El Cezeri
Journal of Science and Engineering
ISSN: 2148-3736

Research Article

Tea Stem Waste Powders for Epoxy-Based Composites: Processing, Mechanical Properties, and Fracture Behavior

Hüseyin Kaya[1a], Mürsel Ekrem[2b], Yasin Uslugil[3c]
[1] Department of Machine and Metal Technologies, Technical Sciences Vocational School, Karamanoğlu Mehmetbey University, Karaman, 70100, Türkiye
[2] Department of Mechanical Engineering, Necmettin Erbakan University, Konya, 42005, Türkiye
[3] Department of Mechatronics, Trade and Industry Vocational School, KTO Karatay University, Konya, 42010, Türkiye

yasin.uslugil@karatay.edu.tr

DOI : 10.31202/ecjse.1766935
Received: 16.08.2025 Accepted: 30.09.2025
How to cite this article:
Hüseyin Kaya, Mürsel Ekrem, and Yasin Uslugil, " Tea Stem Waste Powders for Epoxy-Based Composites: Processing, Mechanical Properties, and Fracture Behavior", El-Cezeri Journal of Science and Engineering, Vol: 12, Iss: 3, (2025), pp.(235-248).
ORCID: [a]0000-0002-2714-8481, [b] 0000-0001-5324-7929,[c]0000-0002-7077-2300.

Abstract This study investigates the potential of mechanically processed tea stem waste powders as an environmentally friendly reinforcement for epoxy-based composites. Tea stem waste, a lignocellulosic byproduct of tea processing, was reduced in size by high-speed mixing and planetary ball milling without any chemical pretreatment, and incorporated into an epoxy system at varying weight fractions (1–10 wt.%). The composites were fabricated under controlled conditions and subjected to tensile, fracture toughness, hardness, thermal, and morphological evaluations. Mechanical testing revealed that tensile strength increased with filler loading up to an optimum of 7.5 wt.% (72.16 MPa), corresponding to a 14.8% improvement over neat epoxy, while higher content(10 wt.%) showed minor reduction due to potential filler agglomeration. The elastic modulus displayed a non-linear trend, decreasing slightly at low filler levels but reaching 1.30 GPa at 10 wt.%. Fracture toughness (Kıc) exhibited a continuous increase, with a maximum of 2,41 MPa√m at 10 wt.%—a 26.4% improvement compared to neat epoxy. Shore D hardness also improved progressively, with up to a 23% increase at the highest filler content. Thermal analysis confirmed that the addition of tea stem powder did not significantly affect the glass transition or initial degradation temperatures, while an increased char yield indicated enhanced thermal stability. FE-SEM fractography showed that fiber orientation strongly influenced crack propagation: longitudinal fibers remained intact, while transverse fibers fragmented and dissipated energy, thereby enhancing fracture resistance. These results confirm that tea stem waste powders, processed solely by mechanical means, can effectively reinforce epoxy matrices. Although samples made with 7.5% filler yield the best results in tensile testing, it has been observed that 10% content yields higher results in fracture toughness and hardness tests. The results of the study showed that the mechanical strength of the composite material made from this tea stem waste was higher than that of pure epoxy in all content variations tested here.

Keywords: Epoxy composites, Fracture toughness, Hardness, Natural fibers, Tea stem waste

1.INTRODUCTION

The rapid depletion of fossil-based resources and the environmental burden caused by petroleum-derived polymers have become critical global concerns. These synthetic polymers, despite offering excellent mechanical performance, thermal stability, and ease of processing, are inherently non-biodegradable and persist for decades in landfills and natural ecosystems, thereby contributing significantly to long-term environmental pollution and microplastic accumulation [1], [2]. In addition, the extraction and processing of petroleum feedstocks for polymer synthesis result in substantial greenhouse gas emissions and high energy consumption, further exacerbating global climate change. Such challenges have shifted the focus of materials science toward the development of sustainable and renewable alternatives that can provide comparable mechanical integrity, thermal performance, and durability while simultaneously minimizing ecological impact and supporting circular economy principles[3].

Synthetic fibers are increasingly being replaced with the natural fiber-reinforced composites, which are an attractive combination of mechanical and environmental properties [4]. Their inherently low density contributes to high specific strength, enabling lightweight yet mechanically robust materials suitable for demanding structural applications [5], [6]. Additionally, natural fiber composites exhibit excellent vibration damping characteristics and acoustic properties, making them ideal for noise reduction applications[7], [8]. Natural fibers composites in real-world applications have seen to be utilized the most in sectors like automotive, construction and consumer goods. The integration of fibers derived, for example, from flax, hemp, and jute were part of car interior and structural components enabling less weight in vehicles and better fuel economy [9], [10], [11].

Epoxy resins are well-known for their great strength and excellent resistance to chemicals, being used in high wear and tear industries as well as top performance demands [12]. With high thermal stability and good adhesion and moldability, these thermosetting polymers are employed in industrial applications for various types of composites (industrial and structural [13], [14]). However, conventional epoxy resins are almost entirely petroleum-derived—typically synthesized from precursors like epichlorohydrin and bisphenol A—raising critical sustainability concerns due to reliance on non-renewable feedstocks and associated environmental impacts [15].

Natural fiber-reinforced epoxy composites have been extensively studied for their mechanical and thermal properties, with significant findings across various fibers and fillers. The incorporation of bamboo fiber with nano SiO_2 in a hybrid composite showed a reduction in tensile, flexural, impact, and hardness strengths at 4 wt. % silica, highlighting an optimal limit of 3 wt.% [16]. Banana fiber composites with 0.75 wt% functionalized g-C_3N_4 exhibited enhanced tensile strength (126.8 MPa), flexural strength wt. %.9 MPa), and impact strength (26.9 J/m), suggesting potential for PCB applications [17]. Optimized banana/coir hybrid epoxy composites with a 65:35 fiber/matrix volume proportion, 10 mm fiber length, and 75:25 banana/coir ratio demonstrated superior tensile, compressive, impact, and flexural properties [18]. Cotton fiber composites achieved the highest tensile strength at 52.81 MPa, while coir composites showed the lowest at 15.34 MPa, with coir also exhibiting the highest specific heat (26.313 MJ/m^{3}K) for energy-saving insulation [19]. In another study, basalt-bamboo hybrid composites toughened with 4 vol.% echinoidea spike β-chitin absorbed the highest impact energy (17.4 J at 3 m/s), indicating improved impact resistance and thermal stability [20]. In a related work, it was found that adding areca fiber and pectin biofiller to epoxy led to a clear increase in tensile strength, rising from 65 MPa in neat epoxy to 107 MPa with areca fiber, and further up to 147 MPa when 3 vol.% pectin was included, before decreasing again at higher filler content due to agglomeration [21]. In the study by Singh and Rout (2023), the tensile strength of Borassus leaf fiber + rice husk nanoparticle epoxy composites increased from 9.24 MPa in raw fiber composites to 14.0 MPa at 0.45 wt. % RHN. Similarly, the tensile modulus improved from 471 MPa to 489 Mpa [22].

Tea processing generates substantial amounts of fibrous waste, including leaves, stems, buds, and other residues originating from harvesting, pruning, drying, and brewing operations [23]. In Türkiye, tea cultivation is carried out exclusively in the Eastern Black Sea Region, primarily in Rize, Trabzon, Artvin, Giresun, and Ordu with an annual fresh tea production of 1.357 million tonnes in 2023. During processing, approximately 203,483 tonnes of ITW(Industrial Tea Waste) are generated each year, corresponding to about 17.6% of the total harvested fresh tea mass [24]. Among the various components of ITW, tea stem waste stands out due to its high fiber content and complex lignocellulosic structure, rich in cellulose, hemicellulose, and lignin, which are conducive to reinforcement applications in composite materials[25]. The fibrous nature and rigidity of these stems offer inherent potential for mechanical reinforcement, such as enhancing stiffness and impact resistance when incorporated into polymer matrices [26]. However, despite these favorable properties, research on the utilization of tea stem waste, especially in the form of fiber-containing powders for thermosetting matrix composites remains scarce, highlighting a significant gap in valorization strategies for this underutilized agricultural residue[27].

In one study, industrial waste tea leaf fiber (WTLF) treated with 5% NaOH showed a tensile strength of 132–165 MPa. When 5 wt% WTLF was incorporated into hybrid epoxy composites, tensile, flexural, and impact strengths reached 70.81 MPa, 181.08 MPa, and 96 kJ/m², respectively. Higher WTLF contents reduced performance due to fiber agglomeration and poor resin wetting [28]. In another study, industrial tea waste (ITW) powder (1–250 µm) added to jute–unsaturated polyester composites showed the best performance at 9 wt%, with tensile and flexural strengths of 26.58 MPa and 40.58 MPa, respectively. Higher ITW contents caused agglomeration, reducing mechanical properties [29].

Waste tea fiber (WTF) reinforced bio-based epoxy composites, modified with maleic anhydride (MAH), achieved tensile strengths up to 16.1 MPa with elongation at break of 23.8%, and maintained good thermal stability up to 244 °C. These composites also showed remarkable durability, with tensile strength increases of up to 65.8% after heat aging and 49.5% after UV exposure[26]. A recent study utilized tea stem fiber (TSF) to reinforce polypropylene (PP) through conventional and microcellular injection molding. The best balance of strength and weight was obtained with 30 wt% TSF in foamed composites, delivering tensile, flexural, and impact strengths of 32.84 MPa, 51.31 MPa, and 6.94 kJ/m², alongside a ~6% density reduction. The enhanced performance was attributed to improved interfacial bonding from alkali treatment and maleic anhydride grafting, combined with a uniform microcellular structure [27].

The conversion of wastes into micro- and nanoparticles using a ball mill is relatively rare in the literature. For instance, Sharma et al. reported that rice straw remaining after harvest was ground with a ball mill to produce cellulose nanofibers (CNF), and although no composite was fabricated, the fiber microstructure was examined in detail [30]. Similarly, in a study with jute composites, ball milling not only reduced particle size but also activated the fiber surface, enhancing matrix adhesion and providing a low-cost alternative to chemical pretreatments [31]. While CNF production from agricultural residues often involves chemical pretreatments to enhance surface charge and dispersion, the present study demonstrates that mechanically processed tea stem waste reduced in size via high-speed mixing and subsequent ball milling without any chemical modification can still serve as an effective reinforcement in epoxy composites, combining simplicity, low cost, and environmental friendliness.

In this study, tea stem waste was selected as a lignocellulosic agricultural residue with potential for reinforcement in polymer composites. The waste was mechanically processed through high-speed mixing and planetary ball milling to obtain fiber-containing powders without the use of chemical pretreatments. These powders were then incorporated into epoxy matrices at different weight ratios to fabricate composite specimens. A systematic characterization program was conducted, including tensile testing, fracture toughness measurements, Shore D hardness evaluation, thermal analyses (DSC and TGA), and morphological examinations using FE-SEM. Through this comprehensive approach, the study aimed to establish a framework for the utilization of tea stem waste as a sustainable filler in epoxy-based composites, contributing to the broader valorization of agricultural residues in advanced material applications.

2. MATERIALS

The tea stem wastes used were obtained from a tea factory in Rize (Turkey) and no chemical treatment was applied. The tea stems photographed in Figure 1 are approximately 3 cm long and 2.5 mm in diameter.

Figure 1.Tea stem waste in its original form as supplied from the factory.

The epoxy system employed in this study consisted of two components supplied by Momentive. The A-component, designated as MGS-L285, is a lamination-grade resin with a viscosity of 600–900 mPa·s, formulated as a biphasic mixture containing 80–90 wt% diglycidyl ether of bisphenol A and 10–20 wt% aliphatic diglycidyl ether. The B-component, MGS-H285 curing agent, comprises 70–90 wt% cycloaliphatic amine and 10–30 wt% polyoxyalkyl amine, with a viscosity of 50–100 mPa·s. When mixed at the manufacturer-recommended ratio of 100:40 (by weight), the system exhibits a gelation time of approximately 2–3 h at room temperature.

3. METHODS

This section describes the preparation of tea stem waste powder, the fabrication of epoxy-based composites with varying filler contents, and the subsequent mechanical and microstructural characterizations performed to evaluate their performance.

3.1. PREPARATION OF TEA STEM WASTE POWDER

The raw stems, approximately 3 cm in length and 2.5 mm in diameter, were first dried at 70 °C for 6 h to remove residual moisture. The dried material was coarsely size reduced using a high-speed mixer (Retsch/GM200) operating at 10000 rpm for 2 min. The resulting coarse particles were then milled in a planetary ball mill (Retsch PM100) at 300 rpm for a total of 2 h, using 5mm diameter stainless steel balls at a ball-to-powder weight ratio of 10:1.

Figure 2.Tea stem powder after ball-milling operation.

To prevent excessive heat generation from the grinding media, the milling process was performed in cycles of 15 min operation followed by 5 min cooling intervals. In one similar study, it was observed that bamboo-based cellulose nanofibers were produced at 500 rpm for 4 hours [32]. In yet another study, bamboo fibers were milled at 300 rpm with a 10:1 ball ratio for 3.6 and 12 hours. In this context, the values selected in this study are consistent with and appropriate for the literature[33]. Laser scatter analysis of the tea stem powder revealed an average particle size of 73.20 µm, with a median (D50) of approximately 71.9 µm. The particle size distribution was relatively broad, as 10% of the particles were smaller than 22.4 µm while 90% were below 142.6 µm. The resulting fiber-containing powder was stored in sealed containers at ambient conditions until composite fabrication. Figure 2 shows the resulting powder in the container.

3.2. COMPOSITE FABRICATION

In the composite fabrication stage, the micro-sized waste tea stem powder and epoxy resin (A-component), obtained from the preceding preparation processes, were utilized as the primary constituents. The waste tea powder, dried and sieved to the targeted particle size range, was incorporated into the epoxy resin without the addition of the curing agent, in accordance with the predetermined weight ratios presented in Table 1. This dry–wet blending process was carried out using a laboratory-scale mechanical stirrer with a rated power of 1000 W. The mixing speed was maintained at a constant 500 rpm for a duration of 20 minutes to promote effective wetting of the tea powder particles by the resin and to ensure their uniform dispersion throughout the viscous epoxy phase. The controlled shear provided by the stirrer facilitated the breakdown of any powder agglomerates and minimized localized resin-rich or fiber-rich regions, which are known to cause mechanical property inconsistencies in cured composites.

Table 1. Formulations of Epoxy–Waste Tea Stem Powder Composites

Sample Code	Epoxy Resin A [wt.%]	Hardener B [wt.%]	Waste Tea Stem Powder [wt.%]
E-0	60	40	0,00
E-1.0	59,4	39,6	1,00
E-2.5	58,5	39	2,50
E-5	57	38	5,00
E-7.5	55,5	37	7,50
E-10	54	36	10,00

For each formulation, an independent batch was prepared under identical processing conditions to maintain consistency across experimental variables. The homogenized suspensions, exhibiting a uniform particle distribution and stable viscosity, were then ready for the subsequent steps involving curing agent addition, mold casting, and thermal curing protocols, which ensured complete crosslinking and optimal network formation within the composite matrix.

Prior to casting, the epoxy–waste tea stem powder(Figure 3) mixtures were subjected to ultrasonic dispersion(Figure 3) using a 1000 W ultrasonic processor operated at 20% amplitude for 5 minutes. This additional step was employed to break down agglomerates and promote a more uniform distribution of the tea powder particles within the resin matrix. Ultrasonic energy generates localized high shear and cavitation, which enhances particle wetting and reduces the likelihood of particle clustering, thereby contributing to improved interfacial bonding and more consistent mechanical performance in the final composites.

Figure 3.Waste tea stem powder (left) and ultrasonic dispersion of epoxy–tea powder mixture using a 1000 W ultrasonic processor at 20% amplitude for 5 minutes (right).

3.3. PREPARATION OF TENSILE AND FRACTURE TEST SPECIMENS

The homogenized epoxy–waste tea stem powder mixtures were subsequently poured into steel molds designed for tensile and fracture specimens in accordance with the relevant ASTM standards, as illustrated in Figure 4. These molds were manufactured to meet the specified dimensions and tolerances required for the respective mechanical tests, ensuring consistency and comparability of the results. Care was taken during the pouring process to minimize air entrapment, as voids could serve as stress concentrators and adversely affect the mechanical performance of the cured composites.

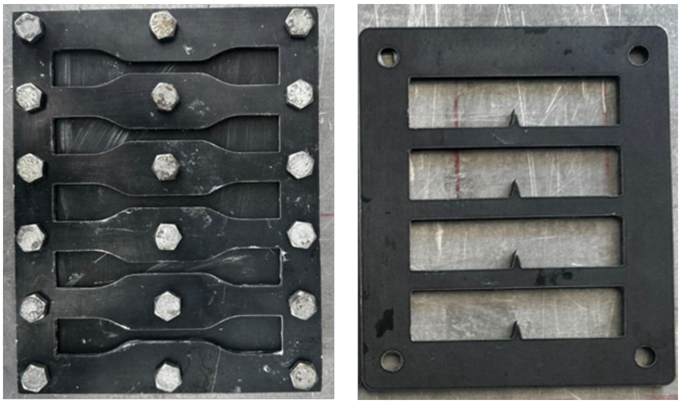


Figure 4.Steel molds for specimen fabrication: tensile test mold (left) and fracture test mold (right), manufactured in accordance with ASTM D638[34] & ASTM D5045[35] standards.

Following the initial casting, the composites were allowed to cure at room temperature for 8 hours to facilitate gelation and partial crosslinking of the epoxy matrix. Subsequently, a post-curing process was carried out in a laboratory oven, where the specimens were maintained at 80 °C for 2 hours to achieve complete crosslinking. This thermal treatment increased the crosslink density of the polymer network, resulting in improved mechanical strength, thermal stability, and overall composite performance. The curing parameters were selected in accordance with the epoxy resin manufacturer's recommended curing profile to ensure optimal material properties.

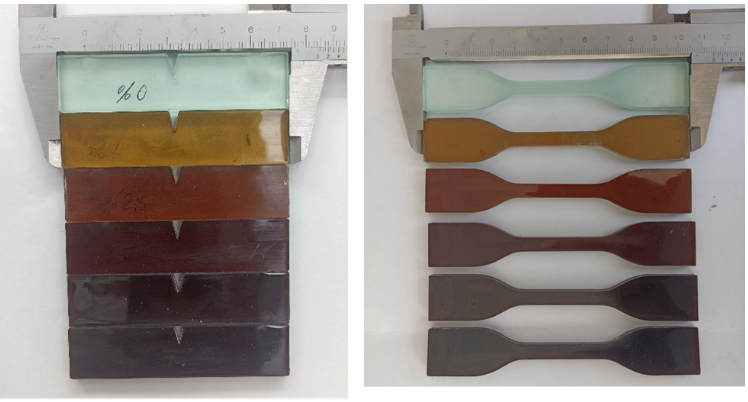


Figure 5.Fracture and tensile specimens with varying tea stem waste powder contents.

Figure 5 presents the sets of fracture (left) and tensile (right) test specimens produced with varying weight percentages of waste tea stem powder, as specified in Table 1. The specimens were fabricated in accordance with the relevant ASTM standards using steel molds to ensure precise geometry and dimensional accuracy. The gradual change in coloration across the series reflects the increasing filler content, which not only influences the visual appearance but is also expected to impact the microstructural characteristics, interfacial bonding, and consequently the mechanical performance of the composites. These specimens represent the final cured state after the complete curing cycle, including room-temperature pre-curing and oven post-curing, and were subsequently used for mechanical testing.

3.4. TENSILE TESTING

Tensile testing of the epoxy–waste tea stem powder composites was conducted using a Shimadzu universal testing machine with a load cell capacity of 10,000 N, as illustrated in Figure 6a. The specimens were mounted in standard wedge grips (Figure 6b, 6c), ensuring proper alignment to minimize bending stresses during loading. Tests were performed at a crosshead speed of 2 mm/min in accordance with ASTM D638, under controlled laboratory conditions at an ambient temperature of 23 ± 2 °C and relative humidity of 50 ± 5%. Load–displacement data were continuously recorded for the determination of tensile strength, elastic modulus, and elongation at break. The machine's high sensitivity and precise displacement control ensured reliable measurements across all composite formulations. The tensile tests were repeated for 3 samples of each tea stem powder loading composition.

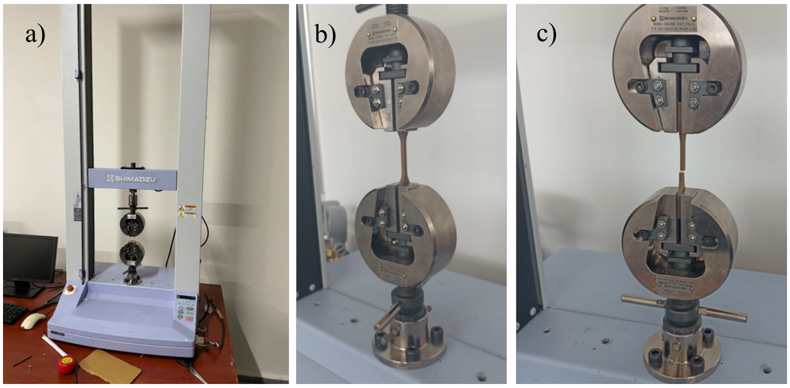


Figure 6.Shimadzu universal testing machine with a 10000N load cell (a), tensile specimen attached to the machine before testing (b), broken tensile specimen after testing (c)

3.5. FRACTURE TEST

Fracture toughness tests were performed in accordance with ASTM D5045-99 for the determination of plane-strain fracture toughness (K_{IC}) of plastic materials. All tests were carried out on a Shimadzu universal testing machine with a maximum load capacity of 10,000 N. Single-edge notched bend (SENB) specimens with a total length of 85 mm, width (W) of 20 mm, thickness (B) of 4 mm, and an initial crack length (a) of 9 mm were used. The notches were pre-formed during molding and subsequently sharpened using a fresh razor blade in a single controlled sliding motion to obtain the sharp crack tip geometry required by the standard.

The support span-to-width ratio (S/W) was maintained at 4, corresponding to an 80 mm span length. Tests were conducted at 23 ± 1 °C and 50 ± 5% relative humidity under displacement control, with a crosshead speed of 1 mm/min. Load–displacement curves were recorded, and the conditional stress intensity factor (K_Q) was calculated according to the equations provided in ASTM D5045-99. Three repeats were performed for each sample set for fracture toughness tests. The validity of the K_Q values was checked against the size requirements of the standard to determine whether they qualified as K_{IC}.

Figure 7.Shore D hardness tester (UVE Measurement Equipment) with Novotest durometer used for composite specimen hardness measurements.

3.6. SHORE-D HARDNESS TEST

Shore D hardness measurements of the composite specimens were conducted using a UVE Measurement Equipment setup equipped with a Novotest Shore D durometer. The test procedure followed the guidelines specified in ASTM D2240[36], which is the standard method for determining the indentation hardness of rigid plastics. Prior to testing, all specimens were conditioned at room temperature (approximately 23 ± 2 °C) and relative humidity of 50 ± 5% for at least 24 hours to ensure thermal and moisture equilibrium. Each specimen was placed on a flat, rigid support surface, and the indenter was applied perpendicularly to the sample surface without impact, maintaining firm contact for 15 seconds before recording the hardness value. For each formulation, a minimum of five measurements were taken at different locations on the specimen surface to account for local variations, and the average Shore D hardness was calculated and reported.

3.7. POST-FRACTURE MORPHOLOGICAL ANALYSIS WITH FE-SEM

The fracture surfaces of the tested specimens were examined using a Zeiss Gemini SEM 500 Field Emission Scanning Electron Microscope (FE-SEM) to investigate the microstructural features and failure mechanisms induced by tensile and fracture toughness testing. The imaging was performed at an accelerating voltage of 5 kV, with a working distance (WD) ranging from 5 to 6 mm depending on the topography of the observed region. High-resolution micrographs enabled the identification of matrix–filler interfacial characteristics, fiber pull-out phenomena, and the extent of fiber–matrix adhesion. Distinct fracture morphologies were observed depending on the waste tea stem powder content, including regions of brittle cleavage, ductile tearing ridges, and resin-rich areas. The dispersion of filler particles and the presence of voids or agglomerates were also analyzed, as these microstructural features can act as stress concentrators and significantly influence the mechanical performance.

3.8. THERMAL ANALYSIS BY DSC AND TGA

The thermal properties of the epoxy-based composites were characterized using DSC and TGA. Two representative specimens were analyzed: neat epoxy (sample mass: 12.91 mg) and epoxy reinforced with 10 wt.% waste tea stem powder (sample mass: 11.07 mg).

Both DSC and TGA measurements were performed under a nitrogen atmosphere with a flow rate of 50 mL/min to prevent oxidative degradation. For DSC, each sample was heated from 25 °C to 250 °C at a constant rate of 10 °C/min, cooled back to room temperature, and then reheated under the same conditions. The glass transition temperature (T_g) was determined from the midpoint of the heat capacity change during the second heating cycle. For TGA, the heating range was from 25 °C to 800 °C at a rate of 10 °C/min. From the TGA curves, the onset decomposition temperature (T_{onset}), the maximum decomposition temperature (T_{max}), and the residual char content at 800 °C were recorded. The total run time for each TGA analysis was approximately 70 minutes.

4. RESULTS AND DISCUSSION

This section presents and interprets the experimental findings obtained from the mechanical, thermal, and morphological characterization of epoxy-based composites reinforced with varying weight percentages of waste tea stem powder. The results are organized to first address the influence of filler content on the tensile and fracture properties, followed by an evaluation of the composites' hardness performance. Subsequently, the thermal stability and glass transition behavior are analyzed through DSC and TGA measurements, providing insights into the effects of the lignocellulosic filler on the thermal performance of the epoxy matrix. Finally, FE-SEM fractographic examinations are discussed to elucidate the underlying failure mechanisms, focusing on fiber–matrix interfacial bonding, fracture surface morphology, and filler dispersion characteristics.

4.1. HARDNESS TEST RESULTS

As can be seen in the box plot presented in Figure 7, the distribution and variation of Shore D hardness values clearly demonstrate the reinforcing effect of tea stem powder on the epoxy matrix. The neat epoxy (E-0) exhibited a median hardness of approximately 79, with relatively narrow variability across replicates. At 1 wt.% filler (E-1.0), the hardness remained nearly unchanged (~80), indicating that such a low content does not provide a significant stiffening effect. A remarkable improvement was observed at 2.5 wt.% (median ~93), representing a ~17% increase compared to the neat epoxy, suggesting effective load transfer and surface densification. At 5 wt.% (median ~91), the hardness slightly decreased but remained significantly higher than the neat epoxy, likely due to localized agglomerates influencing indentation resistance. The 7.5 wt.% and 10 wt.% composites exhibited the highest hardness values, ~96 and ~98 respectively, corresponding to ~20–23% improvements relative to the control. The narrow interquartile ranges for these groups further confirm the uniform reinforcement effect at higher filler contents.

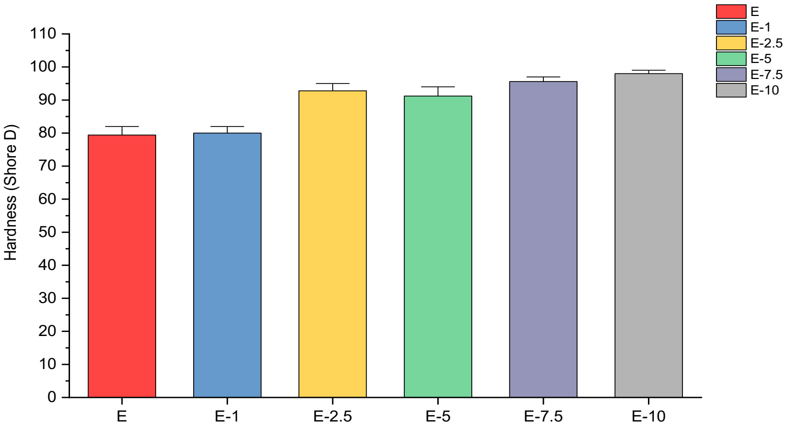


Figure 8.Shore D hardness of epoxy–tea stem composites with varying filler contents.

4.2. TENSILE TEST RESULTS

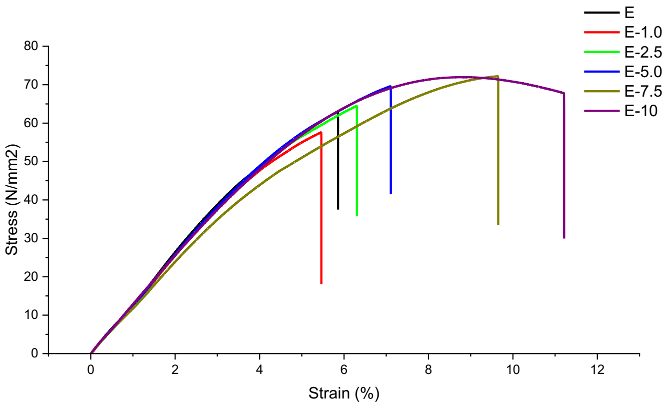


Figure 9.Tensile test results of epoxy-tea stem composite tensile specimens in accordance with ASTM standards.

In tensile test, each composite variation was tested using three replicate specimens, and the representative results presented here correspond to the median values among the triplicates. The average standard deviation across all groups was approximately ±1.2 MPa, indicating good reproducibility of the measurements. As shown on Figure 9, based on the results, the neat epoxy (E-0) exhibited a maximum strength of 62.84 MPa. The incorporation of 1.0 wt.% waste tea stem powder (E-1.0) led to a reduction of approximately 8.5 % in tensile strength, decreasing to 57.51 MPa, likely due to insufficient filler content to effectively reinforce the matrix and potential particle–matrix interfacial imperfections. Conversely, a slight filler loading of 2.5 wt.% (E-2.5) improved the tensile strength to 64.37 MPa, corresponding to a 2.4 % increase compared to the neat epoxy. A more pronounced enhancement was observed for E-5.0, which reached 69.58 MPa (a 10.7 % increase), indicating more efficient stress transfer between the matrix and the dispersed filler. The highest tensile strength was recorded for E-7.5 at 72.16 MPa, representing a 14.8 % improvement over the neat epoxy, suggesting that this composition provides the optimum filler-to-matrix balance for mechanical reinforcement. A slight decrease was observed for E-10 (71.91 MPa, 14.5 % above neat epoxy), which, although still high, may be attributed to the onset of particle agglomeration and micro-defect formation at higher filler contents, partially offsetting the reinforcing effect. These results confirm that 7.5 wt.% is the optimal filler content for maximizing tensile strength in the studied epoxy–waste tea stem composites.

