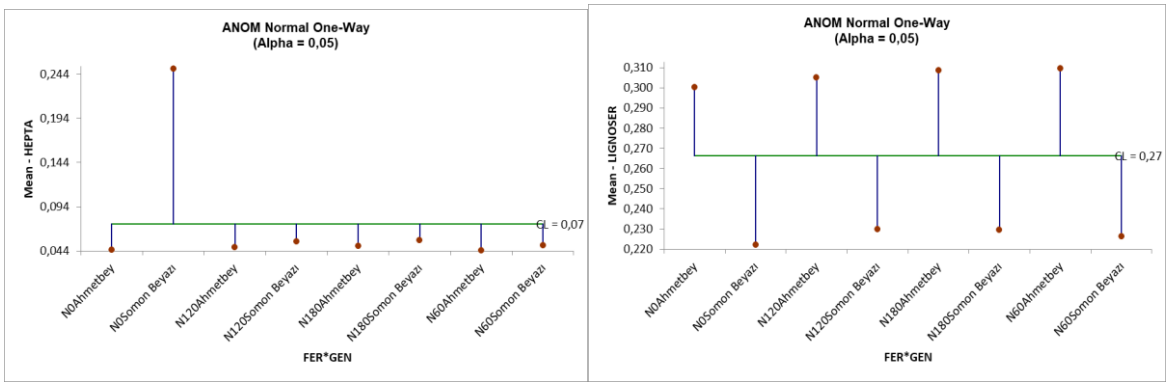
*: P<0.05; **: P<0.01; ***: P<0.001

The fatty acid composition analysis of two confectionery sunflower genotypes under varying levels of nitrogen fertilization is presented in Figure 5. The genotypic response to fertilization doses in the ANOM is visually illustrated in a decision chart resulting from mean comparisons via the Analysis of Mean (ANOM) method. When analysing the fatty acid compositions for different fertilizer doses, it was observed that the Ahmetbey variety had values higher than both the overall mean and the Somon Beyazı genotype for all nitrogen doses. Notably, highest values for fatty acids such as Myristic, Palmitic, Stearic, Linolenic, Tricona, and Lignoceric were recorded at a nitrogen dose of 180 kg/ha-1.

Palmitoleic, Heptadecenoic and Linoleic acids responded negatively to N application, whereas Oleic acid reached its highest concentration at 16 kg/ha-1 of nitrogen. These results are supported by Nanjundappa et al. (2001), Munir et al. (2007), and Boydak et al. (2010), who observed decreases in the composition of this fatty acid with increased N application.

On the other hand, arachidic, gadoleic and behenic acids showed their maximum values at a nitrogen dose of 120 kg/ha^{-1}.





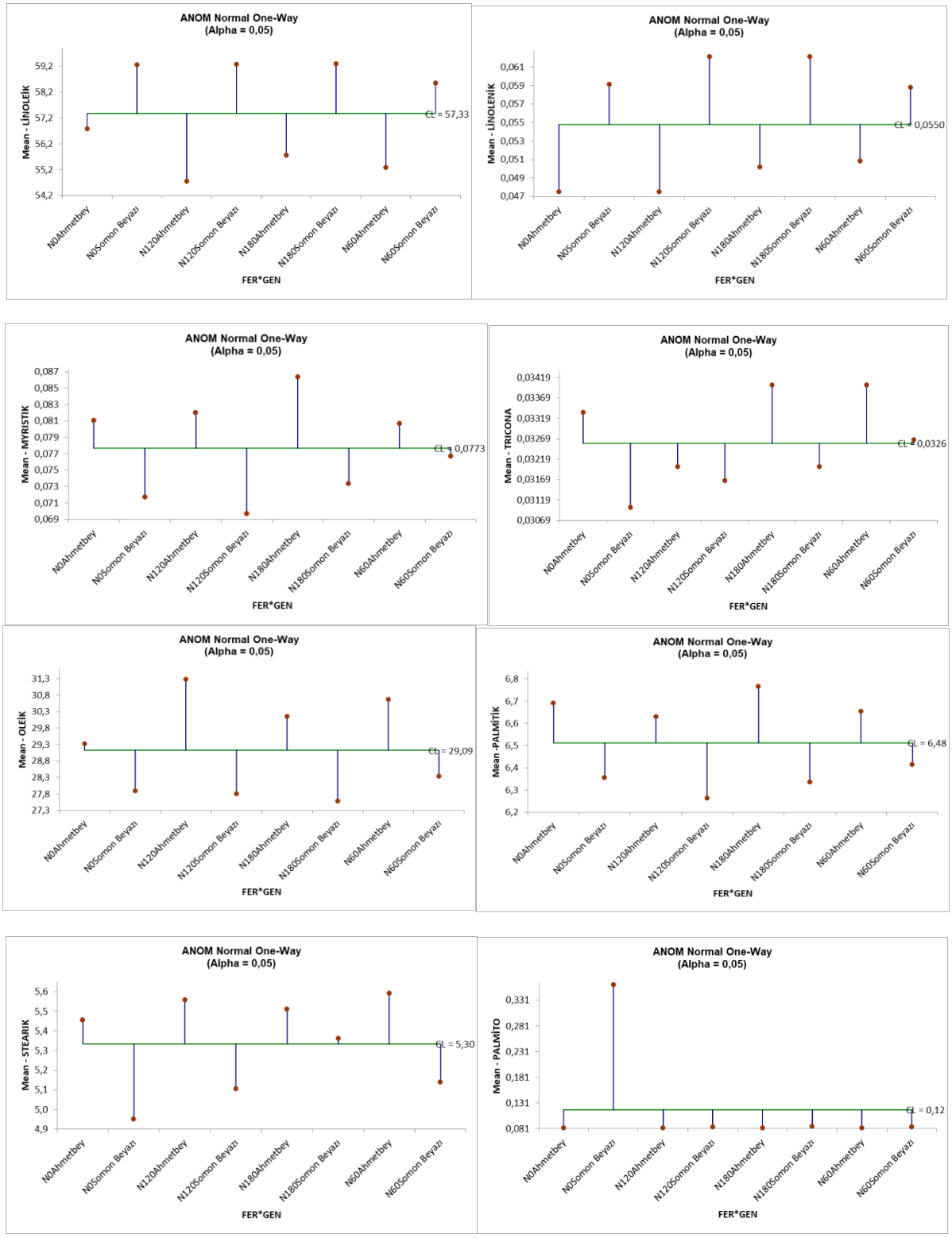


Figure 5- The Anom (Analysis of Means) of fatty acids composition of two confectionary sunflower genotype under different nitrogen fertilization

3.5. Effects of nitrogen fertilization on amino acid profiles

The ANOVA results for the amino acid profile of two confectionary sunflower varieties under nitrogen fertilization doses are presented in Table 3. The analysis revealed significant interactions between N dose*genotype for the levels of glutamine, serine, histidine, glycine, arginine, alanine, phenylalanine, lysine, leucine, and tyrosine. While differences were observed in Tryptophan, Valine, Isoleucine, and proline values across nitrogen doses, no statistically significant variation was found for Asparagine and Cysteine values (Table 4). The analysis demonstrates a diverse distribution of amino acids, with 16 different types detected at varying concentration levels. Particularly noteworthy is the high content of glutamine and Asparagine, followed by cysteine,

arginine, glycine, serine, lysine, valine, and phenylalanine, all exceeding 1 gram per 100 millilitres. Additionally, the presence of Histidine, Tryptophan, alanine, isoleucine, leucine, proline, and tyrosine was also confirmed. Glutamine, serine, histidine, glycine, arginine, alanine, phenylalanine, lysine, leucine, and tyrosine. The optimum nitrogen dose was determined as 120 kg N ha-1 and the variety with black hull colour was higher than the variety with white hull colour in all treatments. The distribution of amino acid composition to nitrogen doses was a quadratic distribution and this result scientifically revealed that the most suitable nitrogen dose was 120 kg N ha^{-1} (Table 4).

The use of reductive amination and transamination processes produces a greater amount of amino acids compared to fatty acids (Prasad 1990). This shift in metabolic focus has implications for oil content, as Singh & Quadri (1984) and Sharma & Gaur (1988) reported that high levels of nitrogen lead to a decrease in oil content (%) due to an increase in impurities. Consequently, managing nitrogen levels is crucial to balancing amino acid and fatty acid production, as well as maintaining oil quality. Protein synthesis is entirely reliant on the amount of nitrogen fertilizer accessible for plant use, as stated by Nasim et al. (2012). The use of nitrogen fertilizer also leads to an increased concentration of protein regardless of the irrigation method utilized, as noted by Jalilian et al. (2012).