4.3. EVALUATION OF TENSILE ELASTIC MODULUS

The elastic modulus values were determined by considering the 0.5–1.0% strain interval within the elastic deformation region. The calculated results for each formulation are presented in Figure 10.

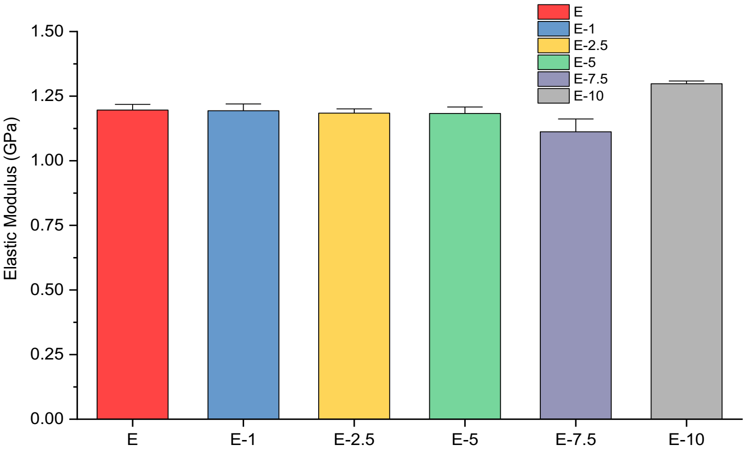


Figure 10.Variation of elastic modulus of epoxy–tea stem powder composites with different filler contents (GPa).

The elastic modulus values of the epoxy–tea stem powder composites exhibit a non-linear trend with increasing filler content. While the neat epoxy (E) shows a modulus of approximately 1.21 GPa, the addition of 1–7.5 wt.% filler results in a slight decrease, indicating a modest reduction in stiffness. At 10 wt.% filler loading (E-10), however, a significant increase in modulus is observed, reaching nearly 1.30 GPa.

Consistent with reports in the literature, the absence of a significant improvement in elastic modulus with fiber addition can be explained by several microstructural and processing-related factors. In short and randomly oriented fiber composites, residual stresses often accumulate at the fiber ends, and local shear stresses develop around the fibers. More importantly, the orientation of fibers tends to change during melt processing, where the material is subjected to both extensional and shear flow. Under extensional flow, fibers rotate toward the direction of extension, whereas in shear flow, some fibers align with the shear direction while others rotate oppositely. Such variations in orientation reduce the effective alignment of fibers along the loading direction, thereby limiting stress transfer efficiency. Combined with the critical fiber length requirement and the dominant role of the matrix in load bearing at low fiber content, these factors collectively explain why the elastic modulus does not increase appreciably despite fiber incorporation [37], [38].

4.4. FRACTURE TEST RESULTS

The fracture toughness (K_{IC}) results (Figure 11) demonstrated that the addition of tea stem powder improved the crack resistance of the epoxy matrix. Compared with neat epoxy, the composite with 2.5 wt.% filler showed a slight increase of about 2%. More pronounced enhancements were observed at higher loadings: 5 wt.% and 7.5 wt.% resulted in gains of roughly 12% and 14%, respectively. The highest increase was achieved at 10 wt.% filler, with an improvement of nearly 26% over the neat epoxy.

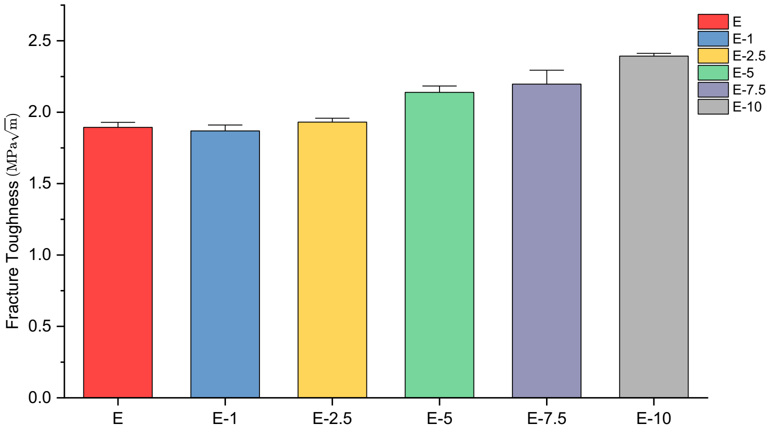


Figure 11. Fracture toughness values of epoxy–tea stem powder composites.

4.5. FRACTURED PARTS MORPHOLOGY ANALYSIS

When the surface morphology of the neat epoxy was examined at 1000× magnification (Figure 12a), a homogeneous curing structure with minimal inclusions or foreign particulates was observed, indicating uniform crosslinking of the polymer matrix. Furthermore, the fracture surface of the neat epoxy at 500× magnification (Figure 12b) revealed predominantly brittle fracture features, characterized by smooth cleavage regions [39].

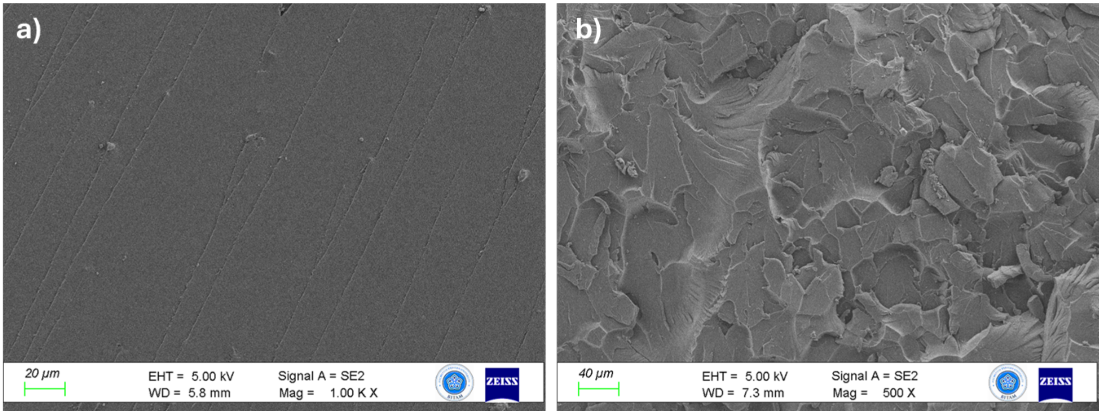


Figure 12. FE-SEM micrographs of neat epoxy: (a) surface morphology of the untested neat epoxy at 1000× magnification, and (b) fractured surface morphology after testing at 500× magnification.

In certain localized zones, mixed-mode features combining tearing ridges[40] and quasi-brittle[41] fracture patterns were also evident, reflecting limited plastic deformation prior to final failure.

Figure 13 shows the fracture surface of the epoxy composite containing 10 wt.% tea stem waste. The fracture morphology reveals distinct lamellar (step-like) structures and river patterns, which are characteristic of crack propagation in brittle polymeric systems. It is observed that fibers aligned parallel to the crack propagation direction largely remain as long and intact fragments, whereas fibers oriented perpendicular to the fracture direction tend to break into multiple shorter segments. This fragmentation increases interfacial friction and energy dissipation within the matrix, thereby hindering crack propagation and contributing to higher fracture resistance.

These microstructural features are consistent with the mechanical results, where the 10 wt.% composite exhibited the highest fracture toughness. The coexistence of intact and fragmented fibers demonstrates that increasing fiber content not only enhanced crack resistance but also supported the strengthening trend observed in tensile performance, confirming the beneficial role of fiber addition in improving both strength and toughness.

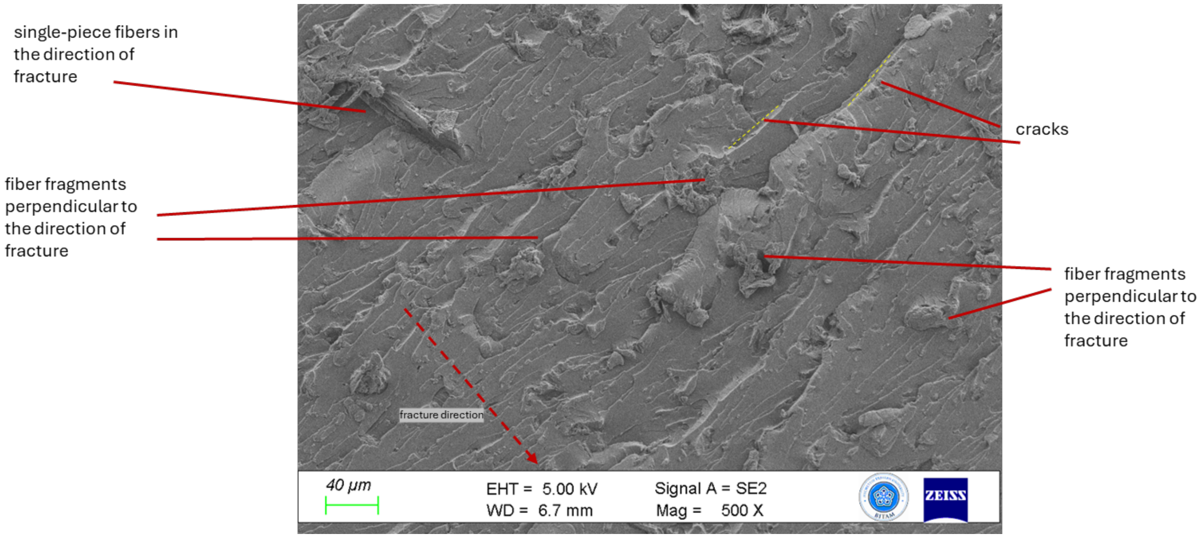


Figure 13.FE-SEM micrograph of the fracture surface of the epoxy composite containing 10 wt.% tea stem waste at 500× magnification.

4.6. THERMAL ANALYSIS RESULTS

The DSC and TGA analyses of the neat epoxy and the 10 wt.% tea stem powder-reinforced composite are presented in Figure 14. The DSC thermograms (upper curves) reveal that both materials exhibit a similar primary endothermic transition, corresponding to the glass transition temperature (Tg), followed by an exothermic region attributable to residual crosslinking reactions of the epoxy matrix. The incorporation of tea stem powder did not significantly shift the Tg, indicating that the filler did not substantially alter the segmental mobility of the epoxy network. However, a slightly higher heat flow was observed in the composite beyond Tg, which can be associated with the presence of lignocellulosic constituents contributing additional reaction enthalpy or altering the effective heat capacity of the system [16], [42], [43].

The TGA curves (lower plots) show comparable onset degradation temperatures for both materials, approximately around 300 °C, confirming that the addition of tea stem powder does not markedly compromise the initial thermal stability of the matrix. Nevertheless, the composite exhibited a marginally faster mass loss in the early stages of degradation, reflecting the lower thermal stability of the lignocellulosic filler compared to the epoxy resin. Furthermore, the composite retained a higher char residue at the end of the test, which is attributed to the inorganic ash content and the enhanced carbonization tendency of the tea stem powder.

The increased char yield observed in the thermogravimetric analysis is not only indicative of improved thermal stability but also has direct implications for application performance. A higher char content can contribute to the formation of a dense, carbonaceous protective layer during combustion, which serves as an effective barrier to heat and oxygen transfer [44]. This phenomenon enhances the fire resistance and flame-retardant behavior of the composites. Such characteristics are particularly desirable in applications where fire safety is critical, including automotive interiors, building materials, and electrical insulation systems [44], [45].

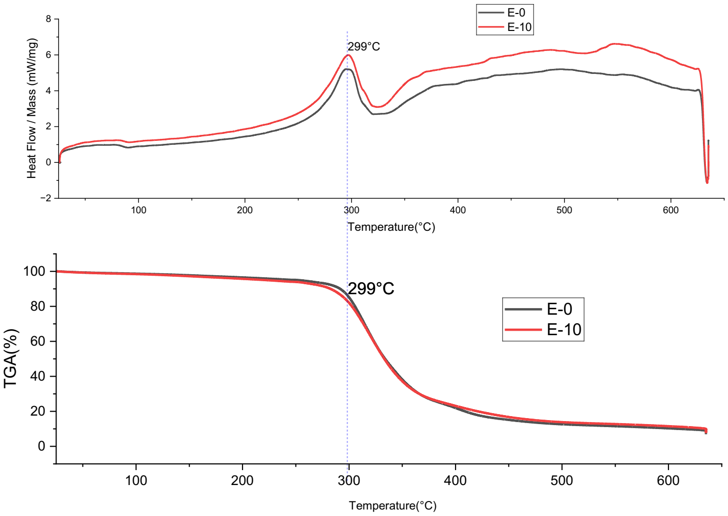


Figure 14.DSC and TGA curves of neat epoxy and 10 wt.% tea stem composites.

5. CONCLUSION

This study demonstrated the feasibility of utilizing mechanically processed tea stem waste powders as a sustainable reinforcement for epoxy-based composites without requiring any chemical pretreatment. Comprehensive mechanical, thermal, and morphological evaluations consistently revealed the reinforcing capability of the filler.

Tensile strength results indicated that the optimum reinforcement level was achieved at 7.5 wt.% filler content, yielding a 14.8% increase compared to neat epoxy, while 10 wt.% still maintained high values but showed slight reductions due to agglomeration effects. Elastic modulus analysis revealed a non-linear trend, with modest decreases at low filler contents but a significant increase at 10 wt.% (≈1.30 GPa), demonstrating the stiffening potential of high filler loading.

Fracture toughness (K_{IC}) values exhibited a clear and continuous improvement with increasing filler content, reaching a maximum enhancement of 26.4% at 10 wt.%, which highlights the role of tea stem fibers in impeding crack propagation. Shore D hardness measurements further confirmed these findings, showing progressive increases with filler content and achieving up to a 23% improvement in indentation resistance at 10 wt.% loading.

Thermal analysis by DSC and TGA demonstrated that filler incorporation did not significantly alter the glass transition temperature or onset decomposition temperature of the epoxy, while the higher char residue confirmed improved thermal stability associated with the lignocellulosic filler. Morphological examination via FE-SEM revealed distinct fracture features: fibers aligned parallel to the crack propagation direction remained intact, whereas transversely oriented fibers fragmented, thereby enhancing crack resistance by increasing interfacial friction and energy dissipation.

The results confirm that waste tea stem powders, processed solely by mechanical means, provide a low-cost, environmentally friendly, and effective reinforcement for epoxy composites. While the optimum composition for tensile strength was identified at 7.5 wt.% filler content, the 10 wt.% loading exhibited superior performance in fracture toughness, hardness, and thermal stability, indicating that the best overall reinforcement effect is achieved at higher filler contents. These findings not only establish a valorization pathway for agricultural residues but also contribute to the development of sustainable polymer composites with reduced environmental footprints.

Beyond these fundamental outcomes, the developed composites show strong potential for use in automotive interior components, furniture, building materials, and other cost-sensitive engineering applications where mechanical robustness and sustainability are both required. For future work, it would be valuable to explore surface modification strategies to further improve interfacial adhesion, hybridization with other natural or synthetic fillers, and scalability studies for industrial processing. Additionally, evaluating long-term durability, water uptake, and fire-retardant behavior would strengthen the applicability of these composites in safety-critical and outdoor environments.

Authors' Contributions

Hüseyin Kaya was responsible for specimen preparation, experimental testing, and data processing. Mürsel Ekrem contributed to the manuscript through critical review, editing, and supervision. Yasin Uslugil carried out the manuscript writing and data analysis.

Competing Interests

The authors declare that they have no conflict of interest.

References

[1] N. Singh, O. A. Ogunseitan, M. H. Wong, and Y. Tang, "Sustainable materials alternative to petrochemical plastics pollution: A review analysis," *Sustainable Horizons*, vol. 2, p. 100016, Mar. 2022, doi: 10.1016/j.horiz.2022.100016.

[2] T. Khan, M. T. B. Hameed Sultan, and A. H. Ariffin, "The challenges of natural fiber in manufacturing, material selection, and technology application: A review," *Journal of Reinforced Plastics and Composites*, vol. 37, no. 11, pp. 770–779, June 2018, doi: 10.1177/0731684418756762.

[3] S. H. Mousavi-Avval, K. Sahoo, P. Nepal, T. Runge, and R. Bergman, "Environmental impacts and techno-economic assessments of biobased products: A review," *Renewable and Sustainable Energy Reviews*, vol. 180, p. 113302, July 2023, doi: 10.1016/j.rser.2023.113302.

[4] K. Karthik *et al.*, "State of the Art: Natural fibre-reinforced composites in advanced development and their physical/chemical/mechanical properties," *Chinese Journal of Analytical Chemistry*, vol. 52, no. 7, p. 100415, July 2024, doi: 10.1016/j.cjac.2024.100415.

[5] R. A. Khan, "A Review on the Properties of Natural Fibers and Manufacturing Techniques of Fiber Reinforced Biocomposites," *MCMS*, vol. 4, no. 4, Nov. 2021, doi: 10.33552/MCMS.2021.04.000592.

[6] F. M. L. Rekbi *et al.*, "Ultrasonic Characterization of the Mechanical Behavior of Epoxy/Date Kernel Powder Biocomposites: A Feasibility Study of the Powder Size Effect," *ECJSE*, vol. 12, no. 2, pp. 176–190, May 2025, doi: 10.31202/ecjse.1572805.

[7] V. Sekar, M. H. Fouladi, S. N. Namasivayam, and S. Sivanesan, "Additive Manufacturing: A Novel Method for Developing an Acoustic Panel Made of Natural Fiber-Reinforced Composites with Enhanced Mechanical and Acoustical Properties," *Journal of Engineering*, vol. 2019, pp. 1–19, Sept. 2019, doi: 10.1155/2019/4546863.

[8] S. Gokulkumar, P. Thyla, L. Prabhu, and S. Sathish, "Characterization and Comparative Analysis on Mechanical and Acoustical Properties of *Camellia Sinensis/Ananas Comosus/* Glass Fiber Hybrid Polymer Composites," *Journal of Natural Fibers*, vol. 18, no. 7, pp. 978–994, July 2021, doi: 10.1080/15440478.2019.1675215.

[9] R. H. Hu, M. H. Jang, Y. J. Kim, Y. J. Piao, and J. K. Lim, "Fully Degradable Jute Fiber Reinforced Polylactide Composites Applicable to Car Interior Panel," *AMR*, vol. 123–125, pp. 1151–1154, Aug. 2010, doi: 10.4028/www.scientific.net/AMR.123-125.1151.

[10] C. Alves *et al.*, "Ecodesign of automotive components making use of natural jute fiber composites," *Journal of Cleaner Production*, vol. 18, no. 4, pp. 313–327, Mar. 2010, doi: 10.1016/j.jclepro.2009.10.022.

[11] G. Thilagavathi, E. Pradeep, T. Kannaian, and L. Sasikala, "Development of Natural Fiber Nonwovens for Application as Car Interiors for Noise Control," *Journal of Industrial Textiles*, vol. 39, no. 3, pp. 267–278, Jan. 2010, doi: 10.1177/1528083709347124.

[12] T. Aziz *et al.*, "The epoxy resin system: function and role of curing agents," *Carbon Lett.*, vol. 34, no. 1, pp. 477–494, Jan. 2024, doi: 10.1007/s42823-023-00547-7.

[13] N. Saba, M. Jawaid, O. Y. Alothman, M. Paridah, and A. Hassan, "Recent advances in epoxy resin, natural fiber-reinforced epoxy composites and their applications," *Journal of Reinforced Plastics and Composites*, vol. 35, no. 6, pp. 447–470, Mar. 2016, doi: 10.1177/0731684415618459.

[14] C. Czaderski, E. Martinelli, J. Michels, and M. Motavalli, "Effect of curing conditions on strength development in an epoxy resin for structural strengthening," *Composites Part B: Engineering*, vol. 43, no. 2, pp. 398–410, Mar. 2012, doi: 10.1016/j.compositesb.2011.07.006.

[15] S. Kumar, S. Krishnan, S. Mohanty, and S. K. Nayak, "Synthesis and characterization of petroleum and biobased epoxy resins: a review," *Polymer International*, vol. 67, no. 7, pp. 815–839, July 2018, doi: 10.1002/pi.5575.

[16] R. Kumar, A. Ganguly, and R. Purohit, "Thermogravimetric analysis of natural fiber reinforced hybrid composites – A review," *Materials Today: Proceedings*, p. S221478532304275X, Aug. 2023, doi: 10.1016/j.matpr.2023.08.025.

[17] D. Sundeep and E. K. Varadharaj, "Mechanical and spectroscopic characterization of functionalized g-C3N4 fillers loaded epoxy reinforced banana natural Fiber composite for PCB applications," *J Polym Res*, vol. 31, no. 12, p. 368, Dec. 2024, doi: 10.1007/s10965-024-04218-7.

[18] M. S. Senthil Kumar, L. Rajeshkumar, S. M. Rangappa, and S. Siengchin, "Mechanical behaviour analysis for banana/coir natural fiber hybrid epoxy composites through experimental modelling," *J Polym Res*, vol. 31, no. 6, p. 163, June 2024, doi: 10.1007/s10965-024-04018-z.

[19] Y. Tasgin, G. Demircan, S. Kandemir, and A. Acikgoz, "Mechanical, wear and thermal properties of natural fiber-reinforced epoxy composite: cotton, sisal, coir and wool fibers," *J Mater Sci*, vol. 59, no. 24, pp. 10844–10857, June 2024, doi: 10.1007/s10853-024-09810-2.

[20] M. K. A. Khan, M. Faisal, and V. R. Arun Prakash, "Damage investigations on natural fiber-epoxy human prosthetic composites toughened using echinoidea spike β-chitin biopolymer," *Biomass Conv. Bioref.*, vol. 15, no. 4, pp. 6175–6185, Feb. 2025, doi: 10.1007/s13399-024-05421-8.

[21] R. A. Gandhi, V. Jayaseelan, S. Sambath, and V. Suyamburajan, "Production and characterization of epoxy based biocomposites using pectin biopolymer derived from Passiflora edulis husk and areca fibre," *Polym. Bull.*, Aug. 2025, doi: 10.1007/s00289-025-05968-0.

[22] J. K. Singh and A. K. Rout, “Study on the physical, mechanical, and thermal behaviour of RHN blend epoxy hybrid composites reinforced by Borassus flabellifer L. fibers,” *Cellulose*, vol. 30, no. 8, pp. 5033–5049, May 2023, doi: 10.1007/s10570-023-05191-y.

[23] Z. Wang, W. Ahmad, A. Zhu, S. Zhao, Q. Ouyang, and Q. Chen, “Recent advances review in tea waste: High-value applications, processing technology, and value-added products,” *Science of The Total Environment*, vol. 946, p. 174225, Oct. 2024, doi: 10.1016/j.scitotenv.2024.174225.

[24] Ş. Efe, “Industrial Tea Waste and Energy Potential,” Jan. 2025, doi: 10.5281/ZENODO.14670859.

[25] A. Donmez Cavdar, H. Kalaycioglu, and F. Mengeloğlu, “Technological properties of thermoplastic composites filled with fire retardant and tea mill waste fiber,” *Journal of Composite Materials*, vol. 50, no. 12, pp. 1627–1634, May 2016, doi: 10.1177/0021998315595113.

[26] N. Borah and N. Karak, “Green composites of bio-based epoxy and waste tea fiber as environmentally friendly structural materials,” *Journal of Macromolecular Science, Part A*, vol. 60, no. 3, pp. 217–229, Mar. 2023, doi: 10.1080/10601325.2023.2177171.

[27] F. Zhao, W. Guo, X. Liu, J. Zhao, and T. Feng, “Injection molded lightweight composites from tea-stem fiber and polypropylene: Effect of fiber loading on forming properties and cell structure,” *Industrial Crops and Products*, vol. 221, p. 119372, Dec. 2024, doi: 10.1016/j.indcrop.2024.119372.

[28] L. Prabhu, V. Krishnaraj, S. Gokulkumar, S. Sathish, and M. Ramesh, “Mechanical, Chemical and Acoustical Behavior of Sisal – Tea Waste – Glass Fiber Reinforced Epoxy Based Hybrid Polymer Composites,” *Materials Today: Proceedings*, vol. 16, pp. 653–660, 2019, doi: 10.1016/j.matpr.2019.05.142.

[29] H. Aftab, G. M. S. Rahman, Md. Kamruzzaman, M. A. Khan, Md. F. Ali, and M. A. A. Mamun, “Physico-Mechanical Properties of Industrial Tea Waste Reinforced Jute Unsaturated Polyester Composites,” *J. eng. adv.*, pp. 42–49, June 2022, doi: 10.38032/jea.2022.02.001.

[30] N. Sharma, B. J. Allardyce, R. Rajkhowa, and R. Agrawal, “Rice straw-derived cellulose: a comparative study of various pre-treatment technologies and its conversion to nanofibres,” *Sci Rep*, vol. 13, no. 1, p. 16327, Sept. 2023, doi: 10.1038/s41598-023-43535-7.

[31] R. Gallego *et al.*, “Green approach for the activation and functionalization of jute fibers through ball milling,” *Cellulose*, vol. 27, no. 2, pp. 643–656, Jan. 2020, doi: 10.1007/s10570-019-02831-0.

[32] Y. K. Verma, A. K. Singh, M. K. Paswan, and P. K. Gurmaita, “Preparation and characterization of bamboo based nanocellulose by ball milling and used as a filler for preparation of nanocomposite,” *Polymer*, vol. 308, p. 127396, Aug. 2024, doi: 10.1016/j.polymer.2024.127396.

[33] N. Jayarambabu, L. S. Rao, A. Akshaykranth, T. V. Rao, and R. R. Kumar, “Study of optical and structural properties of natural bamboo fiber powder prepared by ball milling method,” *Eur. Phys. J. Plus*, vol. 136, no. 10, p. 989, Oct. 2021, doi: 10.1140/epjp/s13360-021-01932-9.

[34] ASTM D20 Committee, *Test Method for Tensile Properties of Plastics*. doi: 10.1520/D0638-14.

[35] ASTM D20 Committee, *Test Methods for Plane-Strain Fracture Toughness and Strain Energy Release Rate of Plastic Materials*. doi: 10.1520/D5045-14.

[36] ASTM D11 Committee, *Test Method for Rubber Property--Durometer Hardness*. doi: 10.1520/D2240-15.

[37] H. Y. Sastra, J. P. Siregar, S. M. Sapuan, and M. M. Hamdan, “Tensile Properties of *Arenga pinnata* Fiber-Reinforced Epoxy Composites,” *Polymer-Plastics Technology and Engineering*, vol. 45, no. 1, pp. 149–155, Feb. 2006, doi: 10.1080/03602550500374038.

[38] R. M. N. Arib, S. M. Sapuan, M. M. H. M. Ahmad, M. T. Paridah, and H. M. D. K. Zaman, “Mechanical properties of pineapple leaf fibre reinforced polypropylene composites,” *Materials & Design*, vol. 27, no. 5, pp. 391–396, Jan. 2006, doi: 10.1016/j.matdes.2004.11.009.

[39] N. L. Ravikumar and D. Sunil, “THE EFFECT OF A STARCH ENVELOPE ON FLY ASH PARTICLES ON THE IMPACT PROPERTIES OF FILLED EPOXY COMP0SITES,” *Advanced Composites Letters*, 2003.

[40] J. Fan, J. Yang, H. Li, J. Tian, M. Wang, and Y. Zhao, “Cryogenic mechanical properties of graphene oxide/epoxy nanocomposites: Influence of graphene oxide with different oxidation degrees,” *Polymer Testing*, vol. 96, p. 107074, Apr. 2021, doi: 10.1016/j.polymertesting.2021.107074.

[41] X. Yao, I. A. Kinloch, and M. A. Bissett, “Fabrication and Mechanical Performance of Graphene Nanoplatelet/Glass Fiber Reinforced Polymer Hybrid Composites,” *Front. Mater.*, vol. 8, p. 773343, Nov. 2021, doi: 10.3389/fmats.2021.773343.

[42] S. Kumar, L. Prasad, P. P. Bijlwan, and A. Yadav, “Thermogravimetric analysis of lignocellulosic leaf-based fiber-reinforced thermosets polymer composites: an overview,” *Biomass Conv. Bioref.*, vol. 14, no. 12, pp. 12673–12698, June 2024, doi: 10.1007/s13399-022-03332-0.

[43] N. M. Nurazzi *et al.*, “Thermogravimetric Analysis Properties of Cellulosic Natural Fiber Polymer Composites: A Review on Influence of Chemical Treatments,” *Polymers*, vol. 13, no. 16, p. 2710, Aug. 2021, doi: 10.3390/polym13162710.

[44] Y. Kim, S. Lee, and H. Yoon, “Fire-Safe Polymer Composites: Flame-Retardant Effect of Nanofillers,” *Polymers*, vol. 13, no. 4, p. 540, Feb. 2021, doi: 10.3390/polym13040540.

[45] E. Madyaratri *et al.*, “Recent Advances in the Development of Fire-Resistant Biocomposites—A Review,” *Polymers*, vol. 14, no. 3, p. 362, Jan. 2022, doi: 10.3390/polym14030362.

El Cezeri
Journal of Science and Engineering
ISSN: 2148-3736

Research Article

A Non-AI Homework Support Tool to Enhance Achievement and Interest in Science Education: BilgeCan Bot

Taner Yılmaz[1a], İlbilge Dökme[2b]
[1]Medical Services and Techniques Department, Cekerek Fuat Oktay Vocational School of Health Services, Yozgat Bozok University, Yozgat, Turkiye
[2]Science Education Department, Education Faculty, Gazi University, Ankara, Turkey

taner.yilmaz@bozok.edu.tr

DOI : 10.31202/ecjse.1726783
Received: 25.06.2025 Accepted: 08.07.2025
How to cite this article:
Taner Yılmaz, İlbilge Dökme, " A Non-AI Homework Support Tool to Enhance Achievement and Interest in Science Education: BilgeCan Bot ", El-Cezeri Journal of Science and Engineering, Vol: 12, Iss: 3, (2025), pp.(249-273).
ORCID: [a] 0000-0003-1164-3549, [b]0000-0003-0227-6193.

Abstract Homework helps students learn and develop independent study skills, but they often need extra guidance and reliable information sources, especially in science where abstract concepts can be challenging. Traditional resources may not match students' cognitive levels and can lead to information overload. To address this, non-AI rule-based (NARB) educational chatbots can provide focused, essential information with minimal cognitive load. This study explores the effects of BilgeCan Bot, a NARB chatbot designed to help middle school students with astronomy homework, based on Cognitive Load Theory. Conducted during the 2022–2023 academic year with 52 fifth-grade students in Türkiye, the research used an explanatory sequential mixed-methods approach. Students were divided into an experimental group using BilgeCan Bot and a control group using textbooks. Data was collected through tests, interest scales, interviews, chatbot records, and teacher feedback. Results showed that students using BilgeCan Bot achieved higher scores and had greater interest in science. They described the chatbot as effective and engaging, helping them grasp difficult concepts. Overall, the findings suggest that NARB chatbots can offer targeted, reliable support for homework in science education without overwhelming students.

Keywords: Non-AI Rule-Based educational chatbot, Science education, Academic achievement, Cognitive load theory, Student interest.

I. INTRODUCTION

Chatbots are software programs that simulate human-like conversations through pattern recognition techniques and are increasingly recognized for their educational potential [1]. Educational chatbots (ECs) employ natural language processing algorithms to provide personalized guidance, facilitate both in-class and extracurricular learning, optimize students' learning pace, and enhance their understanding of abstract concepts [2–4]. Chatbots are typically categorized as rule-based or AI-based systems [5]. While AI-based chatbots adapt to student input, rule-based chatbots deliver information through predefined scenarios, offering structured guidance that mitigates cognitive load [5–8]. Particularly in science education, where abstract concepts prevail, structured information delivery enhances learning efficiency and academic achievement [9, 10].

Homework is a critical tool for assessing student progress, reinforcing classroom learning, and fostering independent study skills [11, 12]. When structured support is lacking at home, students seek reliable resources, particularly when facing conceptual difficulties [13–15]. This need is pronounced in science education, which demands abstract thinking and often leads to misconceptions, especially within topics like the "Sun, Earth, and Moon" unit [16, 17]. If unaddressed, these misconceptions may persist across educational stages [18]. Therefore, accurate instruction in abstract scientific concepts is essential for cognitive development [19]. Additionally, challenges such as time management and motivation during homework further highlight the necessity of structured support systems [20].

Students frequently use search engines such as Google and Yahoo to address homework-related challenges; however, these tools often provide unfiltered, context-independent information that limits their pedagogical utility [21, 22]. Consequently, the demand for dialogue systems offering meaningful, natural language feedback has grown [23, 3]. Educational chatbots meet this need by delivering real-time responses, fostering interactive learning, and enhancing students' motivation and depth of understanding [24]. Studies further indicate that educational chatbots strengthen teacher-student interaction and improve the effectiveness of homework support [25, 26]. The use of generative AI (GAI) systems such as ChatGPT for information retrieval among students has increased; however, their pedagogical reliability remains contested. Research indicates that students often struggle to assess online content credibility, with only 7% correctly identifying a fake website [27]. Given that GAI systems frequently draw from anonymous sources, they pose risks of misinformation and cognitive overload, potentially impairing critical thinking and offering contextually inappropriate content [28, 29].

In this context, rule-based educational chatbots facilitate learning by delivering only essential information, thereby reducing excessive cognitive load [30]. According to Cognitive Load Theory (CLT), effective learning requires optimizing intrinsic, extraneous, and germane cognitive loads [31, 32]. While AI-based chatbots may increase extraneous or germane loads through unstructured information, rule-based chatbots mitigate these risks by offering structured, focused content. Consequently, empirical research is needed to assess their effectiveness in teaching abstract subjects such as astronomy. BilgeCan Bot, developed in this study, operates on predefined rules without incorporating machine learning or deep learning techniques, and is thus classified as a non-AI rule-based (NARB) chatbot [33, 23, 34]. Although the effects of AI-supported chatbots on academic achievement and science interest have been widely studied [35, 36], limited research addresses the role of NARB chatbots in guiding homework tasks [37]. Particularly in middle school science education, the impact of NARB chatbots on student achievement and interest remains underexplored. This study aims to evaluate the effectiveness of BilgeCan Bot in supporting middle school students' astronomy homework by providing only necessary information and minimizing cognitive load.

II. BACKGROUND AND RELATED WORK

2.1. Chatbot Design Approaches and Rule-Based Chatbots

The term "chatbot" was first introduced by Mauldin in 1997 [38] and has since been known by terms like "Conversational Agent" and "Virtual Assistant" [39, 40]. Despite the variety, "chatbot" remains the most widely accepted term [41]. Early developments include the Turing Test and ELIZA, which used keyword-based transformation rules [33, 42].

Chatbots are classified based on purpose, domain, platform, and design [43–45]. Madhu et al. (2017) categorized them as AI-based (intelligent) and rule-based (non-intelligent). Rule-based systems, exemplified by ELIZA, are easier to implement and provide more reliable, structured responses [47], while AI-based systems may produce context-free or inaccurate outputs [16].

In education, chatbots like AutoTutor and Duolingo offer structured support [5]. Two main design approaches exist: pattern matching and machine learning [45]. Rule-based chatbots generate responses by matching inputs to predefined templates [48].

BilgeCan Bot, developed for this study, uses regex-based pattern matching without NLP or AIML [49], responding strictly based on predefined structures [50]. While "keyword matching" and "pattern matching" are often used interchangeably, deeper analysis is sometimes distinguished as "keyword recognition" [44, 43].

Most research focuses on AI-supported systems [51, 52], leaving a gap in understanding non-AI rule-based (NARB) chatbots for homework support. This study addresses that gap by exploring the educational potential of simplified NARB systems in middle school science education.

2.2. Cognitive Load Theory and Rule-Based Chatbots

Cognitive Load Theory (CLT) aims to optimize instructional design by considering the limited cognitive capacity of humans during the learning process [31]. According to the theory, three types of cognitive load occur during learning: intrinsic, extraneous, and germane load. CLT particularly emphasizes reducing extraneous load to allow more efficient use of students' cognitive resources [7]. In this context, it is important to design digital learning environments in a simple and guided manner [53]. Rule-based chatbots can help reduce extraneous cognitive load by providing only the necessary information and supporting guided learning [6]. Especially in science courses dealing with abstract concepts, they facilitate load management by delivering structured and directed content [54]. Compared to traditional sources, they offer more focused responses, minimizing unnecessary mental processing [55]. These systems can enhance learning efficiency by providing real-time feedback [56]. Since they operate based on predefined patterns without using natural language processing or artificial intelligence, they deliver information quickly, reliably, and with a low risk of error. When appropriately designed, they can optimize cognitive load by offering content tailored to individual needs [30] and promote active learning through interactive engagement [56].

In conclusion, rule-based chatbots developed within the framework of CLT have the potential to improve the learning experience by preventing information overload, providing guidance, and supporting personalized learning [57]. However, studies directly addressing the effects of rule-based chatbots on cognitive load management in science courses remain limited [6, 54]. This highlights the need for systematic investigation of systems like BilgeCan Bot in the context of science education.

2.3. The Use of Chatbots in Science Education

Science education often presents challenges for students due to the abstract nature of concepts and associated cognitive load [58–60]. Consequently, digital tools are increasingly used to support conceptual understanding and facilitate learning [61]. Among these tools, chatbots have emerged as effective by providing real-time guidance and step-by-step learning support.

Astronomy topics, being unobservable and highly abstract, are particularly difficult for students [62, 63]. Studies show that students often struggle with basic concepts such as the solar system, phases of the Moon, celestial movements, and seasonal changes [64, 65]. Interactive digital tools can help reduce misconceptions by making abstract ideas more concrete. Augmented reality, for example, visualizes celestial movements [5], while chatbots offer real-time responses to student inquiries, guiding conceptual understanding [66].

The interdisciplinary nature of astronomy demands flexible learning tools, as rigid curricula may restrict inquiry-based learning [63]. Chatbots can bridge this gap by explaining scientific concepts, providing visual support, and offering personalized learning opportunities [61]. Rule-based systems like BilgeCan Bot enhance learning efficiency by delivering structured guidance [67].

Research within the framework of Cognitive Load Theory (CLT) shows that material presentation directly affects conceptual understanding [57], and excessive cognitive load hampers learning [68]. Digital guides and chatbots can balance cognitive load by delivering targeted, essential information [32]. Instant feedback and structured guidance help internalize scientific concepts. Overall, chatbots hold strong potential for improving conceptual understanding, academic achievement, and scientific interest, especially in abstract fields like astronomy [69].

2.4. Homework, Academic Achievement, Interest in the Subject, and Chatbot-Supported Learning

Homework reinforces classroom learning, fosters independent study, and boosts academic achievement [70]. However, in abstract subjects like science, students often need additional structured guidance [71]. Stimulating curiosity and engagement is critical for effective science learning [72], and digital tools offering structured support play a key role [73].

Research shows that regular, structured homework enhances retention, self-regulation, and long-term learning [74, 75]. Yet challenges like conceptual misunderstandings, unreliable information access, and lack of timely guidance persist [76]. Traditional resources may lead to information overload [55].

Interactive digital tools providing real-time, structured support ease learning and reduce cognitive load [77, 78]. In this context, chatbots offer targeted information efficiently, improving both learning and motivation [54, 79]. They foster active participation and sustain engagement [36, 16, 3].

Rule-based chatbots, in particular, deepen understanding through structured guidance [80–82]. Research identifies three key benefits of chatbot use in homework: correcting misconceptions [61], providing immediate feedback [116], and enhancing self-regulation [82]. Overall, chatbot-supported environments improve learning efficiency, reduce cognitive overload, and boost achievement and motivation, especially in science education [6, 83, 84].

III. AIM OF THE STUDY

Non-AI rule-based (NARB) chatbots can prevent information overload by providing only targeted information, allowing students to experience a focused and guided learning process [67]. In this context, a non-AI rule-based chatbot named BilgeCan Bot was developed to support middle school students with their astronomy-related homework. The aim of this study is to examine the effects of BilgeCan Bot on students' academic achievement in science courses and their interest in science. Additionally, by analyzing students' perceptions of chatbots, the study seeks to offer insights into how such guidance systems can be used more effectively in educational settings. Based on these aims, the following research questions are addressed:

- RQ1: When controlling for pre-test scores, is there a significant difference in post-test academic achievement and interest in science between students who used BilgeCan Bot (experimental group) and those who used the textbook (control group) while completing homework?
- RQ2: What are the perceptions of the experimental group students regarding BilgeCan Bot after the experimental implementation?
- RQ3: What is the nature of the conversations conducted by the experimental group students with BilgeCan Bot?
- RQ4: What challenges did the control group students face while completing their homework using the textbook?

IV. DESIGNING A HOMEWORK SUPPORT CHATBOT SYSTEM (BILGECAN BOT)

4.1. Overview of BilgeCan Bot

BilgeCan Bot is a web-based, rule-based chatbot system developed to provide homework support for middle school science courses. Designed based on Cognitive Load Theory, the system aims to prevent excessive cognitive load and facilitate learning by offering students only the information they need. Developed with a JavaScript-based front-end, the system can be accessed via an internet browser on either a computer or a mobile device (Figure 1). Users can register by selecting the "Create New Account" option and then log in using their email and password to access the system.

BilgeCan

Oturumunuzu başlatmak için giriş yapın

Figure 1. BilgeCan bot login

4.2. The Architecture and Operational Process of BilgeCan Bot

BilgeCan Bot's architecture consists of three components: data preparation, user interaction, and technological infrastructure (Figure 2). Learning materials are structured into JSON/XML formats via regex-based keyword extraction. Developed with PHP 7.4 and CodeIgniter 3, the system stores data in a MySQL 10.6 database. Students interact with the bot through a text-based web interface, with requests processed by an Apache 2.4 server. Using regex and keyword matching, BilgeCan Bot operates as a lightweight, fast, and efficient rule-based learning assistant.

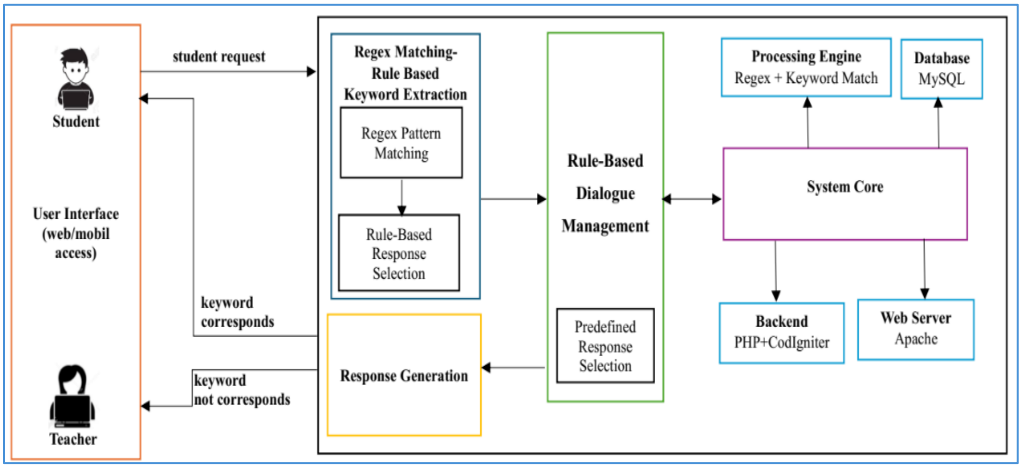


Figure 2. The system architecture of BilgeCan bot

BilgeCan Bot analyzes student questions using regex rules, identifies keywords, and retrieves appropriate astronomy-related answers from its database. If no answer is found, the bot checks teacher availability: if online, the student is directed to the teacher; if not, a notification is displayed. The bot ensures continuity by responding to subsequent questions (Figure 3).

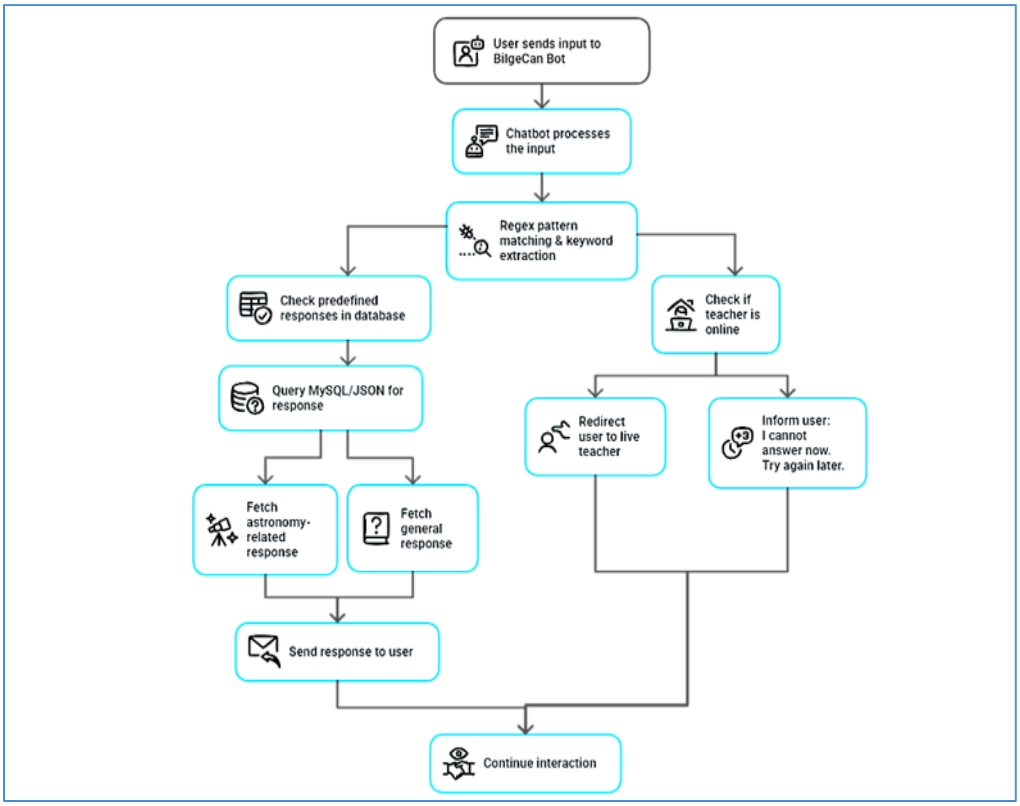
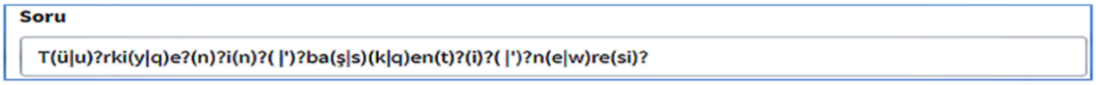


Figure 3. Process of response generation and operation in bilgecan bot

4.3. Dialogue Mechanism of BilgeCan Bot

BilgeCan Bot's dialogue system utilizes RegEx (Regular Expressions). RegEx is a powerful tool for recognizing and matching patterns in strings and is widely used in text processing applications. A similar approach was also employed in the first chatbot, ELIZA [85]. RegEx defines patterns by combining normal and special characters (*, +, ?), allowing the parsing and processing of user inputs [86]. Chatbots generate responses by matching user statements with predefined patterns. Wildcard characters make the patterns more flexible, enabling appropriate responses to questions phrased in different ways [87]. BilgeCan Bot is capable of providing flexible and effective responses using this method. An example question constructed with wildcard characters is shown in Figure 4.

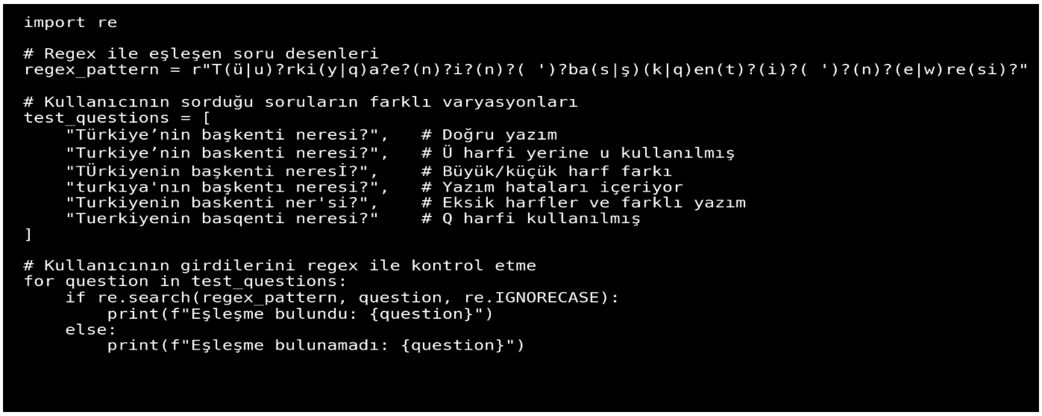


Soru

T(ü|u)?rki(y|q)e?(n)?i(n)?(|')?ba(ş|s)(k|q)en(t)?(i)?(|')?n(e|w)re(si)?

Figure 4. A Sample of wildcards used in the bilgecan bot

The example shown in Fig. 4 is designed to recognize different variations and errors in phrasing the question "What is the capital of Turkiye?" Thanks to the use of wildcard characters, variations such as "What is the capital of Turkiye" or "What is the capital of Turkiye's" can still be matched with the correct response (Figure 5).



```
import re

# Regex ile eşleşen soru desenleri
regex_pattern = r"T(ü|u)?rki(y|q)a?e?(n)?i?(n)?( ')?ba(s|ş)(k|q)en(t)?(i)?( ')?(n)?(e|w)re(si)?"

# Kullanıcının sorduğu soruların farklı varyasyonları
test_questions = [
    "Türkiye’nin başkenti neresi?",   # Doğru yazım
    "Turkiye’nin baskenti neresi?",   # Ü harfi yerine u kullanılmış
    "TÜrkiyenin başkenti neresİ?",    # Büyük/küçük harf farkı
    "turkıya'nın başkentı neresi?",   # Yazım hataları içeriyor
    "Turkiyenin baskenti ner'si?",    # Eksik harfler ve farklı yazım
    "Tuerkiyenin basqenti neresi?"    # Q harfi kullanılmış
]

# Kullanıcının girdilerini regex ile kontrol etme
for question in test_questions:
    if re.search(regex_pattern, question, re.IGNORECASE):
        print(f"Eşleşme bulundu: {question}")
    else:
        print(f"Eşleşme bulunamadı: {question}")
```

Figure 5. Regex-Based keyword matching code structure

4.4. User Interface Features of BilgeCan Bot

The teacher interface of BilgeCan Bot consists of two sections: system management and content control. The first section, shared with the student interface, includes menus such as "User Settings," "Chat," "Messages," "Active Teachers," and "User Guide," allowing teachers to monitor student activity (detailed in the "Student Interface Features" section). The second section, focusing on content and system management, includes the "Questions," "Surveys," "Tests," "System," "Feedback," and "Active Students" menus for content addition, user tracking, and system operations (Figure 6).

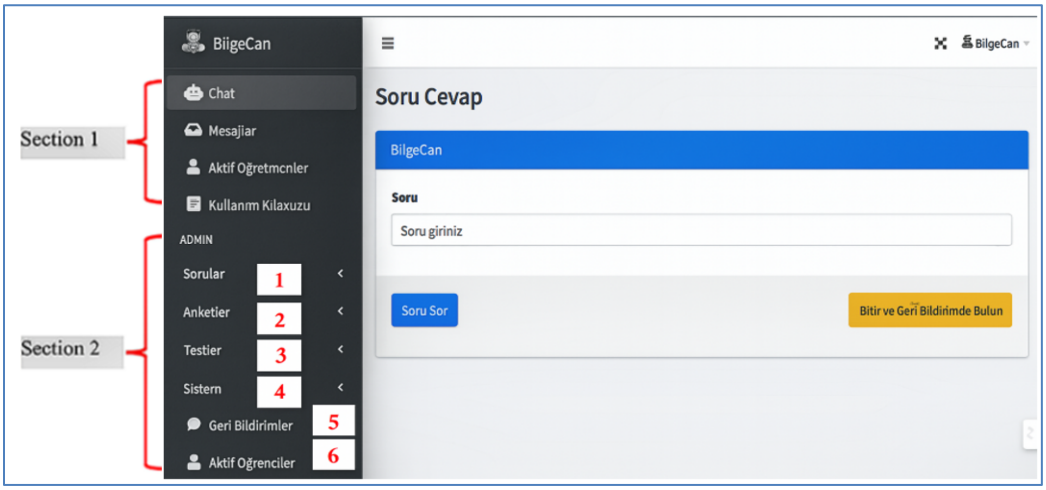


Figure 6. Teacher interface

The second section of the teacher interface includes tools for managing BilgeCan Bot's content and operations. The Questions menu (1) allows teachers to manage inquiries and add new content manually or via CSV files. The Surveys menu (2) enables survey creation and result review, planned for future use. The Tests menu (3) supports the creation of open-ended or multiple-choice assessments. The System menu (4) manages user registrations, roles, and user guide updates. The Feedback menu (5) collects student opinions for system improvement. The Active Students menu (6) allows real-time monitoring of online student activity.

The student interface includes menus for settings, notifications, chat, messages, test results, and active teachers, all supporting the learning process (Fig. 7). In the Chat menu (1), students can ask questions to BilgeCan Bot; if the bot cannot answer, they are directed to an online teacher. The Messages menu (2) displays past conversations. The Active Teachers menu (3) shows which teachers are online. The User Guide menu (4) provides usage instructions. The Finish and Provide Feedback menu (5) allows students to end the session and give feedback. The Enter Your Question field (6) is used to input messages.

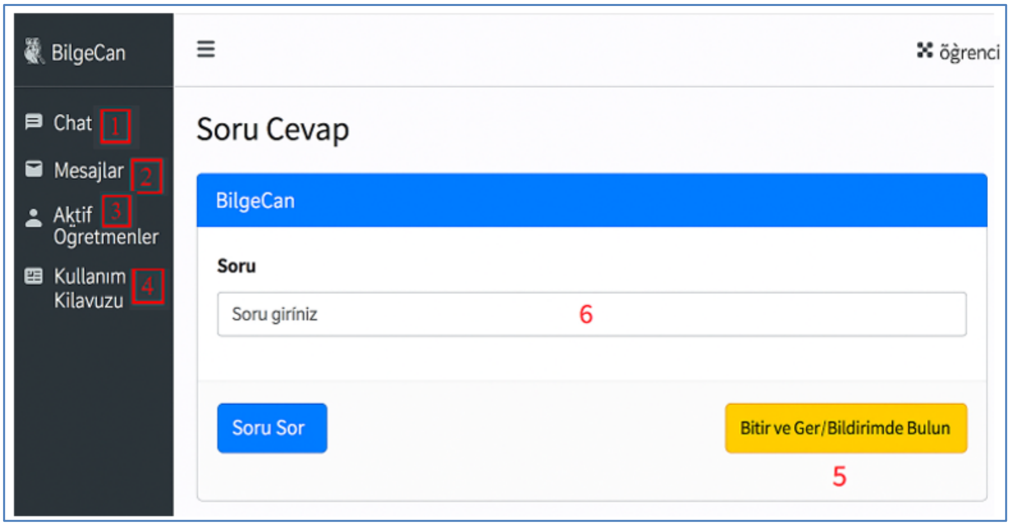
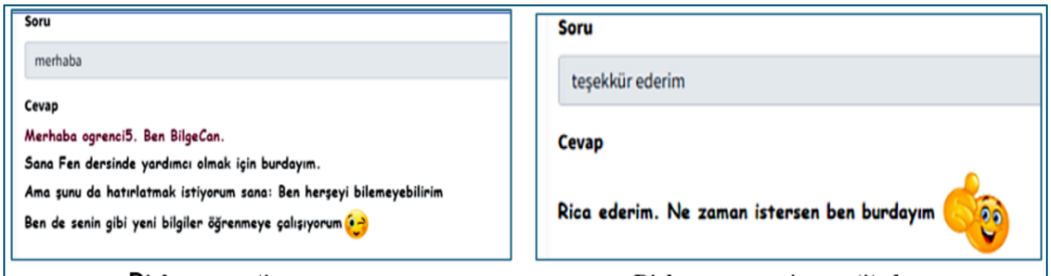


Figure 7. Student interface

4.5. Pilot Implementation with Bilgecan Bot and Preparation of Dialogues

Before the main study, a pilot implementation was conducted with sixth-grade students to test the usability of BilgeCan Bot and enhance student–teacher interaction. The bot was introduced, user registrations were completed, and interactions were maintained for one week. At the end of the process, it was determined that the bot supported conceptual understanding, its language was simplified, and some explanatory texts were added. In addition, short supportive phrases such as greetings, thanks, and guidance were incorporated to establish a friendly and guiding communication with students (Figure 8).



Soru

merhaba

Cevap

Merhaba ogrenci5. Ben BilgeCan.
Sana Fen dersinde yardımcı olmak için burdayım.
Ama şunu da hatırlatmak istiyorum sana: Ben herşeyi bilemeyebilirim
Ben de senin gibi yeni bilgiler öğrenmeye çalışıyorum

Soru

teşekkür ederim

Cevap

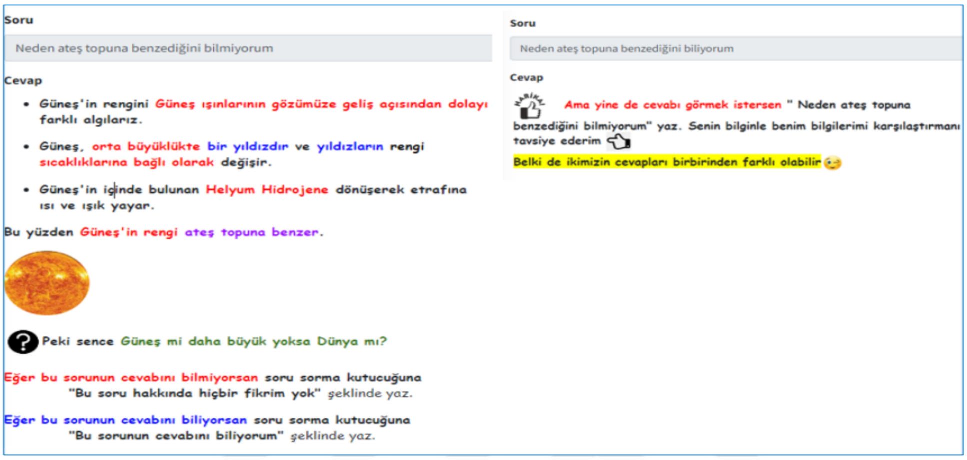
Rica ederim. Ne zaman istersen ben burdayım

Dialogue greeting

Dialogue expressing gratitude

Figure 8. Samples of auxiliary text

The main texts are focused on three key topics from the "Sun, Earth, and Moon" unit of the fifth-grade science curriculum: "The Structure and Features of the Sun," "The Structure and Features of the Moon," and "The Movements and Phases of the Moon." These texts were prepared in accordance with the learning outcomes specified in the 2018 TMoNE (Turkish Ministry of National Education) curriculum. For example, dialogues were developed to address objectives such as explaining the features of the Sun and comparing the size of the Sun with that of the Earth (Figure 9).



Soru

Neden ateş topuna benzediğini bilmiyorum

Cevap

- Güneş'in rengini Güneş ışınlarının gözümüze geliş açısından dolayı farklı algılarız.
- Güneş, orta büyüklükte bir yıldızdır ve yıldızların rengi sıcaklıklarına bağlı olarak değişir.
- Güneş'in içinde bulunan Helyum Hidrojene dönüşerek etrafına ısı ve ışık yayar.

Bu yüzden Güneş'in rengi ateş topuna benzer.

Peki sence Güneş mi daha büyük yoksa Dünya mı?

Eğer bu sorunun cevabını bilmiyorsan soru sorma kutucuğuna
"Bu soru hakkında hiçbir fikrim yok" şeklinde yaz.

Eğer bu sorunun cevabını biliyorsan soru sorma kutucuğuna
"Bu sorunun cevabını biliyorum" şeklinde yaz.