Table 3- ANOVA results of aminoacids profile of two confectionary sunflower under different nitrogen fertilization conditions

Variance Sources	*df*	*F values*							
		Asp	*Glu*	*Ser*	*His*	*Gly*	*Thr*	*Arg*	*Ala*
Replicate	2	1.031	0.219	0.105	1.333	0.281	1.435	1.867	0.018
N Fert.	3	1.506ns	101.882**	334.629**	34.001**	124.350**	90.387**	29.617**	55.265**
Error 1	6	0.024	0.004	0.000	0.000	0.000	0.001	0.001	0.000
Genotype	1	2.053ns	0.631ns	4.012ns	0.127ns	0.033ns	14.003**	0.267ns	0.529ns
N x G	3	0.808ns	19.085**	32.887**	9.485**	37.224**	3.044ns	18.285**	21.678**
Error 2	8	0.025	0.006	0.000	0.000	0.001	0.001	0.001	0.000
$LSD_{N<0.05}$		ns	0.089	0.008	0.015	0.019	0.054	0.053	0.019
$LSD_{G<0.05}$		ns	ns	ns	ns	ns	0.029	ns	ns
$LSD_{NxG<0.05}$		ns	0.141	0.016	0.025	0.059	ns	0.065	0.032

*: P<0.05; **: P<0.01; ***: P<0.001. The values of errors were given as the mean of square. N: Nitrogen fertilization. G: Genotype. Asp: Asparagine. Glu: Glutamine. Ser: Serine. His: Histidine. Gly: Glycine. Thr: Tryptopthan. Arg: Arginine. Ala: Alanine

Table 3(continue)- ANOVA results of aminoacids profile of two confectionary sunflower under different nitrogen fertilization conditions

Variance Sources	*df*	*F values*							
		Cys	*Val*	*Phe*	*IsLe*	*Lys*	*Leu*	*Pro*	*Try*
Replicate	2	0.724	1.510	0.082	0.397	0.086	5.232	8.951	2.319
N Fert.	3	2.865ns	12.760**	18.250**	8.401*	22.958**	134.508**	507.960**	57.114**
Error 1	6	0.020	0.001	0.001	0.001	0.001	0.000	0.000	0.000
Genotype	1	0.317ns	1.103ns	0.116ns	0.004ns	0.238ns	0.629ns	7.948*	0.010ns
N x G	3	3.972ns	3.123ns	19.304**	1.369ns	12.762**	9.439**	3.902ns	9.254**
Error 2	8	0.016	0.001	0.000	0.002	0.001	0.000	0.003	0.000
$LSD_{N<0.05}$		ns	0.035	0.054	0.034	0.043	0.015	0.014	0.010
$LSD_{G<0.05}$		ns	ns	ns	ns	ns	ns	0.049	ns
$LSD_{NxG<0.05}$		ns	ns	0.015	ns	0.050	0.040	ns	0.027

*: P<0.05; **: P<0.01; ***: P<0.001. The values of errors were given as the mean of square. N: Nitrogen fertilization. G: Genotype. Cys: Cysteine. Val:Valine. Phe: Phenylalanine. Le: Leucine. Lys: Lysine. IsLeu: Isoleucine. Pro: Proline. Try:Tyrosine

There was a significant increase in protein contents with the increased N rates. Amjed and Ulah (2012) recorded protein contents for highest (225 kg N ha^{-1}) and lowest (15.33%) in N0 (control). There were similar results in 2011, which are in line with those of Nanjundappa et al. (2001) and Munir et al. (2007), who observed increased protein due to N application. Sunflower hybrids had a significant effect on protein content in achene. 'Ahmetbey' attained significantly higher protein content than 'Somon Beyazı' which might be due to varying genetic potential of the hybrids (Table 4) These results are in accordance with the findings of Roche et al. (2010) and Bukhsh et al. (2011) that different sunflower hybrids exhibit the differential response to protein content in due to their difference of genotype.

Table 4- The changing of amino acids profiles in two different hull color confectionary sunflower seeds by different nitrogen fertilization

Amino Acids (g 100m^{-1})	*0 kg N ha^{-1}*		*60 kg N ha^{-1}*		*120 kg N ha^{-1}*		*180 kg N ha^{-1}*	
	Somon Beyazı	*Ahmet Bey*	*Somon Beyazı*	*Ahmet Bey*	*Somon Beyazı*	*Ahmet Bey*	*Somon Beyazı*	*Ahmet Bey*
Asparagine	2.081	2.109	2.001	2.067	2.150	2.200	2.168	2.161
Glutamine	4.881	4.921	4.705	4.894	5.261	5.486	5.351	4.993
Serine	1.075	1.069	1.023	1.083	1.113	1.131	1.087	0.958
Histidine	0.480	0.475	0.466	0.502	0.526	0.544	0.517	0.476
Glycine	1.477	1.494	1.409	1.544	1.567	1.648	1.598	1.376
Tryptophan	0.720	0.734	0.708	0.647	0.598	0.509	0.419	0.367
Arginine	1.876	1.886	1.838	1.937	2.008	2.091	1.965	1.802
Alanine	0.990	0.996	0.952	1.002	1.027	1.048	0.990	0.893
Cysteine	2.052	2.119	1.878	2.087	2.176	2.019	2.338	2.103
Valine	1.071	0.987	1.008	1.032	1.086	1.116	1.059	1.025
Phenylalanine	1.018	1.042	1.009	1.020	1.135	1.178	1.070	0.983
Isoleucine	0.785	0.758	0.745	0.769	0.804	0.842	0.811	0.772
Lysine	1.361	1.367	1.322	1.378	1.455	1.488	1.401	1.284
Leucine	0.707	0.759	0.741	0.794	0.766	0.727	0.672	0.653
Proline	0.872	0.883	0.861	0.828	0.950	1.074	0.952	1.091
Tyrosine	0.550	0.541	0.537	0.581	0.574	0.577	0.542	0.502

4. Conclusıons

The study investigated how nitrogen fertilization influences the composition of sunflower seeds, focusing on sugar, amino acids, fatty acids, oil content, and α-tocopherol levels across different genotypes and nitrogen doses. It found that nitrogen fertilization did not significantly affect oil content, although variations between genotypes were observed, with the Somon Beyazı genotype showing higher oil content compared to Ahmet Bey. Both genotypes exhibited reduced α-tocopherol levels with increased nitrogen application, but the Ahmet Bey genotype maintained consistently higher α-tocopherol levels. Sucrose content also declined with higher nitrogen doses in both genotypes, with Somon Beyazı having higher sucrose levels than Ahmet Bey. Fatty acid compositions differed notably among genotypes and in response to nitrogen fertilization, indicating the influence of both genetic factors and nutrient management. Additionally, the amino acid profiles revealed significant variations, with high levels of glutamine and asparagine. Overall, the study provides valuable insights into the intricate relationships between genotype and nitrogen fertilization, suggesting that optimizing cultivation practices can enhance the nutritional value and quality of sunflower seeds while emphasizing the importance of integrating genetic and nutrient management factors in crop cultivation.

Acknowledgments

This study was financially supported by ZRF-20022 of Scientific Research Projects Unit at Aydın Adnan Menderes University. We thank to Agrefar – Environmental food analysis laboratory for their assistance in conducting the analysis.

References

Abdel-Sabour M F & Abo El-Seoud M A (1996). Effects of organic-waste compost addition on sesame growth, yield, and chemical composition. *Agriculture, Ecosystems & Environment* 60(2-3): 157-164

Açıkgöz N, Akba ME, Moghaddam A, Özcan K (1994). Turkish data-based statistics programmer for PC. *Turkey Field Crops Congress*, Ege University Press, pp. 264-267

Ali A & Ullah S (2012). Effect of nitrogen on achene protein, oil, fatty acid profile, and yield of sunflower hybrids. *Chilean Journal of Agricultural Research* 72(4): 564-569

Beltrán G, Aguilera M P, Del Río C, Sanchez S & Martínez L (2005**).** Influence of fruit ripening process on the natural antioxidant content of Hojiblanca virgin olive oils. *Food Chemistry* 89(2): 207-215

Boydak E D, Karaaslan K & Türkoğlu H (2010). The effect of different nitrogen and irrigation levels on fatty acid composition of peanut oils. *Turkish Journal of Field Crops* 15: 29-33

Bramley P M, Elmadfa I, Kafatos A, Kelly F J, Manios Y, Roxborough H E, Schuch W, Sheehy P J A & Wagner K H (2000). Vitamin E. *Journal of the Science of Food and Agriculture* 80: 913-938

Duthie G G, Gonzalez B M, Morrice P C & Arthur J R (1991). Inhibitory effects of isomers of tocopherol on lipid peroxidation of microsomes from vitamin E-deficient rats. *Free Radical Research Communications* 15: 35-40

Ghani A, Hussain M & Qureshi M S (2000). Effect of different irrigation regimens on the growth and yield of sunflower. *International Journal of Agriculture and Biology* 2: 334-335

Gonzalez Perez P et al. (2015). Characterization of the amino acid composition of soils under organic and conventional management after addition of different fertilizers. *Journal of Soils and Sediments* 15: 890-901

Jalilian J, Modarres-Sanavy S, Saberali S F & Sadat-Asilan K (2012). Effects of the combination of beneficial microbes and nitrogen on sunflower seed yields and seed quality traits under different irrigation regimes. *Field Crops Research* 127: 26-34

Munir M A, Malik M A & Saleem M F (2007). Impact of integration of crop manuring and nitrogen application on growth, yield, and quality of spring planted sunflower (*Helianthus annuus L.*). *Pakistan Journal of Botany* 39: 441-449

Nanjundappa G, Shivaraj B, Janarjuna S & Sridhara S (2001). Effect of organic and inorganic sources of nutrients applied alone or in combination on growth and yield of sunflower (*Helianthus annuus L.*). *Helia* 24(34): 115-120

Nasim W A, Ahmad A, Bano R, Olatinwo M & Usman T (2012). Effect of nitrogen on yield and quality of sunflower (*Helianthus annuus L.*) hybrids under sub-humid conditions of Pakistan. *American Journal of Plant Sciences* 3: 243-251

Naz S, Sherazi S T H & Talpur N (2011). Changes of total tocopherol and tocopherol species during sunflower oil processing. *Journal of the American Oil Chemists' Society* 88: 127-132

Osman E B A & Awed M M M (2010). Response of sunflower to phosphorus and nitrogen fertilization under different plant spacing. *Assiut University Bulletin for Environmental Research* 13: 11-19

Phanindra K S, Toprope V N & Thakur N R (2018). Genetic variability in mutant confectionary sunflower (*Helianthus annuus L.*). *Journal of Research ANGRAU* 46(2): 15-20

Prasad D T (1990). Proteins of the phenolic extracted sunflower meal: I. Simple method for removal of polyphenolic components and characteristics of salt-soluble proteins. *LWT-Food Science and Technology* 23: 229-235.

Roche J, Alignan M, Bouniols A, Cerny M, Mouloungi Z, Vear F & Merah O (2010). Fatty acid profile composition in sunflower seeds as affected by genotypes and environmental conditions. *Food Chemistry* 121: 990-995.

Sharma S K & Gaur B L (1988). Effect of levels and methods of nitrogen application on seed yield and quality of sunflower. *Indian Journal of Agricultural Sciences* 33: 330-331

Singh C N & Quadri S J S (1984). Note on response of sunflower to rates and time of nitrogen application. *Journal of Agricultural Research* 10: 76-77

Steer B T, Hocking P J, Kortt A A & Roxburgh C M (1984). Nitrogen nutrition of sunflower (*Helianthus annuus L.*): Yield components, the timing of their establishment, and seed characteristics in response to nitrogen supply. *Field Crops Research* 9: 219-236

Tan A S & Kaya Y (2019). Sunflower (*Helianthus annuus L.*) genetic resources, production, and research in Turkey. *OCL* 26: 21

Tan A S, Aldemir M & Altunok A (2017). Yield potentials of some confectionery sunflower genotypes under ecological conditions of Menemen, Izmir. *ETAE Dergisi / Anadolu Journal of AARI* 1: 1-16 ((In Turkish)

Tan A S, Altunok A & Aldemir M (2016). Oilseed and confectionary sunflower (*Helianthus annuus L.*) landraces of Turkey. In *19th International Sunflower Conference* (pp. 556-566). Edirne, Turkey

Valchovski I (2002). Influence of heavy rate of nitrogen fertilizer on oil content and fatty acid composition of different varieties and hybrids. *Rasteniev dni Nauk* 39(516): 338-341

Wagner S C (2011). Biological nitrogen fixation. *Nature Education Knowledge* 2(11): 14

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Changes in the Major Antioxidant Compounds of Red Cabbage Under Water Stress Applied at Different Vegetative Growth Periods

Okan Erken[a*], Murat Yildirim[a], Bayram Kizilkaya[b], Umut Mucan[a]

[a]*Department of Irrigation and Farm Structures, Faculty of Agriculture, Canakkale Onsekiz Mart University, Canakkale, TURKEY*

[b]*Department of Aquaculture, Faculty of Marine Science and Technology, Canakkale Onsekiz Mart University, Canakkale, TURKEY*

ARTICLE INFO

Research Article

Corresponding Author: Okan Erken, E-mail: oerken@comu.edu.tr

Received: 11 March 2024 / Revised: 29 June 2024 / Accepted: 01 July 2024 / Online: 14 January 2025

Cite this article

Erken O, Yildirim M, Kizilkaya B, Mucan U (2025). Changes in the Major Antioxidant Compounds of Red Cabbage Under Water Stress Applied at Different Vegetative Growth Periods. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):242-251. DOI: 10.15832/ankutbd.1451013

ABSTRACT

Effective and efficient use of water resources has become an important issue in recent studies, where the impacts of climate change has become more apparent and alternative solutions are discussed. There, however, are limited studies that look at the impacts of water stress at different vegetative periods. For this reason, in this study, different levels of water treatment (0%, 30%, 70%) were applied to red cabbage at two stages of development (early and late vegetative) in a two-year field study. The effect of water stress on the major antioxidant compounds, as well as on yield and some morphological parameters were investigated.

According to the findings, the least yield loss (22%) occurred in the early vegetative period of the second-year trial where 70% irrigation water was applied, while the highest yield loss (56%) was obtained during the early vegetative period of the first-year trial where no irrigation was applied. Biochemical analyses revealed that the highest accumulation of flavonoids, 0.83 mg g^{-1}, and anthocyanins, 1.51 mg g^{-1}, occurred in the early vegetative period with the trial that received no irrigation treatment. The phenolic compound content was determined as 1.62 mg g^{-1}, and the antioxidant capacity was found to be 1.93 mg g^{-1} during the late vegetative period in the trials without irrigation treatment. These findings suggest that in regions with limited water resources, water conservation can be practiced during different vegetative periods in order to get higher biochemical benefits with a lower yield loss when cultivating red cabbage.

Keywords: Vegetative period, Antioxidant, Anthocyanin, Flavonoids, Phenolic compounds

1. Introduction

Food security is a major public concern, which has gotten worse with the increase of the population and pressure caused by environmental stress. Abiotic stress aggravated by the impacts of climate change has further influenced the yield and quality of crops (Haghighi et al. 2020) threatening food security. Approximately 40% of the world's available land is estimated to be affected by drought because of the climate change (Zhang et al. 2014). In summer months, particularly in arid and semi-arid regions, agricultural crops often face water shortages and erratic rainfall, leading to a significant loss of crops. The necessary amount of water required by these crops cannot be met, resulting in adverse effects on agricultural productivity (Kusvuran & Abak 2012). There is a need, therefore, for improvements in irrigation scheduling to maximize productivity under limited water conditions (Dhungel et al. 2023; Semiz et al. 2023).

Water shortage imposes several constraints on agricultural production unless precisely managed through conservation methods (Saha et al. 2021). In the future, there will be even greater pressure on freshwater resources with the additional need to provide food for future generations, in which the world population is predicted to increase by 65% (3.7 billion) by 2050 (Wallace 2000). In addition to the adverse effects of water shortages on crops, there is an increasing consumer demand for food that can be used as functional ingredients that promise to provide better well-being and health (Janabi et al. 2020). According to Siro et al. (2008), consumers are increasingly aware of the direct contribution of food on health.

Brassica species, which are among the fibrous and easily digestible vegetables, are widely consumed vegetables worldwide known for their nutritional benefits, of which vegetables such as cauliflower, broccoli, brussel sprouts and kale possess both antioxidant and anticarcinogenic properties (Cohen et al. 2000; Chu et al. 2002). Red cabbage, in particular, is a vegetable that offers a variety of beneficial effects on human health. It contains vitamins and minerals and is commonly consumed as a raw vegetable salad (Majkowska & Wierzbicka 2008). Red cabbage is also a valuable source of acylated anthocyanins that exhibit

potent antioxidant activity. It can be utilized as a natural food colorant and may contribute to the prevention of diseases linked to oxidative damage (Wiczkowski et al. 2013).