Soru

Neden ateş topuna benzediğini biliyorum

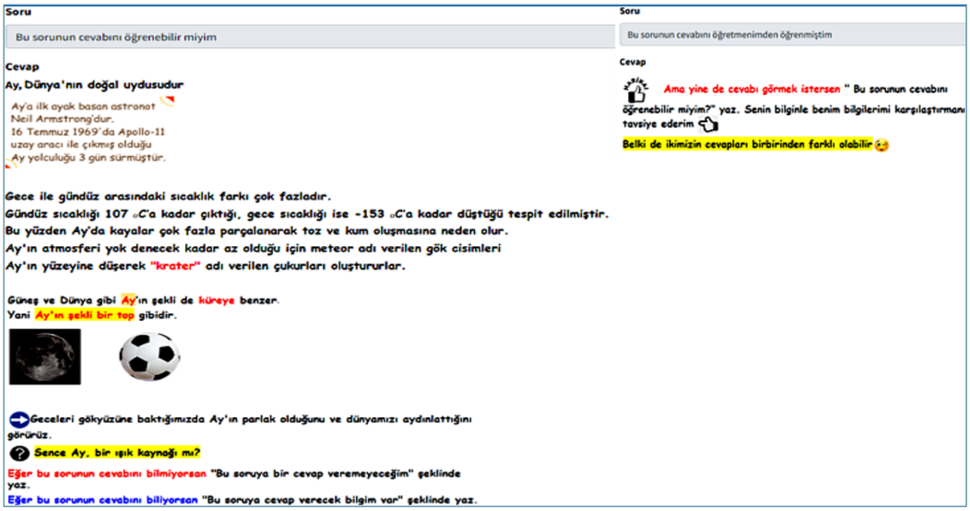
Cevap

Ama yine de cevabı görmek istersen " Neden ateş topuna benzediğini bilmiyorum" yaz. Senin bilginle benim bilgilerimi karşılaştırmanı tavsiye ederim

Belki de ikimizin cevapları birbirinden farklı olabilir

Figure 9. A sample of dialogue on the topic "the structure and properties of the sun"

Dialogues were created for the topics "The Structure and Features of the Moon" and "The Movements and Phases of the Moon" to support students' understanding of the Moon's characteristics, potential for life, and its movements (Figure 10). The dialogues encouraged students to assess their knowledge and compare it with the bot's responses. Regardless of their confidence, students were prompted to review key concepts, with clear explanations provided when needed. Alternative phrasing allowed students to access information through differently worded questions. Guiding questions were also included to enhance engagement and promote natural learning.



Soru

Bu sorunun cevabını öğrenebilir miyim

Cevap

Ay, Dünya'nın doğal uydusudur

Ay'a ilk ayak basan astronot
Neil Armstrong'dur.
16 Temmuz 1969'da Apollo-11
uzay aracı ile çıkmış olduğu
Ay yolculuğu 3 gün sürmüştür.

Gece ile gündüz arasındaki sıcaklık farkı çok fazladır.
Gündüz sıcaklığı 107 ₒC'a kadar çıktığı, gece sıcaklığı ise -153 ₒC'a kadar düştüğü tespit edilmiştir.
Bu yüzden Ay'da kayalar çok fazla parçalanarak toz ve kum oluşmasına neden olur.
Ay'ın atmosferi yok denecek kadar az olduğu için meteor adı verilen gök cisimleri
Ay'ın yüzeyine düşerek "krater" adı verilen çukurları oluştururlar.

Güneş ve Dünya gibi Ay'ın şekli de küreye benzer.
Yani Ay'ın şekli bir top gibidir.

Geceleri gökyüzüne baktığımızda Ay'ın parlak olduğunu ve dünyamızı aydınlattığını görürüz.

Sence Ay, bir ışık kaynağı mı?

Eğer bu sorunun cevabını bilmiyorsan "Bu soruya bir cevap veremeyeceğim" şeklinde yaz.

Eğer bu sorunun cevabını biliyorsan "Bu soruya cevap verecek bilgim var" şeklinde yaz.

Soru

Bu sorunun cevabını öğretmenimden öğrenmiştim

Cevap

Ama yine de cevabı görmek istersen " Bu sorunun cevabını öğrenebilir miyim?" yaz. Senin bilginle benim bilgilerimi karşılaştırmanı tavsiye ederim

Belki de ikimizin cevapları birbirinden farklı olabilir

Figure 10. A sample of a dialogue on "the structure and properties of the moon and the movements and phases of the moon"

V. METHOD

5.1. Research Design

An explanatory sequential (quantitative → qualitative) mixed methods design was used [88]. In the quantitative phase, a pre-test–post-test quasi-experimental design with a control group investigated RQ1 [89], focusing on differences between adjusted post-test means based on pre-test scores. In the qualitative phase, a phenomenological design explored RQ2, RQ3, and RQ4 [90]. Data were collected through semi-structured interviews (Appendix C), BilgeCan Bot conversation records, and control group teacher interviews. Sequential data collection strengthened the interaction between findings, with priority given to quantitative data for systematic analysis and qualitative data for in-depth interpretation.

5.2. Participants

A total of 52 fifth-grade students from a public middle school in Turkey participated in this study. The physical conditions of the classrooms were identical. Students were randomly assigned to experimental conditions. They were divided equally into two groups: the control group (Class A: 26 students; 10 boys and 16 girls) and the experimental group (Class B: 26 students; 12 boys and 14 girls). The students in both groups shared similar socioeconomic backgrounds.

5.3. Experimental Procedure

This study was conducted during the 2022–2023 academic year with fifth-grade students at a public middle school in Turkiye, within the "Sun, Earth, and Moon" unit. Necessary permissions were obtained from the university ethics committee, the Ministry of National Education, and parents through informed consent. Student confidentiality was ensured by assigning unique codes. The study followed the Turkish Ministry of National Education's 2018 science curriculum. The experimental process spanned seven weeks and involved BilgeCan Bot use, data collection, and analysis. Students in the experimental group used the bot outside of school to reinforce lesson topics, receiving predefined responses or being directed to teachers when needed. Teachers were available online on scheduled days, and class hour differences between groups were considered. Lessons for the experimental and control groups were conducted by different teachers using common lesson plans to minimize instructional differences. Classroom activities were monitored through checklists (Appendix A) and teacher interviews, though direct observation of interactions was not conducted. The experimental group teacher introduced BilgeCan Bot, provided usage instructions, and emphasized academic responsibilities. Students were encouraged to use the bot twice a week and to contact their teacher when necessary. Homework aligned with weekly learning objectives was assigned to both groups, with the experimental group supported by the bot and the control group using textbooks. The seven-week implementation schedule is summarized in Table 1.

The weekly implementation schedule was as follows:

- 1st Week: Pre-tests (AAT, SISL) administered; "Structure and Features of the Sun" covered.
- 2nd Week: Homework evaluation; "Comparison of Sun and Earth Sizes" and model-building activities.
- 3rd Week: "Structure and Features of the Moon" and scaling activities.
- 4th Week: "Movements and Phases of the Moon" with modeling, posters, and presentations.
- 5th Week: "Moon’s Phases and Revolution" explored through interactive activities.
- 6th Week: "Sun, Earth, and Moon System" modeled; group work and presentations conducted.
- 7th Week: Post-tests (AAT, SISL) administered; semi-structured interviews conducted with the experimental group and teacher interviews conducted for the control group.

Table 1. Procedure of the Implementation Period of the Research

Week	In-school activities (teaching lessons)				Out-of-school activities (do homework)	
	Experimental and control groups				Experimental group	Control gr
	Learning outcome	Learning activities	Teaching methods and techniques	Teaching materials	Use BilgeCan Bot	Use Textbool
1st	Pre-test and pre-questionnaire (achievement, interest)					
	Explains the properties of the Sun	“I will know the Sun better”	Expository teaching, question and answer technique, brainstorming, watching videos	Textbook, images, videos and animations	Worksheet-1	Worksheet-
2nd	Prepares a model to compare the Sun’s and Earth’s sizes	“We discover the size of the Earth and the Sun” “our earth and sun model”	Project based learning, brainstorming	Textbook, modeling materials	Worksheet-2	Worksheet-
3rd	Describes the size and geometric shape of the Moon	“I Will Know the Moon Better-1”	Expository teaching, question and answer technique, brainstorming, watching videos	Textbook, images, videos and animations	Worksheet-3	Worksheet-
4th	Explains the rotation and orbit movements of the Moon	“I Will Know the Moon Better-2”	Project based learning, Brainstorming Question and answer Technique, Watching videos	Textbook, images, videos and animations	Worksheet-4	Worksheet-
5th	Explains the relationship between the phases of the Moon and the Moon’s orbit around the Earth	“I Will Know the Moon Better-3”		Textbook, images, videos and animations	Worksheet-5	Worksheet-
6th	Prepares a model representing the motion of the Sun, Earth, and Moon relative to each other	“Our Earth and Moon Model”		Modeling materials	Worksheet-6	Worksheet-
7th	Completion of posttests, semi-structured interviews and receive feedback					

5.4. Measurements

Quantitative data in this study were collected using the Academic Achievement Test (AAT) and the Science Interest Scale (SIS). Qualitative data were gathered through a Semi-Structured Interview Form (SSIF) administered to the experimental group students after the intervention to collect their opinions about the chatbot. In addition, weekly conversation records between experimental group students and the chatbot, as well as feedback interviews with the control group's teacher regarding homework, were used as part of the qualitative data.

5.4.1. Academic Achievement Test

An Academic Achievement Test (AAT) was developed for the "Sun, Earth, and Moon" unit of the fifth-grade science course within the scope of this study. The test covers the topics "The Structure and Features of the Sun," "The Structure and Features of the Moon," and "The Movements and Phases of the Moon." Items were selected from various sources based on the learning outcomes specified in the TMoNE (2018) Science Curriculum. Initially, 40 items were prepared, and content validity was ensured using a table of specifications. The test was reviewed by an academician, two science teachers, and one Turkish language teacher, and revised according to expert feedback. In the pilot implementation, three items were eliminated, and the remaining 37 items were analyzed using the TAP software. An additional 10 items were removed based on difficulty and discrimination indices, resulting in a final test with 27 items. Final analyses showed a mean item difficulty of 0.63, an item-total correlation of 0.43, and a KR-20 reliability coefficient of 0.88. These values indicate that the test is both valid and reliable [91–95].

5.4.2. Scale of Interest in Science Lesson

In this study, the 27-item, 5-point Likert-type "Scale of Interest in Science Lesson (SISL)," developed by Laçin Şimşek and Nuhoğlu (2009), was used to determine the science interest levels of fifth-grade students regarding educational chatbots. The scale consists of six sub-dimensions: exploring nature, understanding cause-and-effect relationships, conducting observations, relating concepts to daily life, following mass media, and individual interest. Participants can score between 27 and 135 points on the scale. The Cronbach's alpha reliability coefficient was reported as 0.79 in the original development study, while it was calculated as 0.74 for the pre-test and 0.78 for the post-test in this research. These results indicate that the scale is reliable [91]. Moreover, the findings are consistent with similar studies in the literature [97–99]. However, due to the limited sample size (N=52), a content validity analysis could not be performed, which is noted as a limitation of the study.

5.4.3. Semi-Structured Interview Form

This form included 13 open-ended questions exploring students' experiences, expectations, and opinions about chatbot use (see Appendix B). Topics covered previous chatbot experiences, difficulties, liked and disliked features, improvement suggestions, and impacts on learning. The chatbot's effects on learning, questioning, thinking skills, attitudes toward science, and willingness to use it in other subjects were also evaluated. Its potential to support independent learning at home was examined. Student responses were analyzed using content analysis, with codes, categories, and themes developed. Reliability was ensured using Miles and Huberman's formula [100], with an inter-rater agreement of 86.9%, exceeding the 80% acceptability threshold.

5.4.4. Weekly Chat Records of Experimental Group Students with BilgeCan Bot

The weekly conversation records between experimental group students and BilgeCan Bot were analyzed to identify students' interaction patterns with the chatbot, their information-seeking behaviors, and the challenges they encountered. In this process, both descriptive and thematic analysis methods were used to examine students' learning behaviors and the chatbot's contributions to the educational process. To assess the reliability of the analysis, inter-coder agreement was calculated, and consistency across the coding of student conversation records was determined. According to the classification scale for kappa values proposed by Landis and Koch [101], the inter-rater agreement was $\kappa = 0.85$, 95% CI [0.55, 1.12], $p < .001$, indicating an almost perfect level of agreement between coders. Sample conversations between students and both BilgeCan Bot and the live teacher are presented in Appendix C.

5.4.5. Homework feedback interviews with the control group teacher

Regular interviews were conducted with the teacher responsible for the control group to discuss the difficulties students faced while completing their weekly homework assignments. Feedback collected from students by the teacher was analyzed using thematic analysis. To assess the reliability of the analysis, inter-coder agreement among three coders was calculated, and the consistency of thematic coding processes was determined. Krippendorff's alpha coefficient was calculated as $\alpha = 0.48$, indicating a moderate level of agreement among raters [102].

5.5. Data analysis

5.5.1. Analysis of Quantitative Data

In this study, both quantitative and qualitative data analysis methods were used to examine the effects of BilgeCan Bot on students' academic achievement and interest in the course. For the analysis of quantitative data, descriptive and inferential statistical methods were applied. Within the scope of descriptive statistics, the mean, standard deviation, skewness, and kurtosis values of the students' scores were calculated, and group differences were compared. Subsequently, in the inferential statistics phase, a One-Way Multivariate Analysis of Covariance (MANCOVA) was conducted to examine the effect of BilgeCan Bot on students' academic achievement and interest in scienceMANCOVA allows testing the statistically significant effect of an independent variable on two or more dependent variables [103–105]. In this study, the post-test scores of the Academic Achievement Test (AAT_2) and the Scale of Interest in Science Lesson ($SISL_2$) were defined as dependent variables. During the analysis, the pre-test scores of the Scale of Interest in Science Lesson ($SISL_1$) were controlled as a covariate to more precisely assess the effect of the independent variable on the dependent variables.

5.5.2. Analysis of Qualitative Data

In this study, students' views on chatbot use in science education were examined. Qualitative data were gathered through semi-structured interviews with experimental group students, BilgeCan Bot conversation records, and control group teacher interviews. Student interviews were analyzed through content analysis [106], with students anonymized as S1, S2, etc. BilgeCan Bot records underwent descriptive and thematic analyses. Variables such as Interaction Duration (ID), Number of Questions to the Bot (NQB), and Number of Dialogues with the Live Teacher (NDLT) were calculated based on estimated interaction times. Dialogue content was further examined using TF-IDF and n-gram techniques to extract key concepts. Student questions were classified according to Bloom's cognitive taxonomy [107]. Control group difficulties were thematically analyzed based on teacher interviews. Two researchers independently coded all qualitative data. Initial codes were generated and iteratively refined through discussion, resulting in broader themes. To ensure the reliability of coding, inter-coder agreement was calculated using Cohen's Kappa coefficient (κ = 0.84), indicating substantial agreement.

VI. Results

6.1. Descriptive Statistics

The descriptive statistics for the AAT_1, AAT_2, $SISL_1$, and $SISL_2$ test results of the experimental and control groups are presented in Table 2.

Table 2. Descriptive Statistics Results of the Groups for The Tests

	Experimental group				Control group			
	AAT_1	AAT_2	$SISL_1$	$SISL_2$	AAT_1	AAT_2	$SISL_1$	$SISL_2$
N	26	26	26	26	26	26	26	26
Mean	12.08	20.12	98.62	106.62	11.27	17.27	96.62	94.04
SD	1.47	2.88	8.59	13.40	2.05	2.66	14.73	12.40
Skewness	-.14	.25	.18	-.70	.36	.14	-.14	.03
Kurtosis	-.42	-.84	-.44	-.02	.39	-1.16	-1.05	-1.03

The descriptive statistics for the AAT_1, AAT_2, $SISL_1$, and $SISL_2$ test results of the experimental and control groups are presented in Table 2. AAT scores were evaluated out of 27 points, with each correct answer awarded 1 point. The experimental group's mean AAT_1 score was 12.08, and their mean AAT_2 score was 20.12. In the control group, these values were 11.27 and 17.27, respectively. Accordingly, an 8-point increase was observed in the experimental group, while a 6-point increase was recorded in the control group. SISL scores were evaluated out of 135 points. The experimental group's mean SISL score increased from 98.62 to 106.62, whereas the control group's mean score decreased from 96.62 to 94.04. These findings indicate that the experimental group showed greater improvement in both academic achievement and interest in science compared to the control group. Additionally, the skewness and kurtosis values for AAT_1, AAT_2, $SISL_1$, and $SISL_2$ variables were within the range of -1.5 to +1.5, suggesting that the data were normally distributed [108, 109].

6.2. Inferential statistics

6.2.1 RQ1: Is There a Significant Difference in The Post-Test Scores of Academic Achievement and Science Interest Between Students Who Used Bilgecan Bot (Experimental Group) and Those Who Used The Textbook (Control Group), After Controlling for Pre-Test Scores?

The effects of BilgeCan Bot on students' academic achievement (AAT_2) and science interest ($SISL_2$) were analyzed using MANCOVA and ANCOVA. AAT_1 and $SISL_1$ were initially considered potential covariates, and their correlations with the dependent variables were tested. AAT_1 showed no significant correlation with AAT_2 or $SISL_2$ ($p > .01$), whereas $SISL_1$ was significantly correlated with AAT_2 ($r = .035$, $p < .01$), suggesting its suitability as a covariate. Thus, $SISL_1$ was included in further analyses, while AAT_1 was excluded. MANCOVA assumptions—normality, homogeneity of variance, regression slope homogeneity, multicollinearity, and independence—were tested and met. The analysis showed a significant effect of group on the combined dependent variables (Wilks' Lambda = .66, F(2, 48) = 12.04, $p < .05$, $\eta p^2 = .34$, observed power = .99), indicating that BilgeCan Bot use explained 34% of the variance in AAT_2 and $SISL_2$. To explore effects on each variable, separate ANCOVAs were conducted. The results are presented in Table 3.

Table 3. ANCOVA Results

Source	DV	Type III sum of squares	df	Mean square	F	Sig	Partial Eta squared	Observed power
Treatment	AAT_2	92,75	1	92,75	13.60	.00	.22	.96
	$SISL_2$	2194,20	1	2194,20	13.53	.00	.22	.96

Separate ANCOVA results showed significant differences in favor of the experimental group for both academic achievement (AAT_2: F(1, 48) = 13.60, $p < .05$, $\eta p^2 = .22$, observed power = .96) and science interest ($SISL_2$: F(1, 48) = 13.53, $p < .05$, $\eta p^2 = .22$, observed power = .96), indicating both theoretical and practical significance [110].

Adjusted means revealed that the experimental group outperformed the control group in academic achievement (M = 20.12 vs. 17.26, SE = .52) and science interest (M = 106.81 vs. 93.84, SE = 2.57). Pairwise comparisons confirmed that these differences were statistically significant for AAT_2 (MD = 2.86, SE = .75, $p < .05$, 95% CI [1.36, 4.36]) and for $SISL_2$ (MD = 12.97, SE = 3.69, $p < .05$, 95% CI [5.56, 20.39]).

Taken together, these findings indicate that the use of BilgeCan Bot significantly enhanced students' academic achievement and their interest in science, and that this effect was in favor of the experimental group.

6.2.2 RQ2: What are the Experimental Group Students' Opinions About Bilgecan Bot After the Intervention?

After the implementation, semi-structured interviews (SSIF) were conducted to gather the experimental group students' opinions about BilgeCan Bot. Responses were analyzed using content analysis, with similar expressions coded and grouped into themes, categories, and codes, supported by frequency values. Six of the 13 SSIF questions were associated with the theme "Using Chatbots for Learning," and the related findings are summarized in Table 4 (see Appendix D).

According to Table 4, all students reported that they used a chatbot for the first time, felt comfortable during the interaction, and found BilgeCan Bot easy to use. Sample student responses within this category are as follows:

“"I felt comfortable while using BilgeCan Bot and asked my questions without hesitation.” (S12)

“BilgeCan answered my questions accurately and in detail. I liked that the responses came immediately.” (S3)

“BilgeCan Bot provided information on different topics; it was helpful, and I was able to use it very easily.” (S9)

“It was simple to use, and I could easily enter the questions I wanted to ask, and the bot answered them correctly.” (S5)

According to Table 4, most students reported encountering no problems while using the chatbot, receiving quick responses, finding it enjoyable and friendly, and expressing willingness to use it for other science topics and subjects. Example student responses for this category are presented below:

“No, I did not encounter any problems.” (S5)

“I liked the quick responses of BilgeCan Bot.” (S18)

“Thanks to BilgeCan Bot, I find the science course more interesting and important.” (Ö24)

“It was great that the active teachers sometimes responded to us.” (S4)

“I wish all subjects could be taught with BilgeCan Bot.” (S6)

Students' criticisms of the chatbot mainly concerned difficulties in connecting with active teachers and the limited scope of available information. Example student responses for this category are presented below:

“It doesn't answer some of my questions and cannot connect me to an active teacher.” (S1)

“Sometimes when I log in, it says that there is no available teacher.” (S19)

The last seven questions in the SSIF belong to the theme "Benefits of Chatbots in Learning," and the findings related to this theme are presented in Table 5 (see Appendix E).

According to Table 5, most students reported that BilgeCan Bot supported their learning, improved their inquiry and thinking skills, made science lessons more engaging, and expressed a desire to use it for other topics. Some students also mentioned forming a sense of "small friendships" with the bot during solo study.Several students stated that they could understand complex topics more easily through the bot's simple and clear responses. For instance:

“Thanks to BilgeCan Bot, I learned more about the Sun, Earth, and Moon, and understood the topic better.” (S7)

Others emphasized how the bot encouraged question-asking behaviors and helped improve their critical thinking:

“Yes, it helped. Thanks to BilgeCan Bot, I learned how to ask questions about different concepts.” (S23)

“As I kept asking questions to BilgeCan Bot, I felt that I started thinking differently and asking better questions.” (S19)

The chatbot's role in increasing students' interest and motivation was also highlighted:

“BilgeCan Bot made the science lesson more fun and sparked more curiosity about the subject.” (S6)

“When I was stuck at home, I used BilgeCan Bot. It made me feel like I wasn't alone and helped me continue studying.”

(S15)

"I wish there was a BilgeCan Bot for the living things unit as well." (S3)

In addition, some students reflected on the lack of emotional connection, acknowledging that the bot had limitations in forming human-like bonds:

"BilgeCan Bot cannot make you feel like a real friend because it cannot form an emotional bond." (S9)

"Sometimes I wanted it to say something like 'well done,' but it just gave the answer and moved on." (S11)

Finally, a few students also commented on technical or contextual limitations, such as the bot's inability to understand complex or ambiguous questions:

"Sometimes it didn't understand what I meant. I wish I could ask the teacher at those moments." (S1)

These statements indicate that the students were generally positive about BilgeCan Bot and found it useful as an academic and motivational support tool in science lessons, but were also aware of emotional and contextual limitations. Overall, the fact that students wanted to use the bot in different science subjects reveals that the overall satisfaction level was high.

6.2.3 RQ3: What is the Content of the Conversations that the Experimental Group Students Had with Bilgecan Bot?

The experimental group's interaction frequency and content usage with BilgeCan Bot were analyzed through descriptive and content analysis. Students typically used the bot twice a week during out-of-school hours for homework. Figure 11 shows weekly usage: 11 students in week one, rising to 18 in week two, dropping to 14 in week three, increasing to 16 in week four, peaking at 22 in week five, and falling to 15 in week six. These findings indicate that chatbot usage fluctuated weekly.

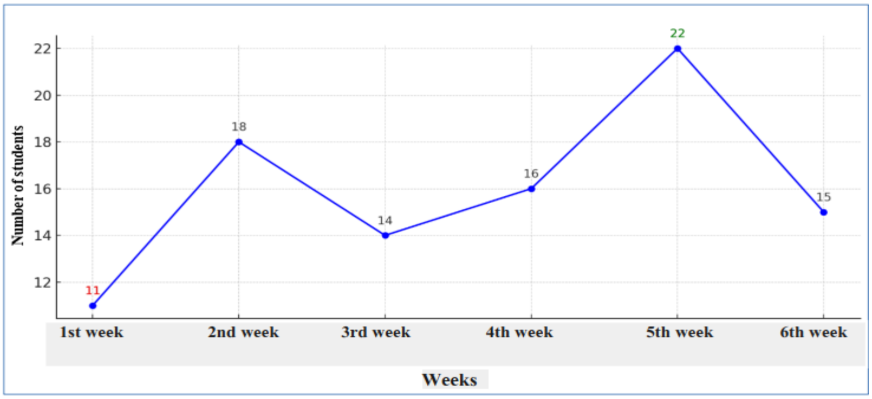


Figure 11. Number of students using bilgecan bot by week

As shown in Figure 12, the amount of time students spent with BilgeCan Bot varied across the weeks.

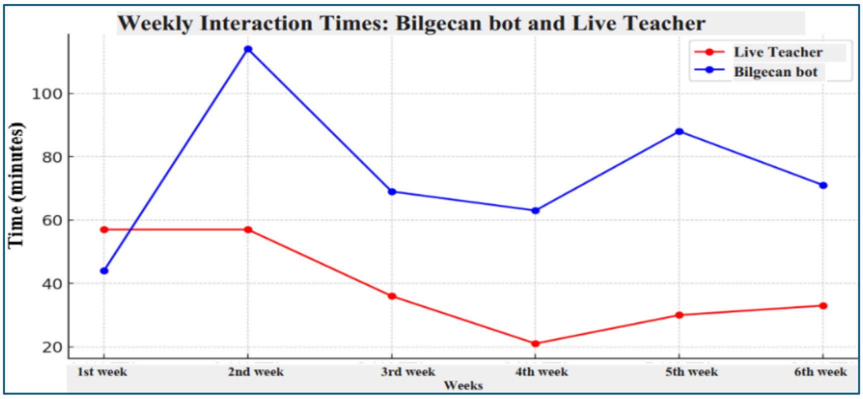


Figure 12. Interaction times per week with bilgecan bot and live tutor

Interaction time with BilgeCan Bot was lowest in the first week, peaked in the second, declined in the third and fourth, rose in the fifth, and dropped again in the sixth week. Interaction with live teachers remained stable initially, decreased mid-study, and increased in the final two weeks. These patterns indicate weekly shifts in students' preferences, with a notable rise in chatbot use during the fifth week.

Weekly interaction data were analyzed through variables such as Interaction Duration (ID), Number of Dialogues with BilgeCan Bot (NDBB), Number of Questions per Dialogue (NQDBB), Number of Questions Asked (NQABB), and Number of Dialogues with the Live Teacher (NDLT). S7 and S10 were the most active users, while S5 and S21 showed lower interaction. Changes in interaction time corresponded with fluctuations in the number of questions asked. Figure 13 displays the weekly comparison of total questions directed to BilgeCan Bot and the live teacher.

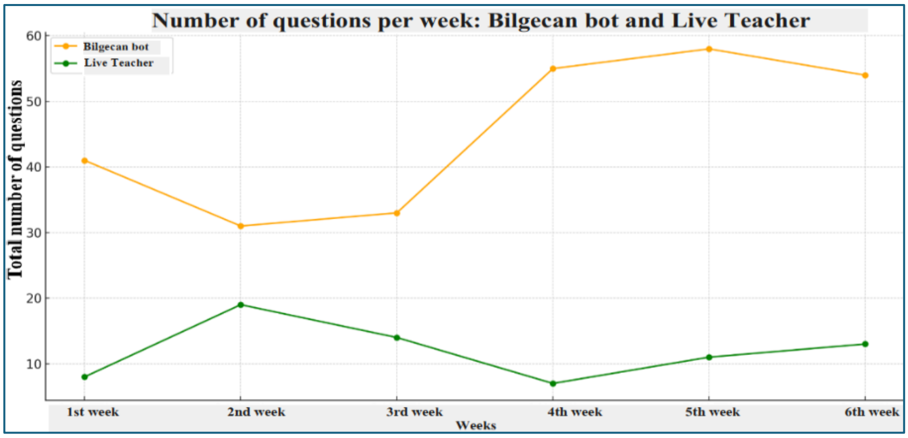


Figure 13. Weekly number of questions asked to BilgeCan bot and live teacher

When examining Fig. 13, it is evident that the number of questions directed to BilgeCan Bot was consistently higher each week than those directed to the live teacher, with a noticeable peak during the fourth and fifth weeks.

To further explore the content of students' interactions with BilgeCan Bot, a TF-IDF (Term Frequency–Inverse Document Frequency) analysis was conducted. Table 6 presents the top 10 keywords with the highest TF-IDF values identified across the six-week period (see Appendix F). According to Table 6, "Sun" dominated the first and second weeks, "Moon" became prominent in the third week, while "Moon's," "Moon," and "appears" were key in the fourth and fifth weeks. In the sixth week, keywords such as "Earth's," "time," and "rotation" stood out.

Additionally, conversation records were categorized according to Bloom's taxonomy (knowledge, comprehension, application, and analysis), and the results are presented in Table 7 (see Appendix G). The analysis showed that while knowledge-level questions increased over time, higher-order questions (comprehension, application, analysis) fluctuated, with notable peaks in certain weeks.

6.2.4 RQ4: What Difficulties did the Control Group Students Encounter While Completing Their Homework Using the Textbook?

Difficulties experienced by control group students during homework completion were identified through post-lesson teacher interviews. Students reported problems such as conceptual misunderstandings, lack of examples, and insufficient guidance. They found textbook explanations inadequate for topics like the Sun's geometric structure, rotational movement, size comparisons between the Sun and Earth, and the Moon's surface features. The phases of the Moon were especially confusing, with students expressing a need for step-by-step support. Similar issues arose in modeling activities due to the absence of guidance. Students also indicated difficulty in understanding textbook content and noted that their questions remained unanswered. These findings suggest that textbooks alone were insufficient for supporting the comprehension of abstract concepts. The results are presented in Table 8 (see Appendix H).

Table 8 shows that students in the control group experienced two main difficulties while completing their homework: challenges in visualization and modeling, and a lack of interaction and support. Students reported struggling to understand abstract concepts and expressed that they did not receive sufficient guidance. These findings highlight the need for more visual and interactive support in instructional materials.

VII. Discussion and Conclusion

This study investigated the effects of BilgeCan Bot, a Non-AI rule-based (NARB) chatbot designed to support science homework, on students' academic achievement and interest in science. It also explored experimental group students' perceptions of chatbot use and analyzed challenges faced by students relying solely on textbooks. Descriptive analyses revealed that students using BilgeCan Bot showed significantly greater improvements in both academic achievement and science interest compared to those using textbooks. MANCOVA results demonstrated that BilgeCan Bot significantly enhanced students' combined achievement and interest scores after controlling for prior interest, and ANCOVA results confirmed higher individual scores in the experimental group. Findings were interpreted within the framework of Cognitive Load Theory (CLT). Given the scarcity of studies focusing specifically on cognitive load in chatbot-based learning [111, 112], this research addresses an important gap. Although most prior work centers on AI-driven pedagogical agents [51, 113, 114], BilgeCan Bot—

despite lacking AI—meets the functional definition of a pedagogical agent by guiding cognitive processes and structuring homework tasks [115].

The findings indicated that BilgeCan Bot’s simple, cognitively streamlined structure contributed positively to learning. Quantitative data showed that students using BilgeCan Bot achieved higher scores than the control group. This result aligns with previous research highlighting that structured digital guidance facilitates learning [51, 113]. Pedagogical agents enhance content interaction and academic performance [116, 117]. Wu et al. (2024) similarly demonstrated that guided learning enriched by student participation improves academic achievement [52]. Thus, BilgeCan Bot optimized students’ cognitive resources, leading to achievement gains.