Red cabbage also has been shown to prevent oxidative stress in the livers and brains of animals exposed to paraquat, as demonstrated in study by Igarashi et al. (2000). Additionally, red cabbage includes various bioactive substances like anthocyanins, flavanols and glucosinolates. These bioactive compounds are known to have a positive impact on human health (Podsedek 2007; Volden et al. 2008; Lobos et al. 2017).

Red cabbage is a cool-season vegetable that is grown in the summer season in semi-arid regions such as Turkey. The changes in climate patterns in the summer season require the more effective management of the water resources while maximizing the nutritional benefits of the crops. Looking at previous research, there are some studies that investigate the impact of water stress on the yield parameters of red cabbage (Hajiboland & Amirazad 2010; Beacham et al. 2017; Shinde et al. 2020; Kishor et al. 2023). Other studies have investigated the impacts of water stress on some of the biochemical properties of the plant (Şahin et al. 2018; Haghighi et al. 2020; Erken 2022).

There, however, is no known study that investigates the impact of water stress applied at different growth stages on the yield, quality and some biochemical properties of red cabbage. This study is a two-year field study that investigates the effect of water stress applied at different growth stages of red cabbage (*Brassica oleracea* L var. Rubra). The more specific objective of this paper is to observe the variation of yield, quality, and some biochemical compounds such as flavonoids, anthocyanins, and antioxidant activity, in red cabbage under water stress applied during different growth stages of the plant.

2. Material and Methods

2.1. Experiment site

In this study, field experiments were carried out both in the summer of 2021 and 2022, where crops were harvested at the end of each year. Experimental plots were set up at the Dardanelles Research Extension Station of Çanakkale Onsekiz Mart located in Dardanos, Çanakkale. The soil at the experiment site was clay-loam with a water holding capacity of 167.7 mm at 90 cm depth. The rainfall data (Table 1) throughout the growth period was taken from the meteorology station at approximately 10 km from the experiment site.

Table 1- Meteorological data for the experiment period of 2021, 2022 and long-term average (MGM 2023)

Months	***Average temperature (ºC)***			***Rainfall (mm)***			***Evaporation (mm)***	
	2021	***2022***	***1940-2022***	***2021***	***2022***	***1940-2022***	***2021***	***2022***
August	28.3	26.2	25.1	0.0	111.6	11.3	479.2	491.9
September	23.1	22.0	21.1	8.9	1.2	24.1	359.7	400.5
October	18.1	17.6	16.2	75.9	0.6	55.2	206.0	301.3
November	15.8	14.8	12.2	26.7	33.2	83.2	156.2	276.8
December	12.2	11.5	8.4	121.0	43.2	103.2	211.1	209.2

The total amount of rainfall from August to December in 2021 was 232.5 mm, while in 2022, the total amount decreased to 189.8 mm (Table 1). During the second year, a significantly higher amount of rainfall occurred during the seedling period, while the total rainfall was lower in the other growing stages of red cabbage. This difference in rainfall amounts offered an important opportunity for comparison.

2.2. Experimental design

The plots were laid out in a randomized complete block design with three replications. Seven different levels of irrigation treatment were applied, of which the experiment layout is provided in Figure 1 for the entire trial. While one plot consisted of a control treatment, three plots were irrigated with 70%, 30% and 0% water in the early vegetative (EV) period and the remaining three plots received 70%, 30% and 0% water in the late vegetative (LV) period.

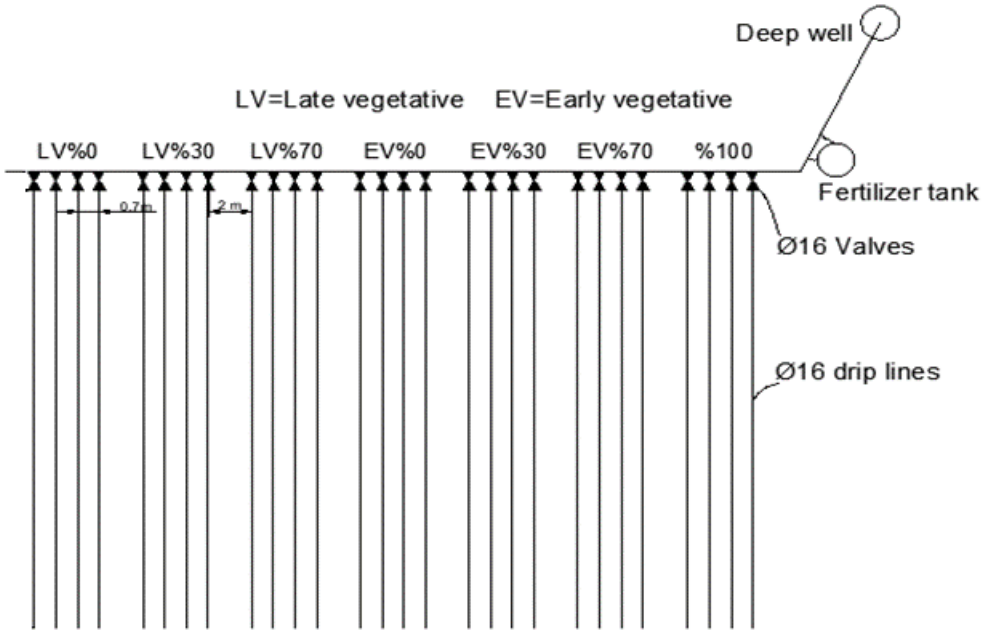


Figure 1- The experimental layout indicating the drip lines for all treatments

In the experiment, a total of 588 seedlings were planted in 7 plots, with 28 seedlings in each repeated application and 84 seedlings in each plot. Both the replications and the plots were kept 2 m apart and the seedlings were planted with 70 cm x 33 cm spacing. The red cabbage seedlings were first transplanted on August 7, 2021 and July 29, 2022. The EV period covered the dates between September 24, 2021 and December 22, 2021 in the first year and lasted from September 7, 2022 to December 13, 22 in the second year of the experiment. The LV period lasted from December 23, 2021 to November 23, 2021 in the first year and from December 14, 2022 to November 14, 2022 in the second year. All water treatments ended at the end date of the LV period. Crops were harvested on January 5 in 2022 and January 9 in 2023.

All experimental plots were applied with the same amount of fertilizer of 10 kg N, 5 kg P and 10 kg of K per decare (1 decare=1000 m^2). Fertilization was carried out three times; first at planting, next in 20 days and last in 15 days after the second application.

Each row was set up with a single drip line where the emitter had a normal discharge of 4L h^{-1} under the pressure of 1-1.5 atm. The irrigation water used in the experiment had an electrical conductivity (ECw) of 0.941 dS m^{-1}, which was measured with an EC59 meter (Martini Institute). This value was in the moderately tolerable range and had previously been used at the site for irrigation purposes.

All experimental plots were irrigated with 4-day intervals throughout the experiment. Following transplantation, each plot was irrigated equally for 20 days until the early vegetative period in order to allow for equal root development.

The irrigation amount was estimated using the following equation (Ertek & Kanber 2000):

$$I = A \times Epan \times Kcp \times P$$

Where: I, irrigation water amount (mm); A, the area of plot (m^2); Epan, the cumulative evaporation at irrigation intervals (mm); Kcp, the crop-pan coefficient; P, the percentage of wetted area (%)

2.3. Yield and quality parameters

Yield and quality parameters were measured from the plants harvested from the center of each treated plot. While a digital balance (±0.01 g) was used to weigh the plants, a digital clipper (±0.01 mm) was used to measure the quality parameters and a refractometer was used to determine the soluble solids. Following the measurement of the fresh plants, random samples from each plant were oven dried for 48 hours at 70 °C.

2.4. Determination of some biochemical properties

Total flavonoids, phenolic content, antioxidant capacity and the anthocyanin content of the red cabbage plants were estimated from fresh samples refrigerated at -18 °C after harvest. All analysis were carried out at the Laboratory of the Center for Plant

and Herbal Products Research-Development located in Istanbul University using international standard (ISO, AOAC) methods.

2.5. Total flavonoids

The flavonoid concentrations were estimated using an aluminium based colorimetric assay as explained by Shraim et al. (2021). 100 µL of leaf extract was mixed with equal amounts of 1 M potassium acetate and 10% aluminium nitrate, as well as 4.4 mL of 96% ethanol, which were then incubated at dark room temperature for 40 minutes. The absorbance values were read at 415 nm using a UV-vis spectrophotometer (Thermo Aquamate). The total flavonoid content was expressed in "mg g^{-1} FW."

2.6. Total phenolic content

The high-throughput assay to assess the Folin-Ciocalteau reducing capacity as described by Magalhaes et al. (2010) was used to measure total phenolic content. 50 µL of gallic acid standard solution and 50 µL of Folin-Ciocalteau reagent diluted in water (1:5, v/v) were mixed. 100 µL of sodium hydroxide solution (0.35 M) was introduced into the mixture. Absorbance was measured at 760 nm at 1-minute intervals until it reached its maximum value (optimally achieved in 3 minutes). In order to assess the intrinsic absorption of the sample, the Folin-Ciocalteau reagent was replaced with 50 µL of 0.4 M acid solution and the reagent blank was prepared with 50 µL of water instead of the standard solution. The total phenolic content of the samples was quantified as GAE (gallic acid equivalents) in mg per g FW (fresh weight).

2.7. Antioxidant capacity

The antioxidant capacity of the plant samples was determined using two methods: the cubric reducing antioxidant capacity (CUPRAC) assay and the 1.1-diphenyl-2-picrylhydrazyl (DPPH) free radical scavenging activity assay.

The CUPRAC assay, developed by Apak et al. (2004) utilizes an electron-transfer method to assess the plant samples' ability to reduce cupric ions (Cu^{+2}). Using this assay, we incubated 40 µL of the plant extract with 0.01 M copper(II) chloride (1 mL), 7.5 × 10–3 neocuproine (1 mL), 1 M pH 7 ammonium acetate (1 mL), and distilled water (1060 µL). After incubating the mixture for 30 minutes at 20 °C, the final volume of the mixture was 4100 µL. The absorbance measured at 450 nm and the antioxidant activity of fresh weight of leaves was quantified using a standard calibration curve, which was expressed in mg trolox 100 g^{-1} FW.

The DPPH radical scavenging activity was evaluated following the method outlined by Brand-Williams et al. (1995) and Ak & Türker (2018). Methanol was used to dilute frozen plant samples until a concentration of 1 mg mL^{-1} were reached. Subsequently, the diluted sample (0.1 mL) was thoroughly blended with the DPPH solution (3.9 mL) at room temperature for 30 minutes to facilitate the reaction between the plant extracts and DPPH. Following the incubation, the absorbance of the samples was measured using a spectrophotometer at 515 nm. The absorbance measurement provided an indication of the level of DPPH radical scavenging activity exhibited by the plant extracts.

The equation used to estimate DPPH radical scavenging activity is provided below:

$$DPPH\ scavenging\ (\ \%\ ) = \left[(Acontrol - Asample) / Acontrol \right] \times 100$$

In this equation, Asample stands for the absorbance of the sample once it reaches a plateau after leaving still for15 minutes. Acontrol stands for the absorbance of DPPH alone. The IC50 inhibition values, which showed the concentration of the compounds that inhibit 50% of the total DPPH radicals, were estimated. Lower IC50 inhibition values indicated higher antioxidant activity.

2.8. Anthocyanin content

The total anthocyanin content of the samples was estimated with the pH differential method (Giusti et al. 1998; Benvenuti et al. 2004). Following this method, two buffer systems were utilized for anthocyanin extraction (0.025 M potassium chloride (pH 1.0) and 0.4 M sodium acetate (pH 4.5)). Anthocyanin extraction was conducted for all four parallel samples obtained from each treatment plot. Each extraction was diffused with a pH 1.0 and a pH 4.5 buffer separately and incubated for 20 minutes in room temperature. This extraction process was replicated three times.

The absorbance measurements were taken at 520 nm and 700 nm using a UV-vis spectrophotometer. The absorbance values and the anthocyanin content were estimated using the following equations: [Please provide the relevant equations for absorbance and anthocyanin content as they were not provided in the initial statement.

$$A = (A520 - A700)\ pH1.0 - (A520 - A700)\ pH4.5$$

$$Monometric\ anthocyanin\ pigment \left(\frac{mg}{lt} \right) = (A \times MW \times DF \times 1000) / (\varepsilon \times 1)$$

In this equation, the symbol A represents the absorbance of the diluted sample, MW stands for the molecular weight, DF denotes the dilution factor, and ε is the molar absorptivity. The MW and ε values for the specific compound cyanidin-3-glucoside were 449.2 and 26,900, respectively.

2.9. Statistical analysis

Statistical analysis was made using the Duncan's Multiple Range Test. Differences were considered statistically significant at the probability level of 5% ($P<0.05$) which the SPSS statistical package software were used to analyse the data.

3. Results

3.1. Effects of water stress on yield

The amount of water used for irrigation and the corresponding fresh head weights for two consecutive years (2021 and 2022) based on the measurements of pan evaporation are provided in Table 2. The corresponding fresh weights of the water stressed red cabbage heads differed statistically significantly from control plants in both years.

Table 2- The amount of irrigation water applied and corresponding yield

Treatments	*Irrigation water (mm)*		*Fresh head weight (g/plant)*	
	2021	*2022*	*2021*	*2022*
$I_{1.0}$	433	495	1514±118.7^{a}	1252.6±41.6^{a}
$EV_{0.7}$	406	444	818.8±41.6^{bc}	972.3±115.5^{ab}
$EV_{0.3}$	370	376	717.2±152.7^{bc}	708±184.6^{b}
$EV_{0.0}$	343	325	661.5±174.4^{c}	650.6±145.6^{b}
$LV_{0.7}$	420	466	975.9±57.4^{b}	923.8±226.9^{b}
$LV_{0.3}$	413	439	951.7±169.4^{b}	769.7±49.5^{b}
$LV_{0.0}$	410	419	825.1±181.7^{bc}	765.5±119.5^{b}

*: Significant differences between treatments in both years are indicated by different letters ($P<0.05$).

According to the results, when water restriction was applied during different growth stages, the yield still decreased significantly in comparison to the fully irrigated ($I_{1.0}$-control treatment) plants. The highest yield, therefore, were obtained from the control plants with 1514 g plant^{-1} in 2021 and 1252.6 g plant^{-1} in 2022. The irrigation water applied during these periods were 433 mm in 2021 and 495 mm in 2022, respectively.

When the meteorological data for the first year of the experiment is examined (Table 1), it can be observed that the autumn season had above-average rainfall. The reason for achieving higher yields with less water usage in the first year could be attributed to the relatively abundant rainfall during the autumn season.

The plant quality indicators measured at the end of harvest are provided in Tables 3 and 4. According to the data, there was no significant difference in diameter development among different treatments in 2021 (Table 3), while this was not the situation in the following year (Table 4). The likely reason for this is that the rainfall was higher in 2021 when compared to 2022, resulting in a less pronounced impact of drought created by irrigation treatments on diameter development.

Table 3- Quality parameters of red cabbage in 2021

Treatments	***Diameter-x (cm)***	***Diameter-y (cm)***	***Height (cm)***	***Circumference (cm)***	***Leaf Number***	***Total Soluble Solid (%)***	***Dry weight Percent (%)***
$I_{1.0}$	12.83^{ns}	12.59^{ns}	16.51^{a}	45.70^{a}	29.40^{ns}	8.77^{ns}	9.26^{bc}
$EV_{0.7}$	10.52^{ns}	10.04^{ns}	13.85^{b}	36.68^{bc}	25.00^{ns}	9.25^{ns}	8.25^{bc}
$EV_{0.3}$	11.58^{ns}	11.00^{ns}	14.50^{ab}	42.06^{ab}	26.56^{ns}	9.43^{ns}	9.79^{b}
$EV_{0.0}$	9.25^{ns}	8.87^{ns}	12.28^{b}	34.84^{c}	25.67^{ns}	9.45^{ns}	7.93^{c}
$LV_{0.7}$	10.86^{ns}	10.73^{ns}	14.30^{ab}	39.01^{bc}	27.53^{ns}	8.93^{ns}	11.41^{a}
$LV_{0.3}$	10.78^{ns}	10.50^{ns}	14.54^{ab}	38.89^{bc}	27.13^{ns}	10.05^{ns}	9.86^{b}
$LV_{0.0}$	10.40^{ns}	10.18^{ns}	13.79^{b}	37.19^{bc}	26.20^{ns}	9.05^{ns}	8.90^{bc}

*: Significant differences between treatments are indicated by different letters ($P<0.05$).