The study also found a statistically significant increase in students’ interest in science. Qualitative data supported this, with students expressing that science lessons became more enjoyable through the bot. For example, one student remarked, “Science lessons were more fun thanks to the bot” (S6). This suggests that the bot fostered positive emotions toward learning. Bowman [118] and Dinçer & Doğanay [119] emphasized that personal, clear language and sustained attention through pedagogical agents are key factors in maintaining student interest. BilgeCan Bot’s communication style likely contributed to these effects.

Regarding homework completion, students completed their homework more independently: “Thanks to the bot, I completed my homework without asking the teacher” (S3). This finding aligns with studies showing that interactive systems enhance engagement and boost homework motivation by reducing feelings of isolation [120, 52, 121]. Structured access to information helped students control their learning, thereby reducing extraneous cognitive load.

BilgeCan Bot’s rule-based design minimized distractions by providing only essential information. Sakellariou et al. noted that highly complex AI systems can overwhelm students, whereas simple, goal-oriented designs support cognitive efficiency [122]. Chen et al. [117] and Mayer & Fiorella [123] also emphasized that overly interactive systems may increase cognitive burden in younger learners, favoring simpler designs.

Qualitative feedback indicated that the bot promoted active cognitive engagement: "As I asked questions, I started thinking about new things" (S23). This aligns with Sweller’s [57] and Mayer & Moreno’s [124] arguments that environments fostering meaningful cognitive processing are crucial for deep learning. Likewise, Ba et al. [125] and Bowman et al. [118] stressed the importance of pedagogical agents in supporting conceptual structure building. By encouraging inquiry, BilgeCan Bot fostered active learning engagement.

However, some students reported that the bot lacked emotional connection: “It’s not like a real friend” (S9). This reflects findings by Wang & Lo [116] and Brandt et al. [126], emphasizing that chatbots may struggle to meet adolescents' expectations for emotional interaction. Therefore, future designs may benefit from incorporating emotional responsiveness to enhance pedagogical effectiveness.

In conclusion, this study demonstrated that BilgeCan Bot, despite operating as a non-AI rule-based system, effectively supported student learning by managing cognitive load, fostering engagement, and improving academic achievement and science interest. Within the framework of CLT, BilgeCan Bot reduced extraneous cognitive load by delivering essential information, balanced intrinsic load through structured explanations, and promoted germane load by encouraging inquiry-based learning. These outcomes indicate that simple, structured chatbots can significantly enhance student learning without relying on complex AI systems, especially in abstract domains like astronomy. Future research could examine the long-term effects of rule-based chatbot support on students’ retention of science concepts, higher-order thinking skills, and intrinsic motivation, providing a broader understanding of its sustained educational impacts. From a policy perspective, it is recommended that national and regional education authorities pilot the use of rule-based chatbot systems, such as BilgeCan Bot, within science curricula to support homework completion, reduce cognitive overload, and enhance equitable access to structured academic guidance.

VIII. Limitations

Although this study offers valuable insights, it has limitations. The sample was limited to 52 students from a single school, reducing generalizability. The chatbot was applied only to the "Sun, Earth, and Moon" unit. As a non-AI system, it lacked flexibility and could not track interaction time. The seven-week duration restricted assessment of long-term outcomes. Additionally, variations in students’ internet access may have affected the level of interaction.

IX. Recommendations

The findings suggest that non-AI rule-based chatbots like BilgeCan Bot can effectively support structured homework in science education. Future studies should explore their application in other subjects and across different age groups. Comparative research with AI-based systems is recommended within the Cognitive Load Theory framework. Integrating multimedia features may enhance user experience. Expanding chatbot functionality to involve teachers and parents could increase impact. Longitudinal studies are needed to assess effects on retention and higher-order thinking. Additionally, in-service training should be provided to help teachers integrate chatbots effectively into instructional practices.

Acknowledgments
This study was adapted from the first author's PhD thesis.

Authors’ Contributions
TY conducted this study under the supervision of İD as part of a doctoral dissertation. TY designed the study, collected and analyzed the data, and drafted the manuscript in consultation with İD. Both authors read and approved the final manuscript.

Competing Interests
The authors declare that they have no conflict of interest.

References

[1] L. Wang, Z. Liu, A. Liu, and F. Tao, "Artificial intelligence in product lifecycle management," *Int. J. Adv. Manuf. Technol.*, vol. 114, no. 3–4, pp. 771–796, 2021, doi: 10.1007/s00170-021-06882-1.

[2] S. S. Barsoum, M. S. Elnagar, and B. M. Awad, "The effectiveness of using a cognitive style-based chatbot in developing science concepts and critical thinking skills among preparatory school pupils," *Eur. Sci. J.*, vol. 18, no. 22, p. 52, 2022, doi: 10.19044/esj.2022.v18n22p52.

[3] R. Wu and Z. Yu, "Do AI chatbots improve students’ learning outcomes? Evidence from a meta-analysis," *Br. J. Educ. Technol.*, vol. 55, no. 1, pp. 10–33, 2023, doi: 10.1111/bjet.13334.

[4] Y. Liang, S. Ma, and C. Lin, "Chatbotification for web information systems: A pattern-based approach," in *Proc. 2022 IEEE 46th Annu. Comput., Softw. Appl. Conf. (COMPSAC)*, pp. 2290–2295, 2024, doi: 10.1109/compsac61105.2024.00368.

[5] J. Q. Pérez, T. Daradoumis, and J. M. M. Puig, "Rediscovering the use of chatbots in education: A systematic literature review," *Comput. Appl. Eng. Educ.*, vol. 28, no. 6, pp. 1549–1565, 2020, doi: 10.1002/cae.22326.

[6] N. L. Schroeder, O. O. Adesope, and R. B. Gilbert, "How effective are pedagogical agents for learning? A meta-analytic review," *J. Educ. Comput. Res.*, vol. 49, no. 1, pp. 1–39, 2013, doi: 10.2190/ec.49.1.a.

[7] J. Schmidhuber, S. Schlögl, and C. Ploder, "Cognitive load and productivity implications in human-chatbot interaction," in *Proc. 2021 IEEE 2nd Int. Conf. Human-Machine Syst. (ICHMS)*, pp. 1–6, 2021, doi: 10.1109/ICHMS53169.2021.9582445.

[8] X. Deng and Z. Yu, "A meta-analysis and systematic review of the effect of chatbot technology use in sustainable education," *Sustainability*, vol. 15, no. 4, p. 2940, 2023, doi: 10.3390/su15042940.

[9] F. R. Baskara, "Chatbots and flipped learning: Enhancing student engagement and learning outcomes through personalised support and collaboration," *Int. J. Recent Educ. Res.*, vol. 4, no. 2, pp. 223–238, 2023, doi: 10.46245/ijorer.v4i2.331.

[10] F. Almasri, "Exploring the impact of artificial intelligence in teaching and learning of science: A systematic review of empirical research," *Res. Sci. Educ.*, vol. 54, no. 5, pp. 977–997, 2024, doi: 10.1007/s11165-024-10176-3.

[11] V. Zenuni-Idrizi and L. Lutfiu-Kadriu, "The impact of regularly checking homework on students’ success in primary school," *Univ. Teach. Educ. J.*, vol. 6, no. 11–12, pp. 178–183, 2024, doi: 10.62792/ut.education.v6.i11-12.p2653.

[12] E. N. Mutayomba, "The impact of homework on teaching and learning in secondary schools: Case of Adventist secondary schools in Goma/D.R. Congo," *Int. J. Soc. Sci. Humanit. Invention*, vol. 10, no. 3, pp. 7823–7831, 2023, doi: 10.18535/ijsshi/v10i03.10.

[13] S. B. Payne and E. Swanson, "Supporting families to motivate their middle school student during homework time," *Teach. Except. Child.*, vol. 55, no. 6, pp. 422–430, 2022, doi: 10.1177/00400599221099656.

[14] L. Corno, "Homework is a complicated thing," *Educ. Res.*, vol. 25, no. 8, pp. 27–30, 1996, doi: 10.3102/0013189x025008027.

[15] J. S. Stephen, "Identifying support systems and resources," in *Springer Texts in Education*, pp. 45–57, 2024, doi: 10.1007/978-3-031-54439-2_4.

[16] J. Lee, T. An, H. Chu, H. Hong, and S. N. Martin, "Improving science conceptual understanding and attitudes in elementary science classes through the development and application of a rule-based AI chatbot," *Asia-Pac. Sci. Educ.*, vol. 9, no. 2, pp. 365–412, 2023, doi: 10.1163/23641177-bja10070.

[17] M. Taşcan, "The effect of science activities developed on astronomy education on fifth-grade students’ spatial skills and academic achievement," Ph.D. dissertation, İnönü Univ., 2019.

[18] F. Kara and N. Kefeli, "Perceptions of fifth-grade secondary school students towards the Sun, Earth, and Moon," *Necatibey Fac. Educ. Electron. J. Sci. Math. Educ.*, vol. 13, no. 2, pp. 1000–1014, 2019, doi: 10.17522/balikesirnef.560113.

[19] N. Demir, "The effect of the 5E learning model on seventh-grade students’ academic achievement and attitudes in teaching astronomy topics," Ph.D. dissertation, Erciyes Univ., 2020.

[20] J. Xu, "Why do students have difficulties completing homework? The need for homework management," *J. Educ. Train. Stud.*, vol. 1, no. 1, 2013, doi: 10.11114/jets.v1i1.78.

[21] A. Dhanani, "Using search engines in the teaching and learning process," *PEOPLE: Int. J. Soc. Sci.*, vol. 6, no. 1, pp. 682–689, 2020, doi: 10.20319/PIJSS.2020.61.682689.

[22] S. Abbasi, H. Kazi, and N. N. Hussaini, "Effect of chatbot systems on students' learning outcomes," *Sylwan*, vol. 163, no. 10, pp. 49–63, 2019. [Online]. Available: https://www.researchgate.net/publication/336373880_Effect_of_Chatbot_Systems_on_Student's_Learning_Outcomes

[23] B. A. Shawar and E. Atwell, "Chatbots: Are they really useful?" *J. Lang. Technol. Comput. Linguist.*, vol. 22, no. 1, pp. 29–49, 2007. [Online]. Available: https://jlcl.org/article/download/88/86

[24] S. S. Shaffie, M. F. Omardin, S. R. Ismail, A. A. Rahman, I. M. Ibrahim, and A. F. Najmuddin, "Attending teaching and learning with chatbot," *Int. J. Acad. Res. Prog. Educ. Dev.*, vol. 13, no. 1, 2024, doi: 10.6007/ijarped/v13-i1/21107.

[25] C. W. Okonkwo and A. Ade-Ibijola, "Chatbots applications in education: A systematic review," *Comput. Educ.: Artif. Intell.*, vol. 2, p. 100033, 2021, doi: 10.1016/j.caeai.2021.100033.

[26] H.-A. Bekkar and Y. Chtouki, "Chatbots in education: A systematic literature review," in *Proc. IEEE Int. Conf. Syst., Commun. Coding*, pp. 637–644, 2024, doi: 10.1109/icscc62041.2024.10690334.

[27] E. Loos, L. Ivan, and D. Leu, "“Save the Pacific Northwest tree octopus”: a hoax revisited. Or: How vulnerable are school children to fake news?" *Inf. Learn. Sci.*, vol. 119, no. 9/10, pp. 514–528, 2018.

[28] B. Lund, W. Ting, N. R. Mannuru, B. Nie, S. Shimray, and Z. Wang, "ChatGPT and a new academic reality: AI-written research papers and the ethics of the large language models in scholarly publishing," *SSRN Electron. J.*, 2023, doi: 10.2139/ssrn.4389887.

[29] E. Feigerlova, H. Hani, and E. Hothersall-Davies, "A systematic review of the impact of artificial intelligence on educational outcomes in health professions education," *BMC Med. Educ.*, vol. 25, no. 1, 2025, doi: 10.1186/s12909-025-06719-5.

[30] C. Papakostas, C. Troussas, A. Krouska, and C. Sgouropoulou, "A rule-based chatbot offering personalized guidance in computer programming education," in *Lect. Notes Comput. Sci.*, pp. 253–264, 2024, doi: 10.1007/978-3-031-63031-6_22.

[31] J. Sweller, "Cognitive load during problem solving: Effects on learning," *Cogn. Sci.*, vol. 12, no. 2, pp. 257–285, 1988, doi: 10.1016/0364-0213(88)90023-7.

[32] F. Paas, A. Renkl, and J. Sweller, "Cognitive load theory and instructional design: Recent developments," *Educ. Psychol.*, vol. 38, no. 1, pp. 1–4, 2003, doi: 10.1207/s15326985ep3801_1.

[33] J. Weizenbaum, "ELIZA—A computer program for the study of natural language communication between man and machine," 1966. [Online]. Available: https://web.stanford.edu/class/cs124/p36-weizenabaum.pdf

[34] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 2nd ed. Pearson Education, Inc., 1995.

[35] J. Yin, T. Goh, and Y. Hu, "Interactions with educational chatbots: The impact of induced emotions and students’ learning motivation," *Int. J. Educ. Technol. High. Educ.*, vol. 21, no. 1, 2024, doi: 10.1186/s41239-024-00480-3.

[36] H. Güldal and E. O. Dinçer, "Can rule-based educational chatbots be an acceptable alternative for students in higher education?" *Educ. Inf. Technol.*, 2024, doi: 10.1007/s10639-024-12977-5.

[37] M. Černý, "Educational psychology aspects of learning with chatbots without artificial intelligence: Suggestions for designers," *Eur. J. Invest. Health Psychol. Educ.*, vol. 13, no. 2, pp. 284–305, 2023, doi: 10.3390/ejihpe13020022.

[38] O. V. Deryugina, "Chatterbots," *Sci. Tech. Inf. Process.*, vol. 37, no. 2, pp. 143–147, 2010, doi: 10.3103/s0147688210020097.

[39] S. Roos, "Chatbots in education: A passing trend or a valuable pedagogical tool?" M.S. thesis, 2018. [Online]. Available: https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-355054

[40] Y. V. Lakshmi and I. Majid, "Chatbots in the education system," *Univ. News: A Wkly. J. High. Educ.*, vol. 60, no. 8, pp. 15–18, 2022. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4039535

[41] İ. Özkol, K. Doğan, and G. Köseali, "Artificial intelligence supported chatbot usage in ERMS applications," 2019. [Online]. Available: https://www.researchgate.net/publication/338336314_EBYS_Uygulamalarinda_Yapay_Zeka_Destekli_Chatbot_Sohbet_Robotu_Kullanimi_-_Artificial_Intelligence_Supported_Chatbot_Usage_in_ERMS_Applications

[42] L. C. Klopfenstein, S. Delpriori, S. Malatini, and A. Bogliolo, "The rise of bots: A survey of conversational interfaces, patterns, and paradigms," in *Proc. 2017 Conf. Designing Interactive Systems*, pp. 555–565, 2017. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/3064663.3064672

[43] S. Hussain, O. A. Sianaki, and N. Ababneh, "A survey on conversational agents/chatbots classification and design techniques," in *Adv. Intell. Syst. Comput.*, pp. 946–956, 2019, doi: 10.1007/978-3-030-15035-8_93.

[44] H. Chen, X. Liu, D. Yin, and J. Tang, "A survey on dialogue systems," *ACM SIGKDD Explor. Newsl.*, vol. 19, no. 2, pp. 25–35, 2017, doi: 10.1145/3166054.3166058.

[45] E. Adamopoulou and L. Moussiades, "Chatbots: History, technology, and applications," *Mach. Learn. Appl.*, vol. 2, p. 100006, 2020, doi: 10.1016/j.mlwa.2020.100006.

[46] D. Madhu, C. J. N. Jain, E. Sebastain, S. Shaji, and A. Ajayakumar, "A novel approach for medical assistance using trained chatbot," in *Proc. 2017 Int. Conf. Inventive Commun. Comput. Technol. (ICICCT)*, pp. 243–246, 2017, doi: 10.1109/ICICCT.2017.7975195.

[47] K. F. Hew, W. Huang, J. Du, and C. Jia, "Using chatbots to support student goal setting and social presence in fully online activities: Learner engagement and perceptions," *J. Comput. High. Educ.*, vol. 35, no. 1, pp. 40–68, 2023, doi: 10.1007/s12528-022-09338-x.

[48] S. A. Abdul-Kader and J. Woods, "Survey on chatbot design techniques in speech conversation systems," *Int. J. Adv. Comput. Sci. Appl.*, vol. 6, no. 7, 2015, doi: 10.14569/ijacsa.2015.060712.

[49] A. Badii, A. Adetoye, D. Patel, and K. Hameed, "Efficient FPGA−Based Regular Expression Pattern Matching," 2015.

[50] L. Bradeško and D. Mladenić, "A survey of chatbot systems through a Loebner Prize competition," in *Proc. Slovenian Lang. Technol. Soc. 8th Conf. Lang. Technol.*, vol. 2, pp. 34–37, 2012. [Online]. Available: https://nl.ijs.si/isjt12/proceedings/isjt2012_06.pdf

[51] R. Moreno and R. Mayer, "Interactive multimodal learning environments," *Educ. Psychol. Rev.*, vol. 19, no. 3, pp. 309–326, 2007, doi: 10.1007/s10648-007-9047-2.

[52] C. Wu, Y. Li, H. Yang, X. Wang, X. Li, and B. Jing, "The greater the interaction, the stronger the learning performance? Examining pedagogical agents' interactive presence in instructional videos," *Appl. Cogn. Psychol.*, vol. 38, no. 5, 2024, doi: 10.1002/acp.4242.

[53] P. Ayres and F. Paas, "Cognitive load theory: New directions and challenges," *Appl. Cogn. Psychol.*, vol. 26, no. 6, pp. 827–832, 2012, doi: 10.1002/acp.2882.

[54] R. Winkler and M. Söllner, "Unleashing the potential of chatbots in education: A state-of-the-art analysis," *Acad. Manag. Proc.*, vol. 2018, no. 1, p. 15903, 2018, doi: 10.5465/ambpp.2018.15903abstract.

[55] P. A. Kirschner, J. Sweller, and R. E. Clark, "Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching," *Educ. Psychol.*, vol. 41, no. 2, pp. 75–86, 2006, doi: 10.1207/s15326985ep4102_1.

[56] B. K. Mungai, P. K. K. Omieno, M. Egessa, and P. N. Manyara, "AI chatbots in LMS: A pedagogical review of cognitive, constructivist, and adaptive principles," *Eng. Technol. J.*, vol. 9, no. 8, 2024, doi: 10.47191/etj/v9i08.15.

[57] J. Sweller, P. Ayres, and S. Kalyuga, *Cognitive Load Theory*. Springer New York, 2011, doi: 10.1007/978-1-4419-8126-4.

[58] J. J. G. Van Merriënboer and J. Sweller, "Cognitive load theory and complex learning: Recent developments and future directions," *Educ. Psychol. Rev.*, vol. 17, no. 2, pp. 147–177, 2005, doi: 10.1007/s10648-005-3951-0.

[59] S. Gumilar, "Conceptual understanding and visual representation: Are they still ordinary or extraordinary issues?" *Res. Phys. Educ.*, vol. 2, no. 2, pp. 1–2, 2023, doi: 10.31980/ripe.v2i2.29.

[60] S. Mi, S. Lu, and H. Bi, "Trends and foundations in research on students' conceptual understanding in science education: A method based on the structural topic model," *J. Baltic Sci. Educ.*, vol. 19, no. 4, pp. 551–568, 2020, doi: 10.33225/JBSE/20.19.551.

[61] X. Huang, "Application of artificial intelligence app in quality evaluation of primary school science education," *Educ. Stud.*, vol. 50, no. 6, pp. 1215–1235, 2022, doi: 10.1080/03055698.2022.2066462.

[62] J. P. Adams and T. F. Slater, "Astronomy in the national science education standards," *J. Geosci. Educ.*, vol. 48, no. 1, pp. 39–45, 2000, doi: 10.5408/1089-9995-48.1.39.

[63] P. Hubber, "Year 8 students' understanding of astronomy as a representational issue: Insights from a classroom video study," 2009. [Online]. Available: http://dro.deakin.edu.au/eserv/DU:30021664/rentschler-year8students-2009.pdf

[64] P. M. Sadler, "Astronomy's conceptual hierarchy," in *Astronomy Education: Current Developments, Future Coordination* (ASP Conf. Ser., Vol. 89), J. R. Percy, Ed., p. 46, 1996.

[65] E. Gazit, D. Chen, and Y. Yair, "Developing understanding of basic astronomical concepts by using a virtual solar system," in *Proc. Int. Conf. Learn. Sci.*, p. 601, 2004. [Online]. Available: https://repository.isls.org/handle/1/4019

[66] M. A. Kuhail, N. Alturki, S. Alramlawi, and K. Alhejori, "Interacting with educational chatbots: A systematic review," *Educ. Inf. Technol.*, vol. 28, no. 1, pp. 973–1018, 2023, doi: 10.1007/s10639-022-11177-3.

[67] S. Wollny, J. Schneider, D. Di Mitri, J. Weidlich, M. Rittberger, and H. Drachsler, "Are we there yet? A systematic literature review on chatbots in education," *Front. Artif. Intell.*, vol. 4, p. 654924, 2021, doi: 10.3389/frai.2021.654924.

[68] P. Chandler and J. Sweller, "Cognitive load theory and the format of instruction," *Cogn. Instr.*, vol. 8, no. 4, pp. 293–332, 1991, doi: 10.1207/s1532690xci0804_2.

[69] N. D. Septiyanti, M. I. Luthfi, and D. Darmawansah, "Effect of chatbot-assisted learning on students' learning motivation and its pedagogical approaches," *Khazanah Informatika: J. Ilmu Komput. Informatika*, vol. 10, no. 1, pp. 69–77, 2024, doi: 10.23917/khif.v10i1.4246.

[70] H. Cooper, "Synthesis of research on homework," *Educ. Leadersh.*, vol. 47, no. 3, pp. 85–91, 1989. [Online]. Available: https://files.ascd.org/staticfiles/ascd/pdf/journals/ed_lead/el198911_cooper.pdf

[71] U. Trautwein, A. Niggli, I. Schnyder, and O. Lüdtke, "Between-teacher differences in homework assignments and the development of students' homework effort, homework emotions, and achievement," *J. Educ. Psychol.*, vol. 101, no. 1, pp. 176–189, 2009, doi: 10.1037/0022-0663.101.1.176.

[72] E. Güven Yıldırım and A. N. Köklükaya, "Determination of primary and middle school students' interest levels in science subjects," *Amasya Univ. J. Educ. Fac.*, vol. 5, no. 1, pp. 1–22, 2016. [Online]. Available: https://dergipark.org.tr/en/pub/amauefd/issue/24319/257699

[73] R. Mintz, S. Litvak, and Y. Yair, "3D-virtual reality in science education: An implication for astronomy teaching," *J. Comput. Math. Sci. Teach.*, vol. 20, no. 3, pp. 293–305, 2001. [Online]. Available: https://www.academia.edu/download/31690325/JCMST203293.pdf

[74] H. Cooper, J. C. Robinson, and E. A. Patall, "Does homework improve academic achievement? A synthesis of research, 1987–2003," *Rev. Educ. Res.*, vol. 76, no. 1, pp. 1–62, 2006, doi: 10.3102/00346543076001001.

[75] S. Dettmers, U. Trautwein, and O. Lüdtke, "The relationship between homework time and achievement is not universal: Evidence from multilevel analyses in 40 countries," *Sch. Eff. Sch. Improv.*, vol. 20, no. 4, pp. 375–405, 2009, doi: 10.1080/09243450902904601.

[76] J. Xu and H. Wu, "Self-regulation of homework behavior: Homework management at the secondary school level," *J. Educ. Res.*, vol. 106, no. 1, pp. 1–13, 2012, doi: 10.1080/00220671.2012.658457.

[77] J. Hattie and H. Timperley, "The power of feedback," *Rev. Educ. Res.*, vol. 77, no. 1, pp. 81–112, 2007, doi: 10.3102/003465430298487.

[78] D. C. Moos and A. Ringdal, "Self-regulated learning in the classroom: A literature review on the teacher's role," *Educ. Res. Int.*, vol. 2012, pp. 1–15, 2012, doi: 10.1155/2012/423284.

[79] K. A. Renninger and S. Hidi, "Interest development and learning," in *Cambridge University Press eBooks*, pp. 265–290, 2019, doi: 10.1017/9781316823279.013.

[80] A. Kuzmin, A. Dukhanov, and S. Kraev, "Approach to intelligent support in educational program design using X-matrices and a domain-specific language," in *Proc. 2022 Int. Russ. Autom. Conf. (RusAutoCon)*, pp. 920–925, 2024, doi: 10.1109/rusautocon61949.2024.10694424.

[81] Y. Lin and J. Ye, "Development of an educational chatbot system for enhancing students' biology learning performance," *J. Internet Technol.*, vol. 24, no. 2, pp. 275–281, 2023, doi: 10.53106/160792642023032402006.

[82] R. Chocarro, M. Cortiñas, and G. Marcos-Matás, "Teachers' attitudes towards chatbots in education: A technology acceptance model approach considering the effect of social language, bot proactiveness, and users' characteristics," *Educ. Stud.*, vol. 49, no. 2, pp. 295–313, 2021, doi: 10.1080/03055698.2020.1850426.

[83] A. S. Al Darayseh, "Acceptance of artificial intelligence in teaching science: Science teachers' perspective," *Comput. Educ.: Artif. Intell.*, vol. 4, p. 100132, 2023, doi: 10.1016/j.caeai.2023.100132.

[84] F. T. M. Ayasrah, K. Alarabi, and H. A. A. Fattah, "A secure technology environment and AI's effect on science teaching: Prospective science teachers," *Migration Lett.*, vol. 20, no. S2, pp. 289–302, 2023. [Online]. Available: https://migrationletters.com/index.php/ml/article/view/3687

[85] B. Kohli, T. Choudhury, S. Sharma, and P. Kumar, "A platform for human-chatbot interaction using Python," *IEEE Xplore*, 2018, doi: 10.1109/ICGCIoT.2018.8753031.

[86] O. A. Mehdi, H. Ibrahim, and L. S. Affendey, "Instance-based matching using regular expression," *Procedia Comput. Sci.*, vol. 10, pp. 688–695, 2012, doi: 10.1016/j.procs.2012.06.088.

[87] M. McTear, Z. Callejas, and D. Griol, *The Conversational Interface*. Springer, 2016, doi: 10.1007/978-3-319-32967-3.

[88] S. Toraman, "Mixed methods research: A brief history, definitions, perspectives, and key elements," *Qual. Soc. Sci.*, vol. 3, no. 1, pp. 1–29, 2021, doi: 10.47105/nsb.847688.

[89] L. Cohen, L. Manion, and K. Morrison, *Research Methods in Education*, 6th ed. London: Routledge, 2007.

[90] J. W. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 2nd ed. Thousand Oaks, CA: Sage, 2003.

[91] Ş. Büyüköztürk, *Handbook of Data Analysis for Social Sciences: Statistics and Research Design*. Ankara: Pegem Academy, 2018.

[92] J. M. Cortina, "What is coefficient alpha? An examination of theory and applications," *J. Appl. Psychol.*, vol. 78, no. 1, pp. 98–104, 1993, doi: 10.1037/0021-9010.78.1.98.

[93] J. C. Nunnally and I. H. Bernstein, *Psychometric Theory*, 3rd ed. New York: McGraw-Hill, 1994.

[94] K. S. Taber, "The use of Cronbach's alpha when developing and reporting research instruments in science education," *Res. Sci. Educ.*, vol. 48, no. 6, pp. 1273–1296, 2018, doi: 10.1007/s11165-016-9602-2.

[95] M. F. Turgut, *Measurement and Evaluation in Education*, 9th ed. Ankara: Saydam, 1992.

[96] C. Laçin Şimşek and H. Nuhoğlu, "Development of a valid and reliable interest scale towards science topics," *Sakarya Univ. J. Educ. Fac.*, no. 18, pp. 28–41, 2009. [Online]. Available: https://dergipark.org.tr/en/pub/sakaefd/issue/11214/133927

[97] E. Ş. Karalar, "Investigation of middle school students' interest in science subjects in terms of various variables," M.S. thesis, Muğla Sıtkı Koçman Univ., 2018.

[98] H. Kendirli, "The effect of Educational Informatics Network (EBA)-supported science applications on seventh-grade students' interest in science," M.S. thesis, Niğde Ömer Halisdemir Univ., 2017.

[99] F. Şimşek, "The effect of STEM activities on students' attitudes, interest, scientific process skills in science and student opinions," *Turk. J. Comput. Math. Educ.*, vol. 10, no. 3, pp. 654–679, 2019, doi: 10.16949/turkbilmat.470261.

[100] M. B. Miles and A. M. Huberman, *Qualitative Data Analysis: An Expanded Sourcebook*, 2nd ed. Sage Publications, 1994.

[101] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, p. 159, 1977, doi: 10.2307/2529310.

[102] K. Krippendorff, "On the reliability of unitizing continuous data," *Sociol. Methodol.*, pp. 47–76, 1995.

[103] J. Fraenkel, N. Wallen, and H. Hyun, *How to Design and Evaluate Research in Education*, 10th ed. New York: McGraw-Hill Education, 1993.

[104] J. Pallant, *SPSS Survival Manual: A Step-by-Step Guide to Data Analysis Using IBM SPSS*, 6th ed. New York: Open University Press, 2016.

[105] V. Sönmez and G. F. Alacapınar, *Illustrated Scientific Research Methods*. Ankara: Anı Publishing, 2019.

[106] A. Yıldırım and H. Şimşek, *Qualitative Research Methods in the Social Sciences*, 5th ed. Ankara: Seçkin, 2006.

[107] V. Clarke and V. Braun, "Thematic analysis," *J. Positive Psychol.*, vol. 12, no. 3, pp. 297–298, 2016, doi: 10.1080/17439760.2016.1262613.

[108] Ş. Büyüköztürk, *Handbook of Data Analysis for Social Sciences: Statistics, Research Design, SPSS Applications, and Interpretation*. Ankara: Pegem Academy, 2014.

[109] B. G. Tabachnick and L. S. Fidell, *Using Multivariate Statistics*, 7th ed. New York: Allyn & Bacon/Pearson Education, 2013.

[110] J. T. E. Richardson, "Eta squared and partial eta squared as measures of effect size in educational research," *Educ. Res. Rev.*, vol. 6, no. 2, pp. 135–147, 2011, doi: 10.1016/j.edurev.2010.12.001.

[111] J. Yin, T.-T. Goh, and Y. Hu, "Using a chatbot to provide formative feedback: A longitudinal study of intrinsic motivation, cognitive load, and learning performance," *IEEE Trans. Learn. Technol.*, vol. 17, pp. 1404–1415, 2024, doi: 10.1109/tlt.2024.3364015.