Table 4- Quality parameters of red cabbage in 2022

Treatments	*Diameter-x (cm)*	*Diameter-y (cm)*	*Height (cm)*	*Circumference (cm)*	*Leaf Number*	*Total Soluble solid (%)*	*Dry weight percent (%)*
$I_{1.0}$	14.02^{a}	13.96^{a}	15.33^{ns}	41.96^{a}	31.44^{ns}	8.77^{ns}	9.26^{bc}
$EV_{0.7}$	11.92^{ab}	11.31^{ab}	14.30^{ns}	38.67^{ab}	32.67^{ns}	8.17^{ns}	8.37^{c}
$EV_{0.3}$	10.67^{bc}	10.48^{bc}	13.33^{ns}	34.46^{bc}	33.56^{ns}	9.67^{ns}	9.40^{bc}
$EV_{0.0}$	8.10^{c}	8.05^{c}	12.19^{ns}	33.35^{c}	33.33^{ns}	9.63^{ns}	9.43^{bc}
$LV_{0.7}$	12.08^{ab}	11.85^{ab}	13.33^{ns}	37.91^{bc}	32.56^{ns}	9.90^{ns}	9.76^{bc}
$LV_{0.3}$	11.37^{ab}	11.39^{ab}	13.00^{ns}	35.95^{bc}	32.89^{ns}	9.85^{ns}	11.07^{b}
$LV_{0.0}$	11.61^{ab}	10.97^{bc}	12.91^{ns}	35.77^{bc}	31.22^{ns}	10.17^{ns}	13.82^{a}

*: Significant differences between treatments are indicated by different letters (P<0.05)

In 2022, however, lower amount of rainfall during the plant growth period indicated the effects of water stress on diameter development highlighting a significant difference in plant diameter growth. Water restriction applied during the EV period had a more pronounced negative impact on plant growth compared to the LV period. This indicates that implementing water restriction during the late vegetative period is more suitable for plant development, because in the EV period, compared to the LV period, the diameters were the smallest with 8.10 cm in x, and 8.05 cm in y directions. Compensating full water demand of the crops in the control treatment ensured that the highest values in diameter were obtained as 14.02 cm in x and 13.96 cm in y directions.

Significant differences were observed in height values in 2021, while no significant difference was observed in 2022. In both years, plant height was the highest in control treatment and the lowest values were obtained in the treatment that water restrictions applied in EV period. Limited irrigation also significantly affected the circumference of red cabbage in all phases in 2021. In 2022, plant circumference was higher in the control treatment than the $EV_{0.0}$ and all LV treatments. Differences in irrigation had no significant effect on the leaf number and total soluble solids in red cabbage in both years.

With respect to dry weight values, in 2021, the highest value was achieved with 11.42% in the LV0.7 stage and in 2022, the highest value was obtained with 13.82% in the $LV_{0.0}$ period. The situation can be explained as follows; in red cabbage, when the plant's water demand is fully met until the end of the EV period, there is a significant increase in fruit weight. In the LV period, however, there is no statistically significant difference in fruit weight increase in both years, as observed in Table 2.

3.2. Effects of water stress on major antioxidant compounds

The amount of flavonoids, phenolic compounds, antioxidant capacity using the CUPRAC and DPPH methods, and anthocyanin levels in red cabbages exposed to seasonal water stress during the two-year experiment are given in Table 5.

Table 5- Biochemical changes under different irrigation treatments

Treatments	*Flavonoid (mg g^{-1})*		*Phenolic (mg g^{-1})*		*CUPRAC (mg g^{-1})*		*DPPH (mg g^{-1})*		*Anthocyanin (mg g^{-1})*	
	2021	*2022*	*2021*	*2022*	*2021*	*2022*	*2021*	*2022*	*2021*	*2022*
$I_{1.0}$(Control)	0.25^{d}	0.47^{bc}	0.73^{c}	0.84^{c}	1.08^{c}	1.38^{b}	2.05^{a}	3.01^{a}	0.6^{d}	0.5^{d}
$EV_{0.7}$	0.34^{ab}	0.53^{bc}	1.14^{b}	1.03^{bc}	1.47^{b}	1.38^{b}	1.94^{a}	2.86^{ab}	0.96^{ab}	1.05^{c}
$EV_{0.3}$	0.35^{ab}	0.59^{b}	1.14^{b}	1.09^{bc}	1.49^{ab}	1.54^{b}	1.97^{a}	2.64^{b}	0.79^{c}	1.21^{bc}
$EV_{0.0}$	*0.36^{a}*	*0.83^{a}*	1.03^{b}	1.11^{b}	1.47^{b}	1.54^{b}	1.63^{b}	2.73^{ab}	*1.04^{a}*	*1.51^{a}*
$LV_{0.7}$	0.29^{c}	0.38^{c}	1.2^{b}	1.19^{bc}	1.4^{b}	1.35^{b}	1.69^{b}	2.62^{b}	0.88^{bc}	1.15^{bc}
$LV_{0.3}$	0.32^{bc}	0.49^{bc}	1.42^{a}	1.25^{b}	1.61^{ab}	1.55^{b}	1.52^{b}	2.56^{b}	0.88^{bc}	1.33^{ab}
$LV_{0.0}$	0.28^{cd}	0.52^{bc}	*1.54^{a}*	*1.62^{a}*	*1.7^{a}*	*1.93^{a}*	*1.45^{b}*	*2.57^{b}*	0.92^{bc}	1.45^{a}

*: Significant differences between treatments in both years are indicated by different letters (P<0.05); **: Numbers in bold represent the highest value within each column

When considering the flavonoid levels in red cabbage plants, those exposed to EV water stress showed significant differences compared to both the control group and the experimental groups subjected to LV water stress in both years. Overall, it was found that water stress experienced by red cabbage during early development leads to increased flavonoid levels.

In the 2022 early vegetative period, the highest flavonoid content of 0.83 mg g^{-1} Fresh Weight (FW) was recorded in the treatment ($EV_{0.0}$) where no irrigation was applied. It is worth noting, however, that this treatment also resulted in a considerable decrease in fruit weight of 650 grams, indicating an approximate 50% yield loss. Thus, while water stress enhanced flavonoid accumulation, it also caused economically significant yield reductions. Addressing this issue, implementing a 30% water saving irrigation practice during the EV period ($EV_{0.7}$) resulted in an average fruit weight of 972 grams corresponding to a 22% decrease in yield, enabling producers to achieve an economically viable, also producing a high flavonoid content of 0.53 mg/g, beneficial

for human health. Moreover, applying a 70% water saving treatment ($EV_{0.3}$) during the EV period in red cabbage led to higher flavonoid content of 0.59 mg g^{-1} compared to other treatments. This treatment also showed an average yield reduction of 43%, resulting in a yield of 708 grams, which remains statistically and economically viable.

Looking at the phenolic compound levels, a variation between 0.73 and 1.54 mg g^{-1} in 2021 and 0.84 and 1.62 mg g^{-1} in 2022 was observed. The lowest value in terms of phenolic compounds was obtained in the control treatment where the plant water needs were fully met. It was determined that the highest accumulation of phenolic compounds occurred in red cabbage with the water restrictions applied during LV period. The highest values were obtained during the LV period, with 1.42 mg g^{-1} for the treatment with 30% fulfillment of the plant's water needs ($LV_{0.3}$) and 1.54 mg g^{-1} for the treatment where no water was provided ($LV_{0.0}$). In the subsequent year, these values were obtained as 1.25 for the $LV_{0.3}$ treatment and 1.62 mg g^{-1} for the $LV_{0.0}$ treatment.

The antioxidant levels in the experimented red cabbage plants (Table 5) were also determined using two methods, namely the CUPRAC and DPPH methods. The highest antioxidant level, as determined by the CUPRAC method, was found in the $LV_{0.0}$ experimental group in both years. Similarly, the DPPH method showed lower radical scavenging activity in the $LV_{0.0}$ treatment, indicating higher antioxidant content in the water stressed plants in both years.

The study results also showed variations in anthocyanin levels among experimental groups subjected to different levels of water stress during different growth stages. The highest anthocyanin synthesis of 1.04 mg g^{-1} and 1.51 mg g^{-1} were obtained from the $EV_{0.0}$ experimental group in the first year (2021) and second year (2022), respectively. The lowest anthocyanin amount was obtained from the control plants in both years.