[112] Q. N. Nguyen, A. Sidorova, and R. Torres, "User interactions with chatbot interfaces vs. menu-based interfaces: An empirical study," *Comput. Human Behav.*, vol. 128, p. 107093, 2021, doi: 10.1016/j.chb.2021.107093.

[113] S. Park, "The effects of social cue principles on cognitive load, situational interest, motivation, and achievement in pedagogical agent multimedia learning," *Educ. Technol. Soc.*, vol. 18, no. 4, pp. 211–229, 2015.

[114] N. L. Schroeder, "The influence of a pedagogical agent on learners' cognitive load," *Educ. Technol. Soc.*, vol. 20, no. 4, pp. 138–147, 2017.

[115] A. L. Baylor, "Expanding pre-service teachers' metacognitive awareness of instructional planning through pedagogical agents," *Educ. Technol. Res. Dev.*, vol. 50, no. 2, pp. 5–22, 2002.

[116] Y. Wang and C. Lo, "The effects of response time on older and young adults' interaction experience with chatbot," *BMC Psychol.*, vol. 13, no. 1, 2025, doi: 10.1186/s40359-025-02459-9.

[117] B. Chen, G.-H. Hwang, and S.-H. Wang, "Gender differences in cognitive load when applying game-based learning with intelligent robots," *Educ. Technol. Soc.*, vol. 24, no. 3, pp. 102–115, 2021, doi: 10.30191/ETS.202107_24(3).0008.

[118] C. D. D. Bowman, "Student use of animated pedagogical agents in a middle school science inquiry program," *Br. J. Educ. Technol.*, vol. 43, no. 3, pp. 359–375, 2011, doi: 10.1111/j.1467-8535.2011.01198.x.

[119] S. Dinçer and A. Doğanay, "The effects of multiple-pedagogical agents on learners' academic success, motivation, and cognitive load," *Comput. Educ.*, vol. 111, pp. 74–100, 2017, doi: 10.1016/j.compedu.2017.04.005.

[120] M. A. S. N. Nunes, L. L. Dihl, L. M. Fraga, C. R. Woszezenki, L. Oliveira, D. J. Francisco, and M. Notargiacomo, "Animated pedagogical agent in the intelligent virtual teaching environment," *Interact. Educ. Multimedia*, vol. 4, pp. 53–60, 2002.

[121] F. S. Al-Hafdi and S. M. AlNajdi, "The effectiveness of using chatbot-based environment on learning process, students' performances and perceptions: A mixed exploratory study," *Educ. Inf. Technol.*, vol. 29, no. 15, pp. 20633–20664, 2024, doi: 10.1007/s10639-024-12671-6.

[122] S. Sakellariou, A. Molohidis, and E. Hatzikraniotis, "On the effectiveness of pedagogical agents," *Int. J. Teach. Learn. Sci. (IJTLS-102)*, 2024, doi: 10.47991/2024/IJTLS-102.

[123] R. E. Mayer and L. Fiorella, *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press, 2022.

[124] R. E. Mayer and R. Moreno, "Nine ways to reduce cognitive load in multimedia learning," *Educ. Psychol.*, vol. 38, no. 1, pp. 43–52, 2003, doi: 10.1207/S15326985EP3801_6.

[125] S. Ba, D. Stein, Q. Liu, T. Long, K. Xie, and L. Wu, "Examining the effects of a pedagogical agent with dual-channel emotional cues on learner emotions, cognitive load, and knowledge transfer performance," *J. Educ. Comput. Res.*, vol. 59, no. 6, pp. 1114–1134, 2021, doi: 10.1177/0735633121992421.

[126] Å. Brandt, M. P. Jensen, M. S. Søberg, S. D. Andersen, and T. Sund, "Information and communication technology-based assistive technology to compensate for impaired cognition in everyday life: A systematic review," *Disabil. Rehabil.: Assist. Technol.*, vol. 15, no. 7, pp. 810–824, 2020, doi: 10.1080/17483107.2020.1765032.

APPENDİX

Appendix A

Lesson Implementation Checklist

Number	Processes	Done	Not Done
1	Checking the classroom's physical environment (lighting, heating, arrangement) before lesson begins		
2	Checking the seating arrangement suitability for lesson delivery		
3	Ensuring classroom quietness by the teacher		
4	Conducting attendance and identifying absent students		
5	Establishing a free and open environment for student expression		
6	Conducting lessons in accordance with the planned curriculum		
7	Encouraging student participation during lessons		
8	Ensuring accessibility of lesson materials		
9	Providing equal opportunities for students to ask questions and participate		
10	Effectively using technological tools during the lesson		
11	Allocating time for students to summarize and reflect on learning at the end		
12	Allowing time for students to share and reflect their understandings		
13	Giving feedback opportunities related to learning processes		
14	Clearly communicating learning objectives at the beginning and end of lessons		

Appendix B

Semi-Structured Interview Form (SSIF) Regarding the Use of Educational Chatbots (Ecs)

1. Have you ever used chatbot technology before?

2. Did you encounter any problems while using the chatbot (BilgeCan Bot)? Please provide examples.

3. Did you feel comfortable while asking questions to the chatbot (BilgeCan Bot)? Please explain your thoughts.

4. What features of the chatbot (BilgeCan Bot) did you like? Please provide examples.

5. What features of the chatbot (BilgeCan Bot) did you dislike? Please provide examples.

6. What features would you like the chatbot (BilgeCan Bot) to have? Please explain and provide examples.

7. Did using the chatbot (BilgeCan Bot) help you learn the topics in the unit "The Sun, Earth, and Moon"? Please explain your thoughts.

8. Did using the chatbot (BilgeCan Bot) contribute to improving your questioning skills? Please explain your thoughts.

9. Did using the chatbot (BilgeCan Bot) contribute to improving your critical thinking skills? Please explain your thoughts.

10. Did using the chatbot (BilgeCan Bot) change your perspective towards the Science course? Please explain your thoughts.

11. Would you like to use the chatbot (BilgeCan Bot) to learn other units in the Science course? Please explain why, providing examples.

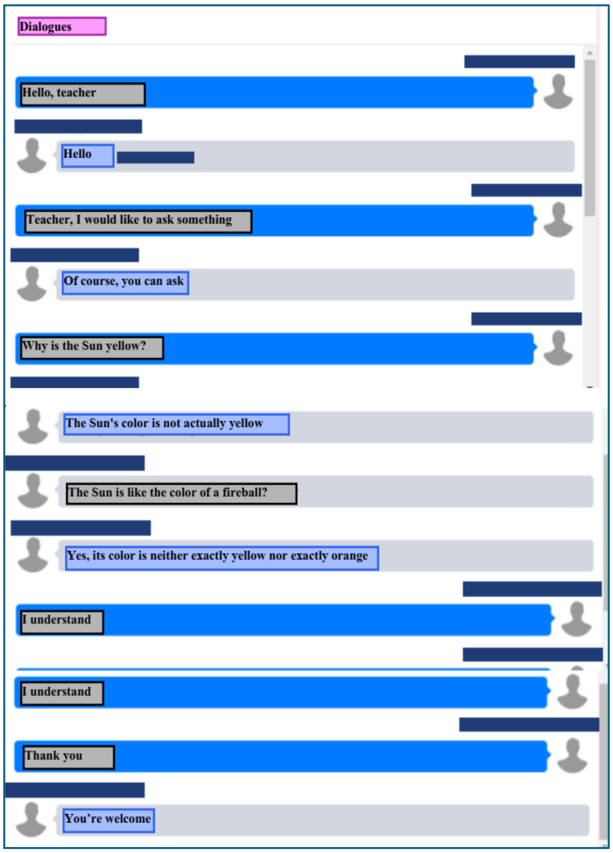
12. Would you like to use the chatbot (BilgeCan Bot) in other subjects outside the Science course? Please provide examples.

13. Do you think the chatbot (BilgeCan Bot) could be a companion for you when you are alone at home? Please explain your thoughts.

Appendix C

Sample Dialogue Between Students and the Live Teachers

Student S5–Teacher Dialogue



Dialogues

[redacted]

Hello, teacher

[redacted]

Hello [redacted]

[redacted]

Teacher, I would like to ask something

[redacted]

Of course, you can ask

[redacted]

Why is the Sun yellow?

[redacted]

The Sun's color is not actually yellow

[redacted]

The Sun is like the color of a fireball?

[redacted]

Yes, its color is neither exactly yellow nor exactly orange

[redacted]

I understand

[redacted]

I understand

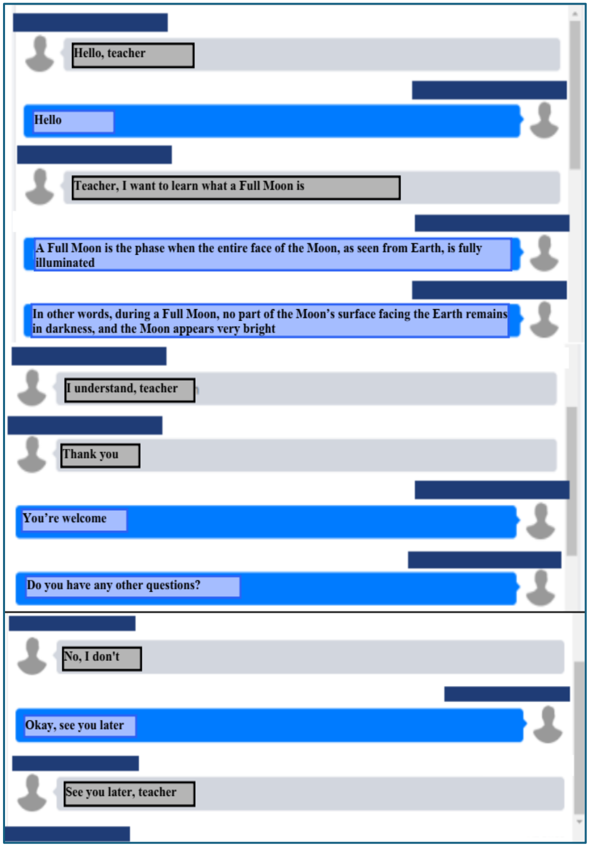
[redacted]

Thank you

[redacted]

You're welcome

Student S12–Teacher Dialogue



[redacted]

Hello, teacher

[redacted]

Hello

[redacted]

Teacher, I want to learn what a Full Moon is

[redacted]

A Full Moon is the phase when the entire face of the Moon, as seen from Earth, is fully illuminated

[redacted]

In other words, during a Full Moon, no part of the Moon's surface facing the Earth remains in darkness, and the Moon appears very bright

[redacted]

I understand, teacher

[redacted]

Thank you

[redacted]

You're welcome

[redacted]

Do you have any other questions?

[redacted]

No, I don't

[redacted]

Okay, see you later

[redacted]

See you later, teacher

[redacted]

Appendix D

Table 4. Thematic Findings of the First 6 Questions in the SSIF

Interview Question	Theme (Using chatbot to learn)	Category	Code	Frequency
1st	Experience History	Use First	No, this is my first time using it	26
2nd	Issues and Challenges	Technical Issues	Delayed Responses	4
			No problem	19
		Content	Missing Content	3
3rd	User Comfort	Comfort	Comfortable to Use	26
4th	Liked Features	Answer Quality	Accurate and Detailed Answers	7
		Speed and Instant Response	Quick Replies	11
		Ease of Use	Easy to use	26
		Active Teachers Menu	Like	5
		Instructive and Entertaining	Cute & Instructive	16
5th	Disliked Features	Answer Quality	Not Responding	3
		Active Teachers	Inability to connect with active teachers	8
		Limited Subject Areas	Restricted Areas of Information	12
6th	Desired Features	Broad Subject Areas	Use in Other Science Subjects	21
			Use in Different Courses	15
		Personalization	Personalized Responses	5
		Fun & Games	Useful Games	9
		Dialogue	Chat	16

1. Have you used chatbot technology before?"
2. Did you encounter any problems while using BilgeCan Bot?
3. Did you feel comfortable asking questions to BilgeCan Bot? Write down your thoughts
4. What features of BilgeCan Bot do you like? Give an example.
5. Which features of BilgeCan Bot did you not like? Give an example.
6. What other features would you like BilgeCanBot to have other than its current features? Give an example

Appendix E

Table 5. Thematic Findings of the Last 7 Questions in the SSIF

Interview Question	Theme (Benefits of Chatbot in Learning)	Category	Code	Frequency
7th	Benefits of Chatbot for Learning	Help to Learn	Provision of Useful Information	22
		Better Comprehension	Better Understanding of the Topic	18
		Motivation and Interest	Motivating and Engaging	13
8th	Development of Question Inquiry Skills	Contribution to Question Inquiry Skills	Contributed	17
			It contributed a little bit	7
			No, it never happened	2
9th	Development of Thinking Skills	Contribution to Thinking Skills	Contributed	23
			It wasn't too much	3
		Sense of Self-Confidence	Self-confidence	8
10th	Change in the Perspective of Science Courses	Interest in Science Course	Interest in Science Courses Is Increasing	18
		Improvement in Learning and Comprehension	Progress in Understanding Science Subjects	14
		Positive perception of the course	Enjoying Science Class	16
11th	Willingness to Learn Other Topics with Chatbot	Positive Attention	Interest and Willingness to Learn Other Units	26
		Fun and Intuitive Learning	Ease and Pleasure of Learning with Chatbot,	16
		Independence from Time and Space	Flexibility of Learning with Chatbot	13
12th	Willingness to Use the Chatbot in Other Courses	Getting General Information	To have knowledge in different courses	26
		Help with Exams and Assignments	Get Help with Exams and Assignments	8
		Practical and Fast Information Access	Fast Learning in Other Lessons	11
		Learning Without Teacher Assistance	Obtaining knowledge without the help of a teacher	16
13th	Making Friends with a Chatbot	Structure and Capabilities of the Chatbot	The Limits of the Chatbot	6
		Human-Chatbot Relationship	Chatbot's ability to make friends	14
		Ability to establish emotional connection	The Chatbot's Lack of Emotion	11

7. Was using BilgeCan Bot helpful in learning about "Sun, Earth and Moon"? Write down your thoughts.
8. Has using BilgeCan Bot helped you improve your questioning skills? Write down your thoughts.
9. Did using BilgeCan Bot help you improve your thinking skills? Write down your thoughts.
10. Did using BilgeCan Bot change your perspective on science class? Write down your thoughts.
11. Would you like to use BilgeCan Bot to learn other subjects in science class? Give an example and explain why.
12. Would you like to use BilgeCanBot in other courses besides science? "Give me an example."
13. Can using BilgeCan Bot accompany you when you are alone at home? Write down your thoughts.

Appendix F

Table 6. Token and TF-IDF Values of Weekly Chats

1st week		2nd week		3rd week		4th week		5th week		6th week	
Token	TFxIDF	Token	TFxIDF	Token	TFxIDF	Token	TFxIDF	Token	TFxIDF	Token	TFxIDF
sun	0,853	sun	0,823	moon	0,589	moon	0,618	moon	0,682	moon	0,408
sun	0,265	sun	0,206	month	0,295	month	0,507	month	0,364	month	0,306
keeping	0,235	sometimes	0,171	surface	0,236	appears	0,265	keeping	0,341	sometimes	0,306
time	0,147	look	0,171	world	0,177	sometimes	0,221	sometimes	0,25	world	0,272
appears	0,118	stains	0,171	moon	0,177	time	0,221	sun	0,227	sun	0,238
down	0,088	same	0,137	craters	0,177	same	0,176	same	0,136	same	0,238
night	0,088	bright	0,137	occurred	0,177	phases	0,132	appears	0,136	time	0,204
spots	0,088	a lot	0,103	around	0,177	around	0,088	time	0,136	around	0,204
same	0,088	time	0,103	rotating	0,177	occurred	0,088	day	0,114	day	0,17
light	0,088	size	0,103	month	0,177	month	0,088	phases	0,091	return	0,17

Appendix G

Table 7. Thematic Findings of the Chat Recordings Between Students and Bilgecan Bot

Week	Question level	Number of questions	Sample questions
1st	Knowledge	11	What are the layers of the Sun?
	Understand	11	How does a solar eclipse happen?
	Application	11	When does a solar eclipse happen and how can I watch it?
	Analysis	8	Why doesn't the sun always look like it's in the same place?
2nd	Knowledge	9	Why does the sun look yellow?
	Understand	8	Why should I wear special glasses while watching a solar eclipse?
	Application	8	How can we see sunspots?
	Analysis	6	Why are Galileo's observations about sunspots important?
3rd	Knowledge	10	Is there an atmosphere on the Moon?
	Understand	8	How does the Moon revolve around the Earth?
	Application	8	How did craters form?
	Analysis	7	Why does the moon sometimes appear completely round at night, sometimes in a half or crescent shape?
4th	Knowledge	14	How long does it take for the Moon to revolve around an Earth?
	Understand	13	How can the moon be so bright at night?
	Application	15	How to distinguish the phases of the Moon?
	Analysis	13	Does the moon always stay in the same place?
5th	Knowledge	12	How does a solar eclipse happen?
	Understand	14	Why does the shape of the Moon keep changing?
	Application	12	How does our rotation around the Sun affect the seasons of the year?
	Analysis	14	Why does the moon sometimes appear big and sometimes small?
6th	Knowledge	15	Do the other planets in the solar system revolve around themselves like Earth?
	Understand	15	Why is the Moon in different phases?
	Application	12	Why does the moon sometimes appear during the day as well?
	Analysis	12	Do other planets in the solar system also have moons like the Moon?

Appendix H

Table 8: Thematic Findings of Control Group Students' Feedbacks

Theme	Category	Code	Explanation
Visualization and Modeling Challenges	Visual Comprehension Difficulties	Geometric shape and rotational motion of the Sun	Difficulty imagining and understanding the shape and motion of the Sun
		Surface structure and atmosphere of the Moon	The complexity of understanding the Moon's surface and atmosphere, the difficulty of expressing thoughts on the possibility of life
		Models of movement of celestial bodies	Uncertainty in modeling the relative motions between the Sun, Earth and Moon
	Modeling and Implementation Challenges	Size comparison of the Sun and Earth	Inability to model the difference in magnitude between the Sun and the Earth
		Rotation of the Moon and its phases	Incomprehension of the relationship between the Moon's rotation around the Earth and its phases
Lack of Interaction and Support	Individual Understanding and Doubts	Lack of comprehension security	Insecurity about fully comprehending the material being read
		The need for interactive support	The need to ask questions and get feedback

El Cezeri
Journal of Science and Engineering
ISSN: 2148-3736

Research Article

Wind Power Prediction Based on Polynomial Regression Method

Hussein H. Saleh[1a]
[1]Ministry of education, Iraq

hunzgz@yahoo.com

DOI : 10.31202/ecjse.1722153
Received: 18.06.2025 Accepted: 19.09.2025
How to cite this article:
Hussein H. Saleh, “Wind Power Prediction Based on Polynomial Regression Method.”, El-Cezeri Journal of Science and Engineering, Vol: 12, Iss: 3, (2025), pp.(274-282).
ORCID: [a]0000-0003-1305-7391.

Abstract Accurate prediction of wind energy output from site-specific wind velocity data is essential for evaluating the feasibility and energy yield of wind farm installations. This study presents a predictive framework for wind power assessment based on daily wind velocity datasets from eight Iraqi cities: Duhok, Mosul, Kirkuk, Baghdad, Najaf, Wasit, Qadisiyyah, and Basra. The proposed method employs polynomial regression (POR) as a nonlinear estimation technique to correlate daily wind speed data with turbine power output. POR is selected for its computational simplicity and adequacy in capturing the nonlinearity inherent in wind power curves. The predictive models are calibrated individually for each city using historical wind data and the manufacturer's specified power curve for the turbine. The results indicate that the POR-based predictions closely align with the actual turbine power curve across all sites, demonstrating low prediction error and strong curve-fitting behavior. Notably, Wasit, Qadisiyyah, and Basra exhibited the highest potential for wind energy generation, with annual predicted outputs exceeding 2000 kWh per turbine, indicating promising conditions for small-scale wind energy exploitation. In contrast, northern cities such as Duhok and Mosul yielded significantly lower outputs (<1000 kWh annually), suggesting limited economic viability under the studied configuration.

Keywords: Artificial intelligence, Machine learning, Wind power, Wind velocity.

I. INTRODUCTION

The slowly growing energy field and increasing demand for electric power are still unreliable and unstable in Iraq. Moreover, Iraqi households receive a daily average of less than twenty-four hours of electricity from the national grid [1]. Iraq is heavily dependent on petroleum products and fossil oil, as well as imported natural gas, to generate energy [2]. In addition, since June 2017, Iraq has lost 4 billion dollars annually due to importing natural gas to support forty-six power plants [3]. In most countries, the development and utilization of new power sources as renewable energy are used to reduce dependence on traditional energy sources, diversify energy sources, and reduce CO2 emissions. Among all forms of renewable energy, wind energy is a sustainable and pollution-free source, environmentally friendly, and the fastest growing in many regions around the world [4]. Nowadays, the use of wind farms is increasing in many countries as a low-cost power system, providing energy without operational fuel, and a significant alternative to traditional power plants that require fuel [5]. The accurate assessment of wind power potential is required before planning to implement a wind farm. Many researchers use several methods to predict wind ability, such as data decomposition, empirical models, and machine learning methods.

A novel wind energy forecasting system to enhance forecasting capability by refining the raw time series using horizontal denoising and vertical granulation methods, The predictive performance was evaluated based on wind energy, speed, and direction data in China and Turkey, the proposed system was contribute to the advancement of wind-power prediction theory with substantial potential for practical applications [6]. Three machine learning methods Bayesian optimization (BO) to optimally tune hyperparameters of the Gaussian process regression (GPR), Support Vector Regression (SVR) with different kernels, and ensemble learning (ES) models have been used to employ and investigate their forecasting performance in wind power generation for three wind turbines in France, Turkey, and Kaggle. The results reveal the benefit of considering lagged data and input variables to better forecast wind power [7]. A new hybrid wind speed prediction model is designed based on the SSA (Singular Spectrum Analysis), EMD (Empirical Mode Decomposition), and CNNSVM (Convolutional Support Vector Machine) to reduce the noise and extract the trend information of the original wind speed data, extract the fluctuation features of the wind speed data, decompose the wind speed time series into several sub-layers and predict each of the wind speed sub-layers to investigate the prediction performance of the proposed model. According to the prediction results, the established model had significantly better performance than the seven comparison models from 1-step to 3-step wind speed predictions [8]. A novel wind speed forecasting framework for stationary data based on a hybrid decomposition method and deep learning models to achieve high estimating accuracy for the Dumat Al-Jandal wind farm in Al-Jouf, Saudi Arabia. According to statistical indicators, the results demonstrated the superiority of the proposed model in single and multiple wind prediction. The average means absolute error, root mean square error, and R-squared error [9]. In Spain, a triple-stage multi-step wind power forecasting approach is proposed by applying attention-based deep residual gated recurrent unit (GRU), network combined with ensemble empirical mode decomposition (EEMD), and crisscross optimization algorithm (CSO), the results demonstrate that the proposed hybrid model has significant advantage over other state-of-the-art models involved in study [10]. Other techniques, such as the Long Short-Term Memory (LSTM) network, were proposed to fit the nonlinearity between data variables and wind power of a wind farm in Hunan, China, during the four seasons of 2020. This method proved the accuracy of wind power prediction and the effectiveness for power dispatching [11]. Several regression methods, namely, Deep neural network (DNN), k-nearest neighbor (KNN) regressor, long short-term memory (LSTM), averaging model, random forest (RF) regressor, bagging regressor, and gradient boosting (GB) regressor are constructed to achieve the prediction of wind power generation, using dataset includes 4 features and 50530 instances. The experimental results illustrated that the proposed optimization of LSTM achieved the best results in forecasting the wind power values [12]. The prior review literature demonstrates the reliability and efficiency of machine learning techniques in modeling wind power. Furthermore, these techniques are complex, less interpretable, require large labeled datasets, and more computational resources, especially for deep networks [13]. As a result, the author suggests simplifying and obtaining higher performance prediction via a polynomial technique for regression. This methodology is easily interpretable, suitable for simpler non-linear data, and requires fewer computational resources. This study provides new insights into using turbine model parameters, presenting a novel wind power prediction model. In contrast to traditional models, the proposed model has reduced implementation complexity and computational burden, as well as ease of real-time implementation.

In Iraq, the number of wind farms for energy generation is almost non-existent. Moreover, the amount of research on wind energy forecasting is limited in circumstances of acute energy deficiency and increasing demand.

To avoid the lack of energy problem, obtain the maximum benefit of renewable energy sources, and understand the ability and reliability of wind power, this study designs and develops a wind power forecasting model based on wind turbine output parameters as the learning stage, followed by the POR method aiming to predict wind energy in eight different locations in Iraq.

II. LOCATION DESCRIPTION AND DATA COLLECTION

Iraq has an area of 437,072 square kilometers, geographically located in the western part of the Asia continent and enjoys considerable climatic diversity [14]. Annually, the highest values of wind velocity are recorded in June and July, while the lowest values are in December. Also, the annual averages of wind velocity were divided into three regions: northern (2.7 m/s), middle (3.6 m/s), and southern (4.1 m/s) [15].

The daily wind velocity recorded at 10 Meters in height for the 2023 season in eight cities, as shown in Figure 1, as well as geographical locations presented in Table 1, are obtained from the Iraqi Meteorological Organization and Seismology (IMOS) [16]. In addition, the performance data of the (GhrepowerFD21-50_61.2kW_21.5) turbine model that determines the output parameters of the turbine at a given wind velocity has been provided from the National Renewable Energy Laboratory (NREL). The parameters of this turbine are presented in Table 2 [17]. The power generated by any wind turbine is dependent on the velocity of the wind hitting it and the characteristics of the turbine model. The parameters associated with the POR method are wind velocity values and their corresponding turbine output power, as shown in Figure 2. The minimum wind velocity required to obtain output power for the suggested turbine is 2.9 meters per second (m/s), which is most appropriate with the natural wind velocity of the study areas. All datasets are available at “https://github.com/Dataset-comp-j”.

Table 1. Geographical Location of the Cities Considered.

City	Latitude (°N)	Longitude (°E)
Duhok	36.8632	42.9885
Mosul	36.3489	43.1577
Kirkuk	35.4666	44.3799
Baghdad	33.3152	44.3661
Najaf	32.0107	44.3265
Wasit	32.5165	45.8160
Qadisiyyah	32.0437	45.1495
Basrah	30.5258	47.7738

Figure 1. Location of the cities on the map of Iraq.

III. METHODOLOGY

The present study aims to extract the total annual values of wind energy using output parameters of specific turbine models and assess the performance of the POR method in predicting the annual amounts of wind power in eight cities in Iraq. To achieve these goals, and to establish the model for every location, two parameters and 16 instances of the (GhrepowerFD21-50_61.2kW_21.5) turbine model have been used as input data to produce the learning stage of the prediction model, and 2904 instances of the daily wind velocity dataset are used to estimate the approximate daily power generated at each location. The parameters of the suggested turbine consist of the output of power levels in the range [1 - 60] kWh at a given wind velocity, as shown in Figure 2. These two parameters have been used as the first inputs to the POR technique to perform the learning process (building a learning model). The second input parameter is the daily wind velocity of one season for a specific location. Both first and second input parameters are basic materials to generate models and obtain the desired output value, as shown in Figure 3.

One of the limitations of the POR method is the risk of overfitting, when the dataset contains noise in the input that leads to poor performance and low accuracy. Therefore, in this study, a few instances of the turbine parameters have negative values. The wind velocity dataset contains values that are out-of-range of the turbine wind parameter. In current research, both the negative values and out of range instances have been removed to minimize prediction errors and produce efficient models.

All values of the wind velocity that are lower than the minimum operating values of the proposed turbine have been removed. The percentage of removed values was (24%, 55%, 35%, 39%, 31%, 24%, 26%, and 20%) in Duhok, Mosul, Kirkuk, Baghdad, Najaf, Wasit, Qadisiyyah, and Basra, respectively.

A. POLYNOMIAL REGRESSION

Polynomial regression is a supervised parametric machine learning algorithm that deal with regression problems [18]. It is a form of linear regression method used when there is a nonlinear relationship between the dependent and independent variables, and it can analyze the relationships between variables by fitting higher degree polynomial functions. It is also considered to be a special case of multiple linear regression [19]. The formulation structure consisted of a single equation with the simplest interpretation of learning parameters that is applicable using fewer computational resources. In predictive methods, the model consists of a random output variable y ̂and a set of random input variables *x={x_1…,x_k}*. The following formula represents the polynomial regression fitting between dependent and independent variables:

$$\hat{y} = T_0 + T_1x + T_2x^2 + \dots + T_kx^k \tag{1}$$

The $\hat{y}$ is the output variable, x is the input variable, T_0 is the intercept, $T_1, T_{2,} \dots, T_k$ are polynomial regression coefficients, and k is the degree of the polynomial [20]. The characteristic of the polynomial regression method is that it uses one input variable. Also, the formula has the power of the input variable x, which is different from the linear regression methods, where the simpler first-order model (a "straight line") is adequate in describing the trend in the data. The model has tested the null hypothesis:

$$H_0: T_2 = T_3 = 0 \tag{2}$$

If a polynomial term of a given order is retained, then *all related lower-order terms are also retained*. That is, if a quadratic term (x^2) is deemed significant, then it is standard practice to use this regression function:

$$\mu y = T_0 + T_1 x + T_2 x^2 \tag{3}$$

Whether or not the linear term x is significant, the polynomial model is represented hierarchically [21]. One of the procedures for building a polynomial model is selecting the degree of the polynomial, and the highest power of the variable in the polynomial expression is crucial for various applications. In this work, the degree of the polynomial used is 25, which has been selected based on the visual match between the blue curve of turbine parameters and the red scatter points.

Table 2. The Parameters of the (GhrepowerFD21-50_61.2kW_21.5) Turbine Model.

Item	Value	Units
Rated Power	50	kW
Rated Wind velocity	11.3	m/s
Cut-in Wind velocity	4	m/s
Cut-out Wind velocity	25	m/s
Rotor Diameter	14.9	m
Hub Height	31.1	m

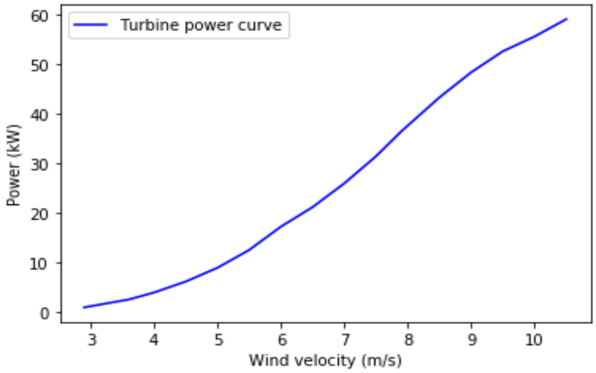


Figure 2. The power curve output of the (GhrepowerFD21-50_61.2kW_21.5) turbine model at a given wind velocity.

IV. RESULT AND DISCUSSION

The POR method was applied as the proposed approach to predict wind energy at eight locations in Iraq using the parameters of the (GhrepowerFD21-50_61.2kW_21.5) turbine model and wind velocity dataset of the 2023 season. To evaluate and prove the performance of the POR models, the parameters of turbine models were plotted against the prediction wind power, Figure 4 illustrate the turbine parameters curve and scatter points of the prediction wind power at all locations, the perfect convergence observed between turbine model curve and scatter points indicate to the successful and high accuracy prediction, as seen from Figure 4 the POR models perform well in predicting wind energy and efficient in analysis the parameters of proposed turbine model. Figure 5 shows the monthly means of wind velocity of the study locations, in the north of the country the wind recorded lowest means specifically in Duhok and Musol, whereas the middle cities Baghdad and Najaf have slightly higher means compared with the northern locations, while the best monthly means were in Wasit, Qadisiyyah, and Basra, especially in June and July with velocity exceeded 5 m/s. Table 3 explains the power annual amounts that can be obtained from a single turbine. The total power was less than 1000 kWh in Duhok and Mosul, while exceeding 1000 kWh in Kirkuk, Baghdad, and Najaf. In other locations, total annual energy was recorded in Wasit and Qadisiyyah, which were limited to 2000 kWh for each turbine. The greatest amounts of wind power are registered in Basra city, which exceeded 3000 kWh annually. The wind energy quantities at the study locations present an incentive to encourage investment in the wind energy field and exploit the wind element to support the energy sector. Figure 6 presents the box plots of the wind power density of all proposed locations. It illustrates the annual distribution of predicted wind power for each model achieved based on the POR technique. The Basra location generally recorded high amounts of wind power density distribution exceeding 20 kWh and maximum of 45.8 kWh, as explained at the beginning and end of the upper line extending from the box, whereas the wind power density distribution values in Najaf, Wasit and Qadisiyyah were less than 10 kWh as well as the fluctuation of the higher values. The remaining locations were worst in term of wind power density distribution due to poor wind velocity in these cities at 10 meters. The single points that fall outside the whiskers on the diagram show the outliers of each location. According to the above outcomes, the study locations contain a wealth of wind energy, which could be one of the basic pillars of support for the energy field in Iraq. In addition, these results proved the effectiveness of the (GhrepowerFD21-50_61.2kW_21.5) turbine model in extracting the best amounts from energy, which makes it recommended by the author.

V. CONCLUSION

To understand the ability of wind power, support the national electricity grid, and add a new power source, this study has designed and developed wind power prediction models based on the polynomial regression method for forecasting wind power amounts in eight cities using (GhrepowerFD21-50_61.2kW_21.5) turbine model parameters as training examples for the models.

Table 3. Total Annual Amounts of the Wind Power for a Single Turbine.

Location	Total wind power (kWh)
Duhok	974.7
Mosul	625.7
kirkuk	1224.3
Baghdad	1036.3
Najaf	1548
Wasit	2926.6
Qadisiyyah	2104.7
Basra	3595.2

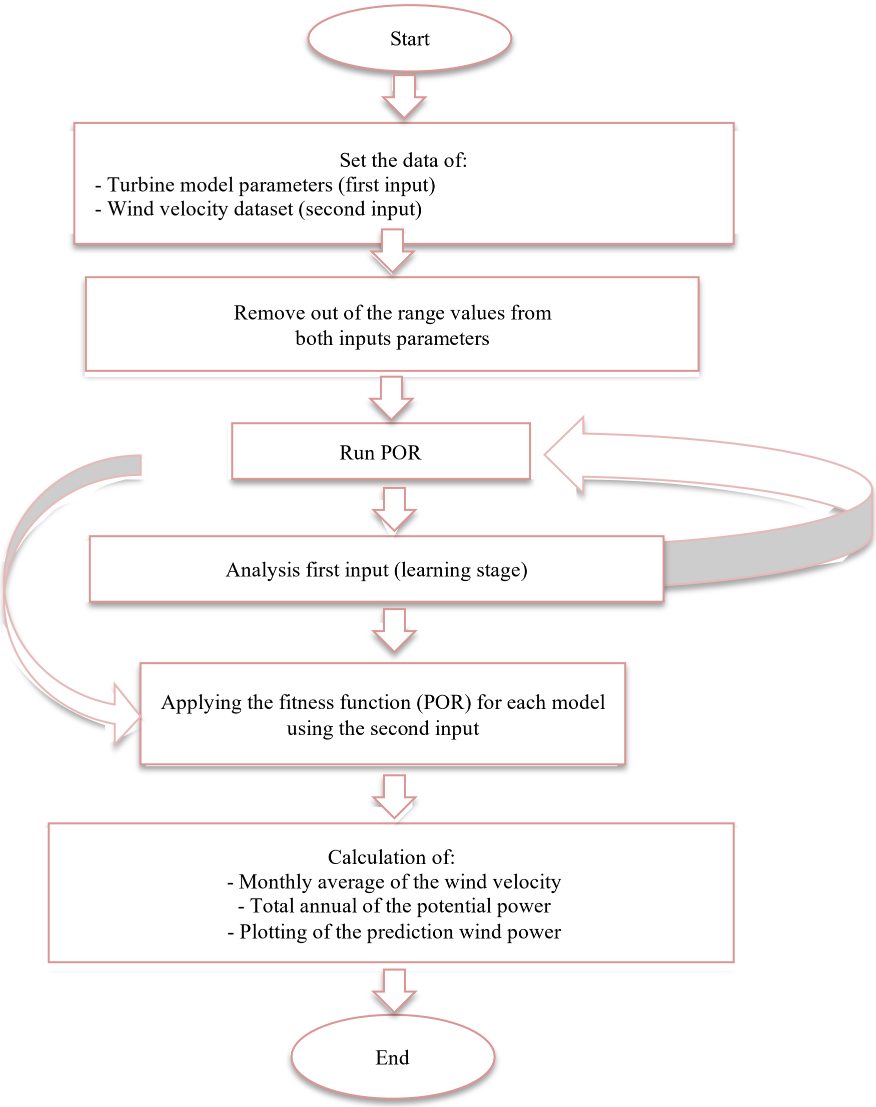


Figure 3. Flowchart of the core structure of the study [22].

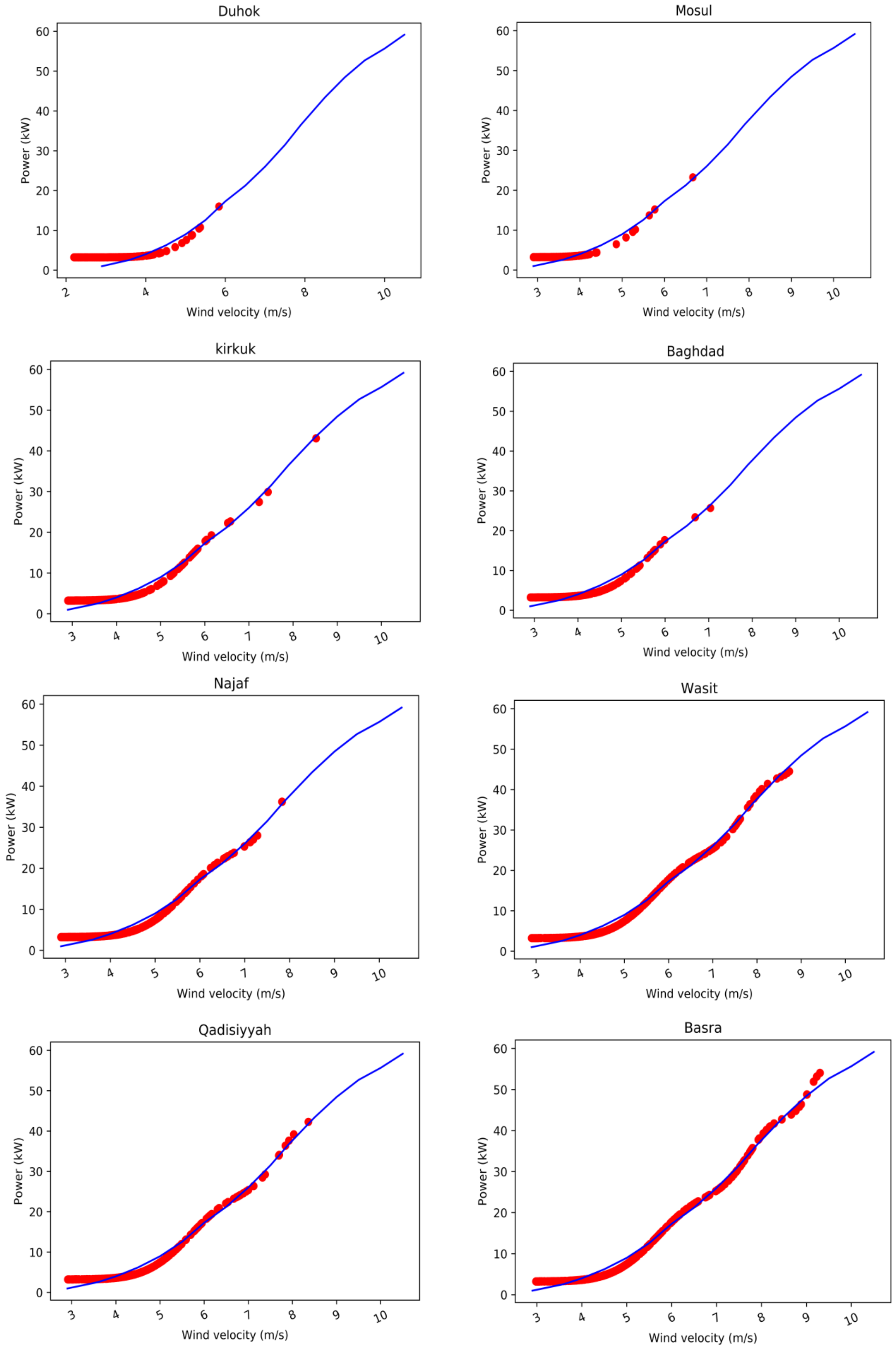


Figure 4. Turbine parameters curve and scatter points of the predicted wind power at all locations.

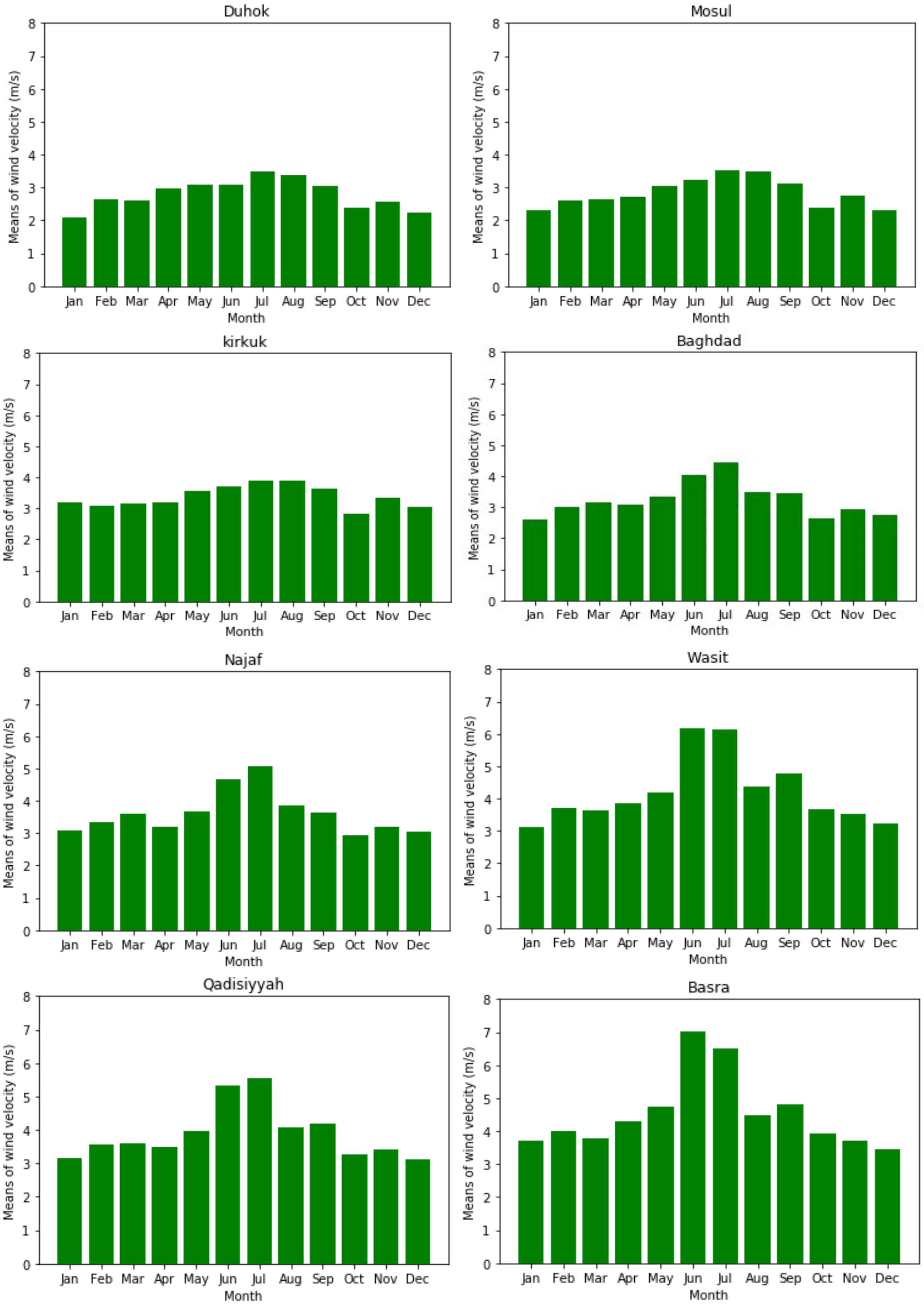


Figure 5. Monthly means of wind velocity of the study locations.

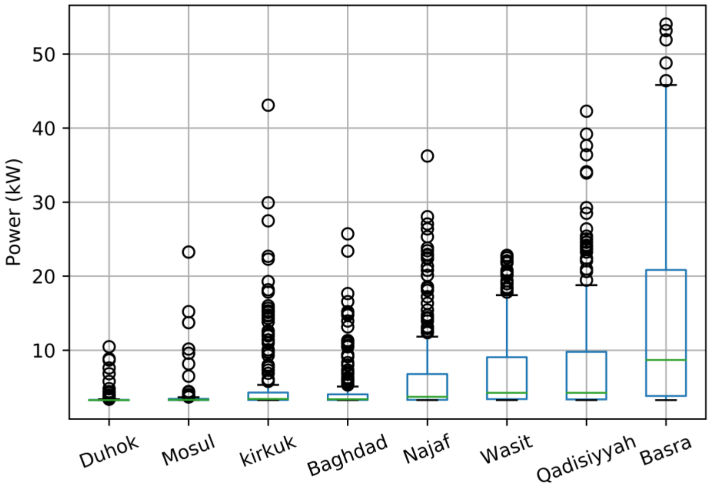


Figure 6. Box plots of wind power density at all proposed locations.

Annually, the monthly means of wind velocity at the southern locations recorded high velocities exceeding 5 m/s, especially in June and July, while the means of wind velocity began to gradually decrease in the middle and northern regions during the season. The results of this study demonstrate variations in the values of wind energy in the north, middle, and south locations. The northern locations have modest quantities of wind power due to the scarcity of wind velocity at 10 meters. However, the best wind velocity averages were in the middle and southern regions, where high quantities of wind power are available. Also, Qadisiyyah and Basra were recognized as powerful producers of wind energy. As a result, Figure 4 proved the accuracy and superior capability of POR as a less complex technique in extracting wind energy potential at different locations and gave better performance. Furthermore, the results of this work encourage the self-sufficient energy, saving billions of dollars spent on importing energy or fuel to operate conventional power plants and investing in clean energy. In addition to developing the energy field. The study confirms the viability of polynomial regression for wind energy estimation when turbine specific performance data is available. However, for future work, the inclusion of additional meteorological variables (e.g., air density variation, turbulence intensity, and diurnal wind patterns) and comparison with more advanced machine learning models (e.g., SVR, ANN) could improve generalizability and robustness. Furthermore, this element could be useful in wind power forecasting and understanding the wind density of each location.

References

[1] Daily Loads. Retrieved March 10, 2023, from Ministry of Electricity: http://www.moelc.gov.iq

[2] H.D.S. Altai, T. A. Faisal, H. Lazim Mohammed, and TH Haider A. Salim, “Analysis of the problems of electricity in Iraq and recommendations of methods of overcoming them, ” *Periodicals of Engineering and Natural Sciences (PEN) 10*, no. 1, 607-614, 2022.

[3] A. Bacci, (2017). Iraq Petroleum 2018—Natural Gas Must Be an Asset for Iraq.

[4] K. Mohammadi, A. Omid and J. G. McGowan, “Use of Birnbaum-Saunders distribution for estimating wind speed and wind power probability distributions: A review,” *Energy Conversion and Management,* 143, 109-122, 2017.

[5] S. Dongran, Y. Yang, S. Zheng, X. Deng, J. Yang, M. Su, WçTang, Xç Yang, L. Huang, and Y. H. Joo. “New perspectives on maximum wind energy extraction of variable-speed wind turbines using previewed wind speeds,” *Energy conversion and management*, 206, 112496, 2020

[6] J. Wang, , Y. Qian, Z. Linyue, W. Kang, and H. Zhang, “A novel wind power forecasting system integrating time series refining, nonlinear multi-objective optimized deep learning and linear error correction,” *Energy Conversion and Management,* 299, 117818, 2024. https://doi.org/10.1016/j.enconman.2023.117818

[7] A. Abdulelah, F. Harrou, and Y Sun, “Efficient wind power prediction using machine learning methods, A comparative study,” *Energies* 15, no. 7, 2327, 2022. https://doi.org/10.3390/en15072327

[8] M. Xiwei, H. Liu, and Y. Li, “Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine,” *Energy conversion and management* 180, 196-205, 2019.

[9] A. Yasmeen, N. Hewahi, and A. Alasaadi, “Wind Speed Forecasting Based on Data Decomposition and Deep Learning Models, A Case Study of a Wind Farm in Saudi Arabia,” *arXiv preprint arXiv*,2412.13356, 2024.

[10] M. Anbo, S. Chen, Z. Ou, W. Ding, H. Zhou, J. Fan, and H. Yin, “A hybrid deep learning architecture for wind power prediction based on bi-attention mechanism and crisscross optimization,” *Energy* 238, 121795, 2022. https://doi.org/10.1016/j.energy.2021.121795

[11] W. Jianing, H. Zhu, Y. Zhang, F. Cheng, and C. Zhou, "A novel prediction model for wind power based on improved long short-term memory neural network," Energy 265, 126283, 2023. https://doi.org/10.1016/j.energy.2022.126283

[12] T. Zahraa, M. Y. Shams, A. M. Elshewey, El-Sayed M. El-kenawy, A. Ibrahim, Abdelaziz A. Abdelhamid, and Mohamed A. El-dosuky, "Wind Power Prediction Based on Machine Learning and Deep Learning Models," *Computers, Materials and Continua,* 75, no. 1, 2023.

[13] V. Cyril, G. Notton, S. Kalogirou, M.-L. Nivet, C. Paoli, F. Motte, and A. Fouilloy, "Machine learning methods for solar radiation forecasting, A review," Renewable energy 105, 569-582, 2017. https://doi.org/10.1016/j.renene.2016.12.095

[14] Al-A. Nadhir, "Topography and climate of Iraq," *Journal of Earth Sciences and Geotechnical Engineering* 11, no. 2, 1-13, 2021.

[15] A. H. Q., and Y. K. Al-Timimi, "GIS techniques for mapping of wind speed over Iraq," *Iraqi Journal of Agricultural Sciences* 50, no. 6, 2019. https://doi.org/10.36103/ijas.v50i6.852

[16] Iraqi Meteorological Organization and Seismology (IMOS). http://www.meteoseism.gov.iq

[17] SWCC Small Wind Certification Council, "ICC-SWCC Summary Report SWCC-15-01-P," Brea, CA,Small Wind Certification Council, 2014. http://smallwindcertification.org/wpcontent/uploads/2019/09/Summary-Report-SWCC-15-01-2019-P.pdf.

[18] O. Eva, "Modelling using polynomial regression," *Procedia engineering,* 48, 500-506, 2012. https://doi.org/10.1016/j.proeng.2012.09.545

[19] S. Maolin, W. Hu, M. Li, J. Zhang, X. Song, and W. Sun, "Ensemble regression based on polynomial regression-based decision tree and its application in the in-situ data of tunnel boring machine," *Mechanical Systems and Signal Processing,* 188, 110022, 2023. https://doi.org/10.1016/j.ymssp.2022.110022

[20] I. Hamza, N. Moneem Al-Abdaly, Mohammed Hammodi Shamsa, A. Shatnawi, Majed Ibrahim, and K. Adam Ostrowski, "Development of prediction model to predict the compressive strength of eco-friendly concrete using multivariate polynomial regression combined with stepwise method,"*Materials* 15, no. 1, 317, 2022. https://doi.org/10.3390/ma15010317

[21] A. Ali, H. Alahmer, A. Handam, and H. Rezk, "Environmental assessment of a diesel engine fueled with various biodiesel blends, Polynomial regression and grey wolf optimization," *Sustainability,* 14, no. 3, 1367, 2022. https://doi.org/10.3390/su14031367

[22] S. Hussein H, "Estimating the amounts of the solar energy, A multiple linear regression approach," *In AIP Conference Proceedings*, vol. 2591, no. 1. AIP Publishing, 2023. https://doi.org/10.1063/5.0125118

El Cezeri
Journal of Science and Engineering
ISSN: 2148-3736

Research Article

A Comparative Analysis of Elastic, Mechanical, and Thermoelectric Properties in Mg_2X (X =Si, Ge and Sn) Semiconducting Nanomaterials

Adwitiya Yadav[1a], Prashant Srivastav[1b], Pramod Kumar Yadawa[1c]
[1]Department of Physics, Prof. Rajendra Singh (Rajju Bhaiya) Institute of Physical Sciences for Study and Research, V. B. S. Purvanchal University, Jaunpur, India

adwitiya94@gmail.com

DOI: 10.31202/ecjse.1695273
Received: 08.05.2025 Accepted: 19.09.2025
How to cite this article:
Adwitiya Yadav, Prashant Srivastav, Pramod Kumar Yadawa, "A Comparative Analysis of Elastic, Mechanical, and Thermoelectric Properties in Mg_2X (X =Si, Ge and Sn) Semiconducting Nanomaterials", El-Cezeri Journal of Science and Engineering, Vol: 12, Iss: 3, (2025), pp.(283-297).
ORCID: [a]0009-0009-0978-6111, [b]0000-0003-4596-6258, [c]0000-0002-9525-2205.

ABSTRACT: The elastic, mechanical, thermodynamic, and ultrasonic properties of Mg_2X (X = Si, Ge, Sn) composite semiconducting nanomaterials were systematically investigated using Lennard-Jones potential analysis. This approach enabled the evaluation of second- and third-order elastic coefficients. Our findings, corroborated through comparative analysis with existing literature, reveal that higher-order elastic constants exhibit a slight initial increase, while subsequent values show minimal variation from Mg_2Si to Mg_2Sn. The elastic constants were further employed to assess the mechanical behavior of hexagonal Mg_2X nanomaterials via key parameters including Poisson's ratio, Young's modulus, bulk modulus, and related thermodynamic properties. Notably, the bulk modulus displays a consistent increase across the compositions. Despite the similar nanomaterial compositions, the calculated G/B ratio, an indicator of ductility versus brittleness, was found to be ~0.976, suggesting predominantly ionic bonding in Mg_2X compounds. Thermal conductivity (k_min) was computed across the series, demonstrating a steady increase from Mg_2Si to Mg_2Sn, accompanied by a consistent hardness profile. Additionally, ultrasonic attenuation and velocity measurements were conducted, revealing composition-dependent variations. The materials exhibit minimum attenuation at initial states, indicating maximum purity, while the lowest attenuation values correspond to enhanced ductility. Previous first-principles studies on Mg_2X (X = Si, Ge, Sn) have primarily focused on electronic and thermoelectric properties, with limited attention to elastic, mechanical, and ultrasonic behavior at the nanoscale. This study fills that gap by providing comprehensive insights into higher-order elastic constants and their impact on the mechanical and thermoelectric performance of Mg_2X nanomaterials.

Keywords: Elastic properties, Mg_2X (X=Si, Ge, and Sn) composition, Mechanical properties, Thermal properties, Ultrasonic properties.

I. INTRODUCTION

The composite semiconducting nanomaterials Mg_2X (X = Si, Ge, Sn) exhibit exceptional properties—such as low electrical resistivity, low thermal conductivity, and high Seebeck coefficients [1,2] which make them highly attractive for integration into both conventional and next-generation high-performance energy storage systems as battery electrode materials. As members of the thermoelectric Mg_2X nanocomposite family, these compounds offer a compelling combination of electrical and thermal characteristics suited for advanced energy applications [3–5].

In addition, Mg_2X compounds possess relatively low band gaps (0.3–0.6 eV), rendering them suitable for optoelectronic applications, including infrared detectors and optical fiber technologies [6]. Magnesium itself is also considered a promising material for hydrogen storage, owing to its high hydrogen uptake capacity (7.6 wt.%). More recently, the Mg_2X system, where X belongs to group IV(B) of the periodic table (i.e., Si, Ge, Sn), has received growing attention in the search for novel intermetallic alloys with enhanced thermodynamic behavior for hydrogen storage, particularly at cryogenic temperatures [7].

The Mg_2X (X = Si, Ge, Sn) compounds are known to crystallize in three primary structures: hP6 (hexagonal), cF12 (cubic), and cF24 (cubic). Among these, recent studies have predominantly focused on the cubic cF12 phase due to its favorable structural and physical properties. Whitten et al. [8] and Chung et al. [9] employed a resonance method as early as 1965 to determine the elastic constants of cF12-Mg_2X in the temperature range of 80–300 K. Using these experimentally derived elastic constants along with optical data, they further computed the lattice vibrational frequencies.

Building on this foundational work, Murtaza et al. [10,11] investigated the thermoelectric, optoelectronic, and elastic properties of cF12-Mg_2X. Their findings revealed that as the composition varies, the bulk modulus increases while the lattice constant decreases—trends that contribute to the materials' narrow bandgaps and high thermal potential, making them highly suitable for optoelectronic and thermoelectric device applications. Furthermore, generalized gradient approximation (GGA) calculations of the thermodynamic and elastic parameters indicated that Mg_2Si exhibits superior structural stability, while cF12-Mg_2Ge demonstrates the strongest alloying capability among the three compositions [12].

Additionally, in cF12-Mg_2X (X = Si, Ge, Sn) compounds, the temperature dependence of the stiffness elastic coefficients Cij(T)C_{ij}(T)Cij(T) has been determined by correlating the equilibrium volume as a function of temperature V(T)V(T)V(T) with the stiffness coefficients Cij(V)C_{ij}(V)Cij(V), which are derived from volume-dependent phonon calculations [13]. First-principles studies on binary magnesium compounds have shown that Mg_2Si exhibits a mixed bonding character, incorporating both ionic and covalent components [14]. Specifically, the bonding in cF12-Mg_2Si is characterized by a complex interplay of metallic, ionic, and covalent interactions. This hybrid bonding nature contributes to the compound's structural stability and intrinsic brittleness, as revealed through detailed thermodynamic, elastic, and electronic structure analyses [15].

The pyroelectric, optoelectronic, and elastic properties of cF12-Mg_2X (X = Si, Ge, Sn, and Pb) reveal that while Mg_2Sn exhibits significant elastic anisotropy, both cF12-Mg_2Si and cF12-Mg_2Ge behave as nearly isotropic materials [16]. One of the key advantages of thermoelectric solid-state power generators lies in their simple architecture, absence of moving parts, and modest efficiency in converting thermal energy directly into electrical power [17]. As early as 1961, Nikitin et al. [18] recognized that Mg_2BIV compounds (BIV = Si, Ge, Sn) possessed a favorable combination of chemical and physical properties, making them promising candidates for the development of efficient thermoelectric materials. Their structural and electronic similarities to elemental group IV semiconductors further support their potential in thermoelectric applications [19, 20].

Among these, magnesium silicide (Mg_2Si), which crystallizes in the antifluorite structure, has been identified as a particularly promising material for thermoelectric energy conversion in the temperature range of 500–800 K. It is a narrow bandgap semiconductor with an indirect bandgap of approximately 0.77 eV [21, 22]. Moreover, Mg_2Si is abundant in the Earth's crust, non-toxic, and environmentally benign—distinct advantages over conventional thermoelectric materials such as PbTe and $CoSb_3$, which may pose environmental or health risks [23].

The anti-fluorite crystal structure (space group $Fm\bar{3}m$) of semiconducting magnesium silicide (Mg_2Si) and magnesium germanide (Mg_2Ge) has positioned these materials as strong candidates for high-performance thermoelectric applications. Their remarkable properties—including low electrical resistivity, low thermal conductivity, and high Seebeck coefficients—make them particularly attractive for energy conversion technologies [24]. Mg_2Si and Mg_2Ge exhibit a broad spectrum of temperature-dependent physical behaviors, largely governed by their lattice dynamical properties, making the understanding of their vibrational characteristics essential for optimizing performance.

Among the Mg_2X compounds (X = Si, Ge, Sn), Mg_2Si stands out as the only stoichiometric silicide with stable existence, distinguishing it from most other silicides. These materials have attracted significant attention in recent years due to their high thermoelectric efficiency, especially in the mid-temperature range of 500–800 K. Their suitability arises from a combination of high electrical conductivity and relatively high melting points—1358 K for Mg_2Si, 1030 K for Mg_2Sn, and 1390 K for Mg_2Ge—which enhance their thermal and mechanical robustness in demanding applications. Furthermore, Mg_2X compounds are lightweight materials exhibiting high specific strength, along with excellent elastic and mechanical properties, particularly in their nanocomposite semiconducting forms [25].