4. Discussion

Results showed that there is slight difference between the yield obtained from plants applied with water stress in the EV and LV periods. Head weights were slightly higher in the plants applied with restricted water in the LV period compared to the EV period. Similarly, a review carried out by Bute et al. (2021) have stated that it is critical to keep red cabbage irrigated when the seedlings have 6-7 leaves and when head formation starts. This indicates that if water restriction is to be applied, it would be more suitable to impose it during the late vegetative period in terms of yield.

Furthermore, the parameters such as head diameter in the x and y directions (cm), circumference (cm), and dry weight (%) indicated that plant development was also negatively affected as the amount of water decreased at different growth phases. Both the yield and quality parameters in the water stressed red cabbage indicate that full irrigation should be applied until the end of the EV period if the objective is to get marketable yield. Alternatively, if marketable yield is not the concern and if water conservation is required, 30% of the plant's water demand could be restricted in the LV period.

Significant increments in flavonoid levels were observed in treatments with water stress during the early vegetative period during both years of the experiment. In comparison to previous studies, Lin et al. (2008) identified the amount of flavonoids in red cabbage between 0.6-2.1 g $100g^{-1}$. These results showed similar outcomes with the control plants of the present study. In another study carried out on tomatoes (Kumar et al. 2015), drought stress increased flavonoid levels statistically significantly although a reduction in yield occurred. This reduction in yield, however, was lowest when water stress was implemented in the vegetative period, rather than the flowering or fruiting periods.

In a study carried out to understand flavonoid response of Mediterranean species, Laoué et al. (2022) have stated that the accumulation of these defensive chemicals is highest in constrained environments due to drought, high temperatures and UV radiations. This mechanism would explain the higher flavonoid levels accumulated in the water stressed EV period of this study, which had the highest temperatures throughout the experiment.

The phenolic compounds also increased statistically significantly, particularly when water stress occurred during the LV period in this research. Comparing to previous studies, while Erken (2022) found an increase in the phenolic content of red cabbage subjected to long-term water-stress, both Shawon et al. (2020) and Šola et al. (2021) reported that the polyphenol content did not differ significantly between the control and drought stressed Chinese cabbage plants. While Shawon et al. (2020) applied short-term drought stress, Šola et al. (2021) implemented longer-term drought stress. The results altogether indicate that polyphenol induction in drought stressed cabbage may occur when stress is applied at the late vegetative period. This result, however, may also be inconclusive due to differences in the species in the abovementioned studies.

Further results of the study showed that the lowest antioxidant content obtained using both the CUPRAC and DPPH methods was in the control plants where the red cabbage plants were fully irrigated. According to Valifard et al. (2017), plants synthesize and accumulate natural antioxidants in response to abiotic stress. The greater antioxidant capacity in the water stressed red cabbage is an indication of its higher ability to stabilize free radicals (Reyes et al. 2017).

Similar to our results, in a two-year study carried out by Hegazi & El-Shraiy (2017), the antioxidant enzyme activity in red cabbage increased statistically significantly under salt stress. Previous studies (Jafari et al. 2019; Erken 2022) have also found that long term water stress influences the biochemical properties of red cabbage significantly.

The results of this study further suggest that, in order to get the highest antioxidant benefits from red cabbage, water restriction can be applied at the LV phase as the highest antioxidant amount was obtained from the $LV_{0.0}$ treatment. These results are similar to the accumulation of phenolic compounds, which increased in red cabbage as water restriction approached harvest time.

The anthocyanin values obtained from the cabbage in this experiment ranged from 0.5 to 1.51 mg g^{-1}. Mazza & Minati (1993) reported that the anthocyanin content for red cabbage lie within the wide range from 25 to 495 mg 100 g^{-1} FW (0.25 to 4.95 mg g^{-1} FW). Ahmadiani et al. (2014) found that the total anthocyanin content of seven red cabbage cultivars at different harvest times ranged from 109 to 170 mg Cy3G 100 g^{-1} FW when harvested in the 13th week, and from 104 to 188 mg Cy3G 100 g^{-1} 1FW when harvested at the 21st week. The amount of anthocyanin decreased in some red cabbage cultivars, while it decreased in others as the harvest time increased.

Similar to the flavonoids, the results of this study showed that the highest accumulation of anthocyanin was measured in red cabbages exposed to water stress in the EV stage. In a research carried out by Erken (2022) the anthocyanin levels increased statistically significantly (from 30.72 to 51.27 mg Cy3G 100g^{-1} FW) with water stress in red cabbage. Hegazi & El-Shraiy (2017) also found increased anthocyanin levels in salt stressed red cabbage.

While there are no known studies that specifically investigate the impacts of water stress at different vegetative phases, this study has shown that the increase in the anthocyanin content of the plants is higher when stress is applied during the late vegetative phase.

5. Conclusions

The findings of this study revealed that yield loss was least with the 30% water saving practice applied in the late vegetative phase (LV0.7). Results also showed that implementing a 30% water saving practice (EV0.7) during the early vegetative period lead to an elevation in flavonoid and phenolic compound levels. The antioxidant and anthocyanin amount in the water stressed red cabbage, on the other hand, were highest in the LV0.0 treatments. These findings could assist irrigation management strategies regarding red cabbage cultivation for different purposes.

Conflicts of interest

The authors declare that there is no conflict of interest to disclose.

Financial Support

This work was supported by Çanakkale Onsekiz Mart University, the Scientific Research Coordination Unit, Project Number: FBA-2021-3687.

References

Ahmadiani N, Robbins R J, Collins T M & Giusti M M (2014). Anthocyanins contents, profiles, and color characteristics of red cabbage extracts from different cultivars and maturity stages. *Journal of Agricultural and Food Chemistry* 62: 7524–7531. https://doi.org/10.1021/jf501991q.

Ak I & Türker G (2018). Antioxidant activity of five seaweed extracts. *New knowledge Journal of science* 7: 149–155

Apak R, Güçlü K, Ozyürek M & Karademir S E (2004). Novel total antioxidant capacity index for dietary polyphenols and vitamins C and E, using their cupric ion reducing capability in the presence of neocuproine: CUPRAC method. Journal of Agricultural and Food Chemistry, 52: 7970–7981. https://doi.org/10.1021/jf048741x.

Beacham A M, Hand P, Pink D A & Monaghan J M (2017). Analysis of *Brassica oleracea* early stage abiotic stress responses reveals tolerance in multiple crop types and for multiple sources of stress. Journal of the Science of Food and Agriculture, 97: 5271–5277. https://doi.org/10.1002/jsfa.8411.

Benvenuti S, Pellati F, Melegari M A & Bertelli D (2004). Polyphenols, anthocyanins, ascorbic acid, and radical scavenging activity of Rubus, Ribes, and Aronia. Journal of Food Science, 69: 164–169. https://doi.org/10.1111/j.1365-2621.2004.tb13352.x.

Brand-Williams W, Cuvelier M E & Berset C (1995). Use of a free radical method to evaluate antioxidant activity. LWT - Food Science and Technology, 28: 25–30. https://doi.org/10.1016/S0023-6438(95)80008-5

Bute A, Iosob G A, Antal-Tremurici A, Brezeanu C, Brezeanu P M, Cristea T O & Ambăruş S (2021). The Most Suitable Irrigation Methods in Cabbage Crops (*Brassica Oleracea* L. var. Capitata): A Review. Sci Papers Series B. Horticulture, 65(1)

Chu Y F, Sun J W & Liu R H (2002). Antioxidant and antiproliferative activities of common vegetables. Journal of Agricultural and Food Chemistry, 50: 6910-6916. https://doi.org/10.1021/jf020665f

Cohen J, Kristal R & Stanford J (2000). Fruit and vegetable intakes and prostate cancer. Journal of the National Cancer Institute 9: 61-68. https://doi.org/10.1093/jnci/92.1.61

Dhungel R, Anderson R, French G A, Skaggs N T, Saber H, Sanchez M C A & Scudiero E (2023). Early season irrigation detection and evapotranspiration modeling of winter vegetables based on Planet satellite using water and energy balance algorithm in lower Colorado basin. Irrigation Science, 1-13. https://doi.org/10.1007/s00271-023-00874-7

Erken O (2022). Some bioactive metabolites' response to long-term water stress in red cabbage. Scientia Horticulturae 293: 110731. https://doi.org/10.1016/j.scienta.2021.110731

Giusti M M, Rodriguez-Saona L E, Baggett J R, Reed L, Durst R W & Wrolstad R E (1998). Anthocyanin pigment composition of red radish cultivars as potential food colorants. Journal of Food Science, 63: 219–224. https://doi.org/10.1111/j.1365-2621.1998.tb15713.x

Haghighi M, Saadat S & Abbey L (2020). Effect of exogenous amino acids application on growth and nutritional value of cabbage under drought stress. Scientia Horticulturae, 272: 109561 p.1-7. https://doi.org/10.1016/j.scienta.2020.109561

Hajiboland R. & Amirazad H (2010). Drought tolerance in Zn-deficient red cabbage (*Brassica oleracea* L. var. capitata f. rubra) plants. Horticultural Science, 37(3): 88-98. https://doi.org/10.17221/64/2009-HORTSCI

Hegazi A M & El-Shraiy A M (2017). Stimulation of photosynthetic pigments, anthocyanin, antioxidant enzymes in salt stressed red cabbage plants by ascorbic acid and potassium silicate. Middle East Journal of Agriculture Research, 6(2): 553-568

Igarashi K, Kimura Y & Takenaka A (2000). Preventive effect of dietary cabbage acylated anthocyanins on paraquat induced oxidative stress in rats. Bioscience, Biotechnology, and Biochemistry, 64: 1600-7. https://doi.org/10.1271/bbb.64.1600

Jafari S, Garmdareh S H E & Azadegan B (2019). Effects of drought stress on morphological, physiological, and biochemical characteristics of stock plant *(Matthiola incana* L.). Scientia Horticulturae, 253: 128–133. https://doi.org/10.1016/j.scienta.2019.04.033

Janabi A H W, Kamboh A A, Saeed M, Xiaoyu L, Bibi J, Majeed F, Naveed M, Mughal M J, Korejo R A, Kamboh R, Alagawany M & Lv H (2020). Flavonoid-rich foods (FRF): A promising nutraceutical approach against lifespan-shortening diseases. Iranian Journal of Basic Medical Sciences, 23(2): 140. doi: 10.22038/IJBMS.2019.35125.8353

Kishor N, Khanna M, Rajanna G A, Singh M, Singh A, Banerjee T, Patanjali N, Singh S, Parihar C M, Prasad S, Manu S M, Kiruthiga B & Arockia A (2023). Red cabbage (*Brassica oleracea*) response to hydrogels under drip irrigation and fertigation regimes. The Indian Journal of Agricultural Sciences, 93 (5): 529–533. https://doi.org/10.56093/ijas.v93i5.132723

Kumar P S, Singh Y, Nangare D D, Bhagat K, Kumar M, Taware P B & Minhas P S (2015). Influence of growth stage specific water stress on the yield, physio-chemical quality and functional characteristics of tomato grown in shallow basaltic soils. Scientia Horticulturae, 197: 261-271. http://dx.doi.org/10.1016/j.scienta.2015.09.054

Kusvuran S & Abak K (2012). Kavun Genotiplerinin Kuraklık Stresine Tepkileri. ÇU Uni Fen ve Müh Bil Der 28(5): 78-87 (In Turkish).

Laoué J, Fernandez C & Ormeño E (2022). Plant flavonoids in Mediterranean species: A focus on flavonols as protective metabolites under climate stress. Plants, 11 (2): 172. https://doi.org/10.3390/plants11020172

Lin J Y, Li C Y & Hwang I F (2008). Characterization of the pigment components in red cabbage (Brassica oleracea L. var.) juice and their anti-inflammatory effects on LPS-stimulated murine splenocytes. Food Chemistry, 109 (4): 771-781. https://doi.org/10.1016/j.foodchem.2008.01.039

Lobos T E, Retamales J B, Ortega-Farías S, Hanson E J, López-Olivari R & Mora M L (2017). Regulated deficit irrigation effects on physiological parameters, yield, fruit quality and antioxidants of *Vaccinium corymbosum* plants cv. Brigitta. Irrigation Science, 36: 49-60. https://doi.org/10.1007/s00271-017-0564-6

Magalhaes L M, Santos F, Segundo M A, Reis S & Lima J L (2010). Rapid microplate high-throughput methodology for assessment of Folin-Ciocalteu reducing capacity. Talanta 83(2): 441-447

Majkowska-Gadomska J & Wierzbicka B (2008). Content of Basic Nutrients and Minerals in Heads of Selected Varieties of Red Cabbage *(Brassica oleracea* var. capitata f. rubra). *Polish Journal of Environmental Studies* 17(2): 295-298. https://doi.org/10.1016/j.scienta.2020.109561

Mazza G & Miniati E (1993). Anthocyanins in Fruits, Vegetables and Grains. CRC Press, Boca Raton, FL.

MGM.0 (2023). Temperature, Precipitation and Evaporation Data. The Turkish State Meteorological Service (in Turkish: Meteoroloji Genel Müdürlüğü). Received February 23, 2023.

Podsedek A (2007). Natural antioxidants and antioxidant capacity of Brassica vegetables: A review. LWT- Food Science and Technology, 40: 1-11. https://doi.org/10.1016/j.lwt.2005.07.023

Reyes L F, Villarreal J E & Cisneros-Zevallos L (2017). The increase in antioxidant capacity after wounding depends on the type of fruit or vegetable tissue. Food Chemistry 101(3): 1254-1262. https://doi.org/10.1016/j.foodchem.2006.03.032

Saha C, Bhattacharya P, Sengupta S, Dasgupta S, Patra S K, Bhattacharyya K & Dey P (2021). Response of cabbage to soil test-based fertilization coupled with different levels of drip irrigation in an inceptisol. Irrigation Science 1-15. https://doi.org/10.1007/s00271-021-00761-z

Semiz G D, Şentürk C, Yildirim A C & Torun E (2023). Modelling Yield Response and Water Use to Salinity and Water Relations of Six Pepper Varieties. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi)* 29(1):188-199. DOI: 10.15832/ankutbd.1017255

Shawon R A, Kang B S, Lee S G, Kim S K, Lee H J, Katrich E, Gorinstein S & Ku Y G (2020). Influence of drought stress on bioactive compounds, antioxidant enzymes and glucosinolate contents of Chinese cabbage (*Brassica rapa*). Food Chemistry 308: 1–9. https://doi.org/10.1016/j.foodchem.2019.125657.

Shinde M G, Pawar D D, Kale K D & Dingre S K (2020). Performance of cabbage at different irrigation levels under drip and micro sprinkler irrigation systems. Irrigation Drainage Published online in Wiley Online Library Cited 02 June 2021. https://onlinelibrary.wiley.com/doi/pdf/10.1002/ird.2557.

Siro I, Kápolna E, Kápolna B & Lugasi A (2008). Functional food. Product development, marketing and consumer acceptance—A review. Appetite 51(3): 456-467

Šola I, Stić P & Rusak G (2021). Effect of flooding and drought on the content of phenolics, sugars, photosynthetic pigments and vitamin C and antioxidant potential of young Chinese cabbage. European Food Research and Technology 247: 1913-1920. https://doi.org/10.1007/s00217-021-03759-1

Şahin U, Ekinci M, Ors S, Turan M, Yıldız S & Yıldırım E (2018). Effects of individual and combined effects of salinity and drought on physiological, nutritional and biochemical properties of cabbage (*Brassica oleracea* var. capitata). Scientia Horticulturae 240: 196–204. https://doi.org/10.1016/j.scienta.2018.06.016

Valifard M, Mohsenzadeh S & Kholdebarin B (2017). Salinity effects on phenolic content and antioxidant activity of Salvia macrosiphon. Iranian Journal of Science and Technology, Transaction A 41: 295–300. https://doi.org/10.1007/s40995-016-0022-y.

Volden J, Borge G I A, Bengtsson G B, Hansen M, Thygesen I E & Wicklund T (2008). Effect of thermal treatment on glucosinolates and antioxidant-related parameters in red cabbage (*Brassica oleracea* L. ssp. capitata f. rubra). Food Chemistry 109: 595-605. https://doi.org/10.1016/j.foodchem.2008.01.010

Wallace J S (2000). Increasing agricultural water use efficiency to meet future food production. Agriculture, Ecosystems & Environment, 82(1-3): 105-119. https://doi.org/10.1016/S0167-8809(00)00220-6

Wiczkowski W, Szawara-Nowak D & Topolska J (2013). Red cabbage anthocyanins: profile, isolation, identification, and antioxidant activity. Food Research International 51: 303-309. https://doi.org/10.1016/j.foodres.2012.12.015

Zhang X, Lu G, Long W, Zou X, Li F & Nishio T (2014). Recent progress in drought and salt tolerance studies in Brassica crops. Breeding Science 64: 60-73. https://doi.org/10.1270/jsbbs.64.60