In this study, we investigate the single-crystal elastic constants—specifically the second-order (SOEC) and third-order (TOEC) elastic constants—along with key thermo-physical, ultrasonic, and thermal properties of Mg_2X (X = Si, Ge, and Sn) using Lennard-Jones potential analysis. The evaluated parameters include Debye velocity, Debye temperature, ultrasonic attenuation coefficient, stiffness constants, thermal relaxation time, Young's modulus (Y), Poisson's ratio, bulk modulus (B), shear modulus (G), and Pugh's ratio (B/G). This study aims to deliver a comprehensive understanding of the elastic and mechanical properties of Mg_2X nanomaterials, offering valuable insights for their application in next-generation energy storage and thermoelectric devices. These comprehensive insights provide valuable guidance for the future design and development of Mg_2X-based compositions for advanced functional and thermoelectric applications.

2. COMPUTATIONAL THEORY

2.1. Second order elastic constant & third-order elastic constant of Mg_2X (X=Si, Ge, and Sn)

To provide a concise description of the material's physical and mechanical properties as well as its anisotropic attenuation behavior, several theoretical approaches are available in addition to experimental methods that can be used to examine the previously mentioned characteristics through formulation. Among the researchers, the well-known Lennard-Jones interaction potential techniques are widely employed because of their high level of accuracy in producing results according to experimental standards. This evaluation provides definitions for the additional-order elastic constants, including the experimental material referred to as second-order (SOEC) and its third-order (TOEC) value, and this potential model's evaluation technique demonstrates how the second-order elastic constants connected to materials with a hexagonal structure are influenced by the lattice parameters.

$$C_{IJ} = \frac{\partial^2 U}{\partial e_I \partial e_J}; \qquad \text{I or J} = 1\ldots6 \tag{1}$$

$$C_{IJK} = \frac{\partial^3 U}{\partial e_I \partial e_J \partial e_K}; \ \text{I or J or K} = 1\ldots6 \tag{2}$$

Also the general potantial farmulation for the Lennard-Jones model is well defined by

$$\varphi(r) = -\frac{a_0}{r^m} + \frac{b_0}{r^n} \tag{3}$$

Where,$a_0 = 4\varepsilon\sigma^m$, and $b_0 = 4\varepsilon\sigma^n$ here m and n are fitting parameters and ε and σ are the inter-atomic potential parameters.

Equations 4 and 5 represent general mathematical simplifications, where the third-order elastic constants (TOECs) are decomposed into ten fundamental parametric forms, and the second-order elastic constants (SOECs) into six. This decomposition facilitates further simplification processes, enabling precise theoretical evaluation of the material properties previously described. These formulations for SOECs and TOECs are detailed in Equations 4 and 5, respectively [26].

$$\left.\begin{aligned} C_{11} &= 24.1\ \mathrm{p}^4 C' \\ C_{12} &= 5.918\ \mathrm{p}^4 C' \\ C_{13} &= 1.925\ \mathrm{p}^6 C' \\ C_{33} &= 3.464\ \mathrm{p}^8 C' \\ C_{44} &= 2.309 p^4 C' \\ C_{66} &= 9.851\ \mathrm{p}^4 C' \end{aligned}\right\} \tag{4}$$

$$\left.\begin{aligned} C_{111} &= 126.9p^2B + 8.853p^4C' \\ C_{112} &= 19.168p^2B - 1.61p^4C' \\ C_{113} &= 1.924p^4B + 1.155p^6C' \\ C_{123} &= 1.617p^4B - 1.155p^6C' \\ C_{133} &= 3.695p^6B \\ C_{155} &= 1.539p^4B \\ C_{144} &= 2.309p^4B \\ C_{344} &= 3.464p^6B \\ C_{222} &= 101.039p^2B + 9.007p^4 \\ C_{333} &= 5.196p^8B \end{aligned}\right\} \tag{5}$$

Here, where p = c/a, is an axial ratio for the hexagonal material or compound and

$$C' = \chi\, \mathrm{a} / \mathrm{p}^5\ , \qquad B\psi\, \mathrm{a}^3 = /p^3,$$
$$\chi = (1/8)[\{n - m)\}/\{\mathrm{a}^{n+4}\}], \qquad \psi = -\chi /\{6\, \mathrm{a}^2 (m + b_0\,(nn + 6)\}$$

Is evaluated using integer values, and the Lennard-Jones potential variable, b_0, is widely known. Additionally, these characteristics may be changed even with a small change in temperature or attenuation because they are incredibly dependent on these factors.

2.2. Ultrasonic Velocity

In order to understand the crystallographic behavior of experimental substances, we have calculated the shear velocity (V_{S2}), quasishear velocity (V_{S1}), and atomic vibratory pattern contingent Ultrasonic longitudinal velocity (V_L) at different propagation inclinations along the unique axis (c-axis). Additionally, we have computed the Debye average velocity (V_D) and Debye temperature (θ_D) using the SOECs and TOECs, the elastic properties of nanomaterials have been a subject of extensive research. For instance, the measurement of elastic properties in epoxy resin/polyvinyl alcohol nanocomposites using ultrasonic wave velocities I. oral et al. [27], provides valuable insights into the methodology and applications of elastic property measurement, which are relevant to our study. Mathematics is used to figure out the direction and propagational angle-dependent ultrasonic shear and longitudinal wave velocity (V_L and V_S) for waves traveling alongside a z-axis in a crystal with a hexagonal shape.

$$\left.\begin{aligned}
V_L^2 &= \{C_{33}\cos^2\theta + C_{11}\sin^2\theta + C_{44}\\
&\quad +\{[C_{11}\sin^2\theta - C_{33}\cos^2\theta\\
&\quad +C_{44}(\cos^2\theta - \sin^2\theta)]^2\\
&\quad +4\cos^2\theta\sin^2\theta\,(C_{13}+C_{44})^2\}^{1/2}\}/2\rho\\
V_{S1}^2 &= \{C_{33}\cos^2\theta + C_{11}\sin^2\theta\\
&\quad +C_{44} - \{[C_{11}\sin^2\theta - C_{33}\cos^2\theta\\
&\quad +C_{44}(\cos^2\theta - \sin^2\theta)]^2\\
&\quad +4\cos^2\theta\sin^2\theta\,(C_{13}+C_{44})^2\}^{1/2}\}/2\rho\\
V_{S2}^2 &= \{C_{44}\cos^2\theta + C_{66}\sin^2\theta\}/\rho
\end{aligned}\right\} \qquad (6)$$

Also from the previously defined value of elastic constants along various tensor directions, we can also define the value of sound velocity in longitudinal and shear direction as [28].

$$V_L = \sqrt{\frac{C_{33}}{\rho}} \qquad (7)$$

$$V_S = \sqrt{\frac{C_{44}}{\rho}} \qquad (8)$$

where V_{S1}, V_{S2}, and V_L stand for longitudinal, quasi, and shear wave ultrasonic velocities, respectively. where the variables (V) stand for the material's density and the angle with the crystal's unique axis, respectively. Also, V_l : Longitudinal wave velocity, associated with compressional waves. V_s : Shear wave velocity, associated with transverse waves. V_D: Debye average velocity, derived from V_l and V_s, and related to the phonon transport and thermal properties of the material.

An essential physical measure for material characterization is the Debye Temperature (T_D), which is closely linked to the Debye Average Velocity (V_D). The formula can be used to determine the density of a material with a hexagonal structure.

$$d = \frac{2\mathrm{Mn}}{3\sqrt{3a^2cN_A}} \qquad (9)$$

M, NA, and n stand for molecular weight, Avogadro number, and atoms per unit cell, respectively. At the low temperature in physics, the associated Debye average velocity is a crucial metric since ultrasonic velocities are dependent on elastic constants. It can be described in general as

$$V_D = \left[\frac{1}{3}\left(\frac{1}{V_L^3} + \frac{1}{V_{s1}^3} + \frac{1}{V_{s2}^3}\right)\right]^{-1/3} \qquad (10)$$

Further, the Deby velocity relative Debye temperature for our experimental hexagonal material can be generally described by

$$T_D = \hbar V_D \frac{6\pi^2 n_a)^{\frac{1}{3}}}{K_B} \qquad (11)$$

Where K_B: Boltzmann constant, n_a: atomic concentration; constant

Due to the advanced engineering application of experimental HCP material in a temperature and composition limit, it becomes important to investigate their thermophysical behavior at those extreme conditions. Utilizing the Debye model of heat capacity, we have calculated the thermal energy densities (E_0) and heat capacity (C_V) [29].

Within the scope of these experiments, the period of time taken during thermodynamic phonons to reassemble themselves following the propagation of an ultrasonic wave is known to be the corresponding thermal relaxation time (τ). It has a direct relationship with specific heat (C_V), Debye average velocity, and thermal conductivity (k) [30].

$$\tau = \tau_S = \tau_L/2 = \frac{3k}{C_V V_D^2} \qquad (12)$$

Based on the experimental analysis of the evaluated physical properties, the material's stability, crystal structure, thermal conductivity, mechanical strength, and hardness can be reliably inferred from the calculated second-order elastic constants (SOECs). Furthermore, the Voigt–Reuss–Hill (VRH) approximation [31] has been employed to investigate variations in key mechanical parameters, including Pugh's ratio (B/G), Young's modulus (Y), Poisson's ratio (σ), bulk modulus (B), and shear modulus (G). The standard relationships for these parameters within the VRH framework are expressed below in Equation (13):

$$M = C_{11} + C_{12} + 2C_{33} - 4C_{13}$$

$$\mathrm{C}^2 = (C_{11} + C_{12})C_{33} - 4C_{13} + C^2{}_{13}$$

$$B_R = \frac{C^2}{M}$$

$$\mathrm{B}_V = \frac{2(C_{11} + C_{12}) + 4C_{13} + C_{33}}{9}$$

$$G_V = \frac{M + 12(C_{44} + C_{66})}{30}$$

$$\mathrm{G}_R = \frac{5C^2C_{44}C_{66}}{2[3B_VC_{44}C_{66} + C^2(C_{44} + C_{66})]}$$
$$Y = \frac{9GB}{G + 3B}$$
$$\mathrm{B} = \frac{B_V + B_R}{2} \tag{13}$$
$$\mathrm{G} = \frac{G_V + G_R}{2}$$
$$\sigma = \frac{\mathrm{3B\text{-}2G}}{2(\mathrm{3B}+G)}$$

The lattice thermal conductivity (κ) was evaluated by following the Morelli and Slack [32] approach, which is given as equation 14:

$$k = AMT_D{}^3\delta^3/\gamma^2Tn^{2/3} \tag{14}$$

Ma (in amu) is the average atomic mass, θ_D is the Debye temperature, T is the temperature, δ3 is the amount of volume for each atom, γ is the Grüneisen number, and A represents the proportional constant. n is the total number of atoms per unit cell. Here the Grüneisien number determines the numerical value for the constant 'A', which is stated as follows.

$$A = 2.43 \times 10^{-8}/(1 - \frac{0.514}{\gamma} + \frac{0.228}{\gamma^2}) \tag{15}$$

γ is the Grüneisen number, which can be determined by $\gamma = \alpha B/(\mathrm{CV}\dot{\rho})$,where α is the thermal volume expansion coefficient, B is bulk modulus, C_V is the heat capacity, and ρ is density of the material.

The relative values of the melting temperatures (T_m) of hexagonal crystals [33] Is also dependent on the value of certain elastic constants C_{11} and C_{33} which are defined as, melting temperatures (T_m)

$$T_m = 354 + \frac{4.5(2C_{11}+C_{33})}{3} \tag{16}$$

Where C_{11} and C_{33} are in GPa, and T_m is in K.

2.3. Attenuation

The ultrasonic attenuations in single crystal solids at high temperatures are mostly due to phonon-phonon interactions (Akhiezer loss) and thermoelastic relaxation mechanisms (thermoelastic loss). The formulation for estimating ultrasonic attenuation due to phonon-phonon interactions was developed by [34] and is given by:

$$(\alpha/f^2)_{Akh}=4\pi 2\tau E_0(D/3)/2\rho V^3 \tag{17}$$

The ultrasonic wave frequency is denoted by f, and the thermal energy density is denoted by E_0.

The thermoelastic loss $(\alpha / f^2)_{Th}$: equation as follows

$$(\alpha/f^2)_{Th} = 4\pi^2 < \gamma_i^j >^2 \frac{kT}{2\rho V_L^5} \tag{18}$$

$$\alpha_{long} = \frac{2\pi^2f^2}{\rho V_l^3}\left(\frac{4}{3}\eta_e + \chi\right) \tag{19}$$

$$\alpha_{shear} = \frac{2\pi^2f^2}{\rho V_S^3}\eta_e \tag{20}$$

The total consequence of ultrasonic attenuation $(\alpha/f^2)_{Total}$, caused by the thermoelastic loss and Akhieser losses exhibited in the longitudinal, shear mode of propagation, can be calculated as [35]:

$$(\alpha/f^2)_{Total} = (\alpha/f^2)_{Th} + (\alpha/f^2)_L + (\alpha/f^2)_S \tag{21}$$

Where$(\alpha/f^2)_{Th}$is the thermoelastic loss, while $(\alpha/f^2)_L$and $(\alpha/f^2)_S$ are the ultrasonic attenuation coefficients for the longitudinal wave and shear wave, correspondingly.

3. RESULTS AND DISCUSSION

3.1. Second order elastic constant & third-order elastic constant of Mg_2X (X=Si, Ge, and Sn)

In the present study on Mg_2X (X = Si, Ge, and Sn) composite semiconducting nanomaterials, we have successfully derived accurate values for various elastic coefficients—specifically six second-order elastic constants (SOECs) and ten third-order elastic constants (TOECs)—using an interaction potential approach. Based on our theoretical framework and comparative analysis, the obtained results show strong agreement with available experimental data, exhibiting minimal deviations and confirming the validity of our model.

Additionally, we have determined the essential lattice parameters for Mg_2X (X = Si, Ge, and Sn), where “a” represents the basal plane lattice constant and “p” denotes the axial ratio (p=c/a)(p = c/a)(p=c/a) for the hexagonal close-packed (HCP) structure [36]. The composition-dependent lattice constants for Mg_2Si, Mg_2Ge, and Mg_2Sn are found to be 4.629 Å, 4.739 Å, and 5.0913 Å, respectively, as illustrated in Figure 1. Corresponding densities of states (DOS) for each composition are also provided. The constant value of Leonard Jones potential constant “b_0” is also evaluated with mathematical simulation manipulation technique and finally taken at $5.07\times e^{-62}$,$6.37\times e^{-62}$,$1.56\times e^{-66}$ erg cm^3, respectively, for different nanomaterials.

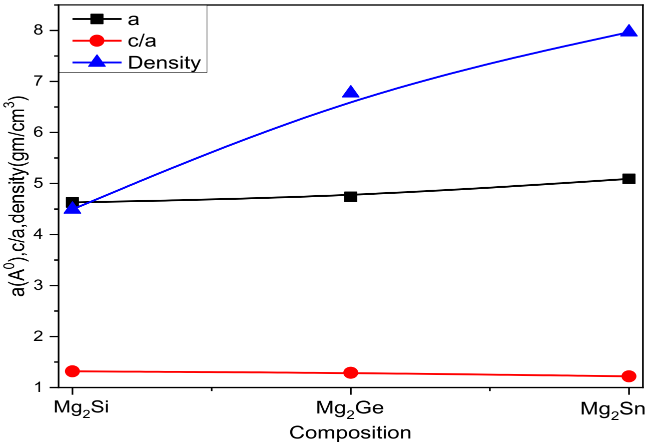


Figure 1. The dedicated value of Lattice constant a (Å), Bessel plane distance c/a, and relative density, with nanomaterial composition Mg$_2$X (X=Si, Ge, and Sn)

Additionally, on the basis of the above-mentioned and evaluated data, the evaluated values of the second and third-order elastic constants of the Mg$_2$X (X=Si, Ge, and Sn) are depicted in Figure 2 and Figure 3, respectively.

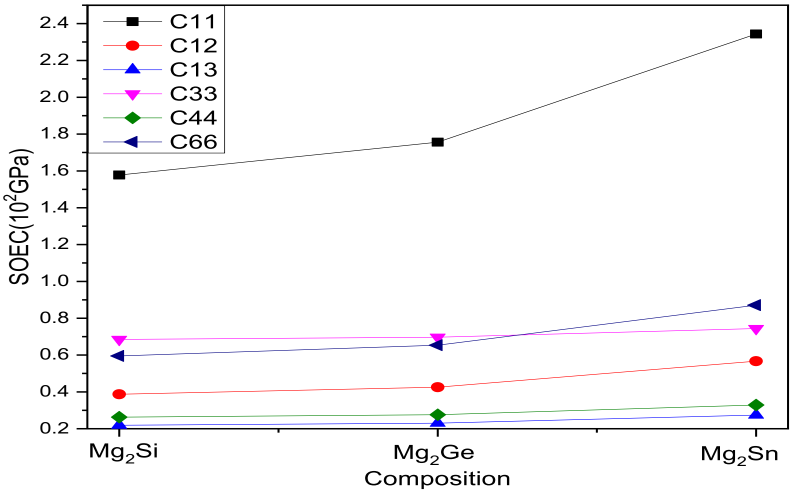


Figure 2. The variation in the respective values of SOECs with nanomaterial composition Mg$_2$X (X=Si, Ge, and Sn)

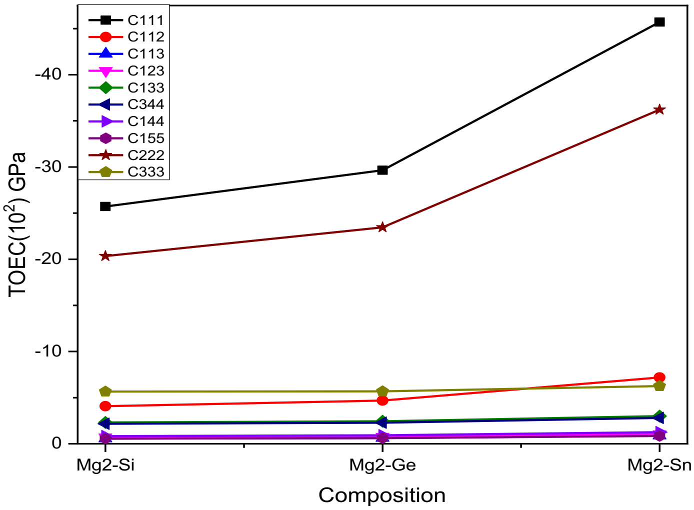


Figure 3. The variation in the respective values of TOECs with nanomaterial composition Mg$_2$X (X=Si, Ge, and Sn)

The above Figure 2 depicts the evaluated variation of SOEC of intrinsic semiconductor material (Si, Ge, and Sn) with the moderate band gap conducting material Mg in the composite format. On the basis of data analysis for the particular HCP complex material Mg_2X (X=Si, Ge, and Sn), we can conclude that the negative values of TOECs indicate the presence of strain within the material Mg in composition. This is aligned with previous investigations on materials with hexagonal structures, affirming the theoretical framework's accuracy in valuing higher-order elastic coefficients. As a result, the theory used to investigate higher-order elastic coefficients is justified using Eqs. 1 and 2. For HCP structure material stability, here the five different SOECs (C_{ij}, namely C_{11}, C_{12}, C_{13}, C_{33}, and C_{44}) satisfy Born-Huang's norms [37]. This Mg_2X (X=Si, Ge, and Sn) is mechanically stable because it is evident that the positive elastic constant values satisfy Born-mechanical Huang's stability constraints defined as.

$$\acute{C}_{44} > 0,\ \acute{C}_{11} - |\acute{C}_{12}| > 0$$

$$(\acute{C}_{11} + \acute{C}_{12})\ \acute{C}_{33} - 2\ \acute{C}^2_{13} > 0 \quad (22)$$

Where $\acute{C}_{jj} = \acute{C}_{jj} - P$, (P=1, 3, 5). $\acute{C}_{12} = C_{12} + P$, $\acute{C}_{13} = C_{13} + P$

3.2. Mechanical Property

As the composition varies from Si to Ge and Sn, the density of the Mg-based compounds increases, while their corresponding atomic volume decreases, as illustrated in Figure 4. The values of bulk modulus (B), Young's modulus (Y), and shear modulus (G) for Mg_2X (X = Si, Ge, and Sn) are observed to increase progressively with changes in composition. This trend reflects an enhancement in the stiffness and bonding strength of the nanomaterials.

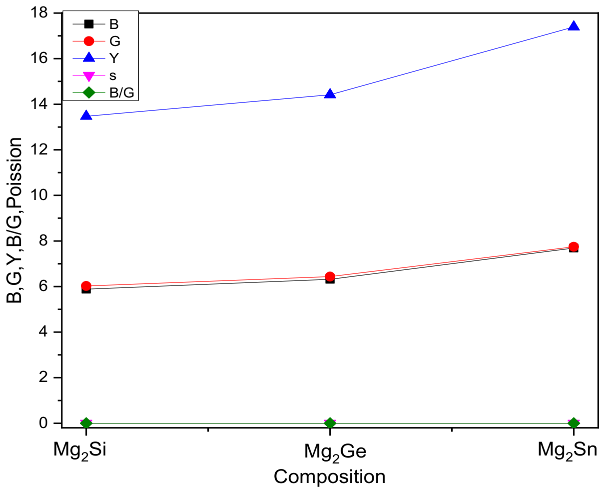


Figure 4. B, G, Y, B/G, σ with nanomaterial composition Mg_2X (X=Si, Ge, and Sn)

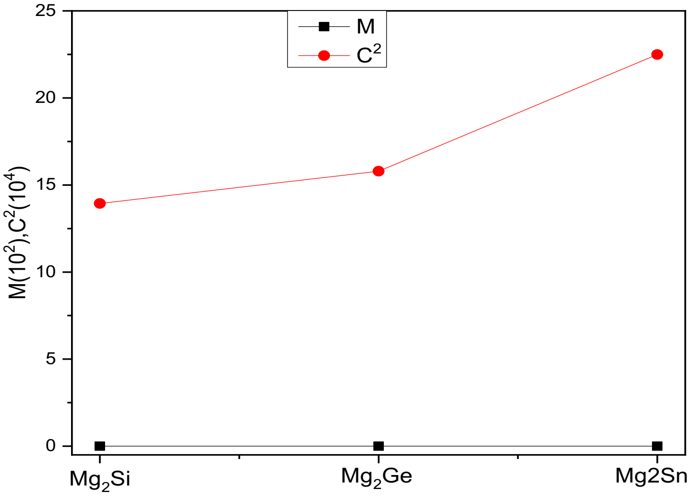


Figure 5. M, C^2 with nanomaterial composition Mg_2X (X=Si, Ge, and Sn)

Additionally, the G/B ratio, which is 0.97, 0.98, and 0.99 for Mg_2Si, Mg_2Ge, and Mg_2Sn respectively, signifies a predominant ionic bonding character within these compounds. This ratio, along with the B/G and Poisson's ratio (σ), provides valuable insights into the mechanical behavior of the materials, particularly in predicting their brittleness or ductility. Materials with $\sigma=0.23\leq 0.26$ and $B/G=1.52\leq 1.75$ are typically considered brittle, or else they are ductile in nature [38]. Our results for B/G indicate that the Mg_2X (X=Si, Ge, and Sn) is brittle in nature with nanomaterial composition because they are under their critical levels. The value of σ, which should perfectly be less than 0.5 for elastic and stable materials, is found to be within an acceptable range for composite. B/G ratio below 1.75 typically signifies brittleness, which is consistent with the behavior observed in Mg_2X compositions. In the below section, we have demonstrated the relative values of Voigt–Reus' constants (M and C^2), Br (x 10^{10} Nm^{-2}), Bv (x 10^{10} Nm^{-2}), Gr (x 10^{10} Nm^{-2}), Gv (x $10^{10}Nm^{-2}$), Y (x $10^{10}Nm^{-2}$) for Mg_2X (X=Si, Ge, and Sn).

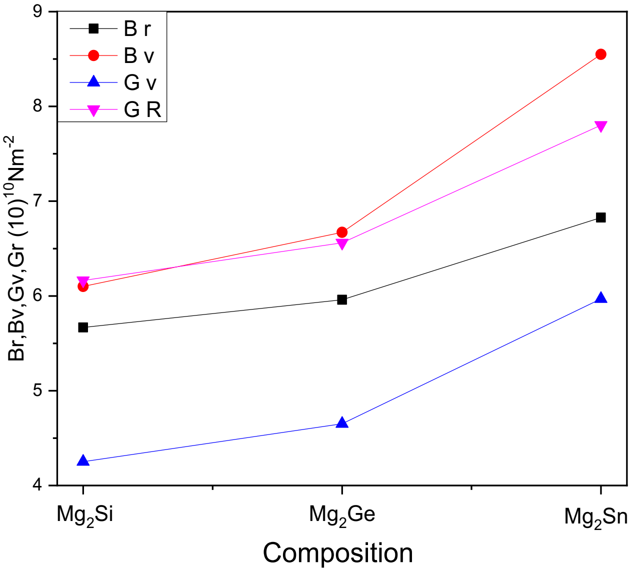


Figure 6. B_r, B_v, G_v, and G_r, with nanomaterial composition. Mg_2X (X=Si, Ge, and Sn)

In addition to the mechanical properties discussed above, the melting temperature serves as a critical parameter in the development of materials with enhanced thermal stability, ductility, and brittleness. The melting temperature (Tm) and hardness (Hv) of Mg_2X (X = Si, Ge, and Sn) have been systematically investigated and are presented in Figure 7, illustrating their variation as a function of composition.

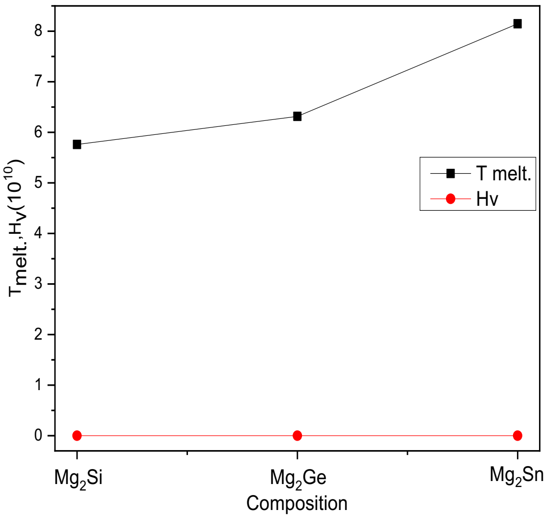


Figure 7. Hardness (H_v), melting temperature (T_m) with nanomaterial composition Mg_2X (X=Si, Ge, and Sn)

Table 1 demonstrates that elastic anisotropy can be effectively described by the unified anisotropic in elasticity may always be explained by the unified anisotropic index (A^U), percent anisotropic (AB and AG), and shear anisotropy factors (A_1, A_2, and A_3).

Table 1 shows that under different compositions of Mg_2X (X=Si, Ge, and Sn), AB has a higher percentage of anisotropy than AG. This suggests that the bulk modulus is less oriented than the shear modulus. A_1, A_2, and A_3 findings with nanomaterial are also displayed in Table 1. If $A_1 = A_2 = A_3 = 1$, the substance is an isotropic crystal. According to the findings, composition Mg_2X (X=Si, Ge, and Sn) is an anisotropic crystal absent in any other circumstances. With nanomaterial, the universal anisotropic index (A^U) deviates from zero, which is known as the highly single crystal anisotropy [39].

Table 1. Various anisotropy constants with their relative poisson's ratio for nanomaterial composition Mg_2X (X=Si, Ge, and Sn)

Composition	A^U	A_1	A_2	A_G	A_3	σ
Mg_2Si	**-1.47**	0.577	0.441	0.395	1.003	0.118
Mg_2Ge	**-1.33**	0.588	0.422	0.395	1.000	0.119
Mg_2Sn	**-0.92**	0.479	0.377	0.395	1.003	0.122

It is clear form Figure 8 that for all temperature, the values of D_S are larger than those of D_l for all composition nanomaterials. It indicates that for the shear ultrasonic wave the transformation of ultrasonic energy into thermal energy is less than for the longitudinal ultrasonic wave.

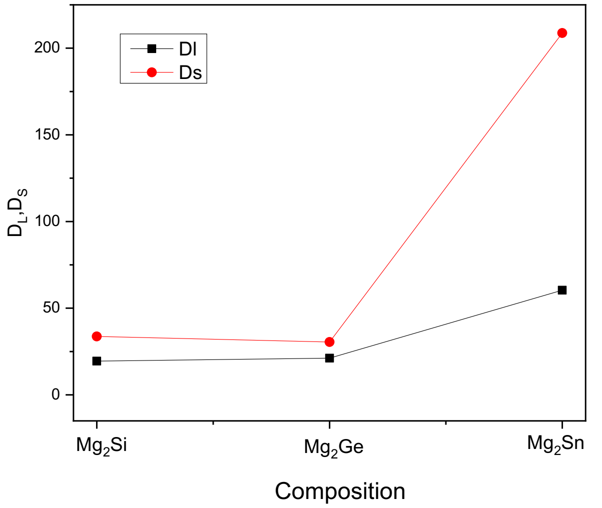


Figure 8. Variation of D_L, D_S with nanomaterial composition Mg_2X (X=Si, Ge, and Sn)

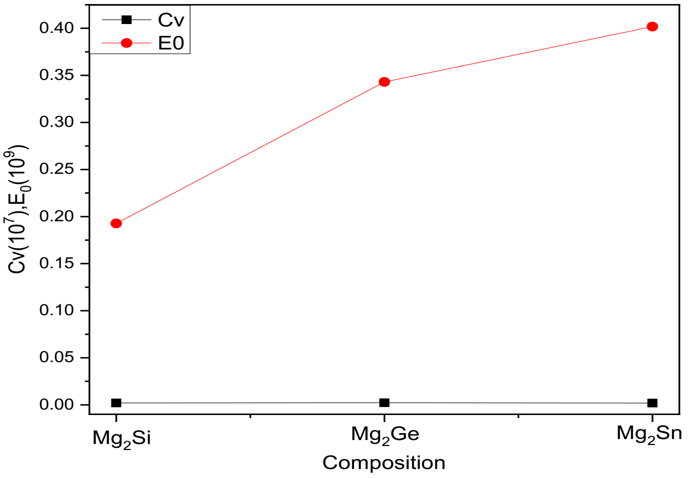


Figure 9. Variation of E_o and C_v with nanomaterial composition Mg_2X (X=Si, Ge, and Sn)

3.3. Ultrasonic velocity and allied parameters

The below Figures 10-13 demonstrate the angular dependency of ultrasonic velocities (V_L, V_{S1}, V_{S2}, V_D) with varying composition along the z-direction for the Mg_2X (X=Si, Ge, and Sn). In Figs. 10 and 11, the z-axis of the Mg_2X (X=Si, Ge, and Sn) correlates directly to the minima and maxima of the ultrasonic velocities V_L and V_{S1}. and the value of V_{S2} is demonstrated

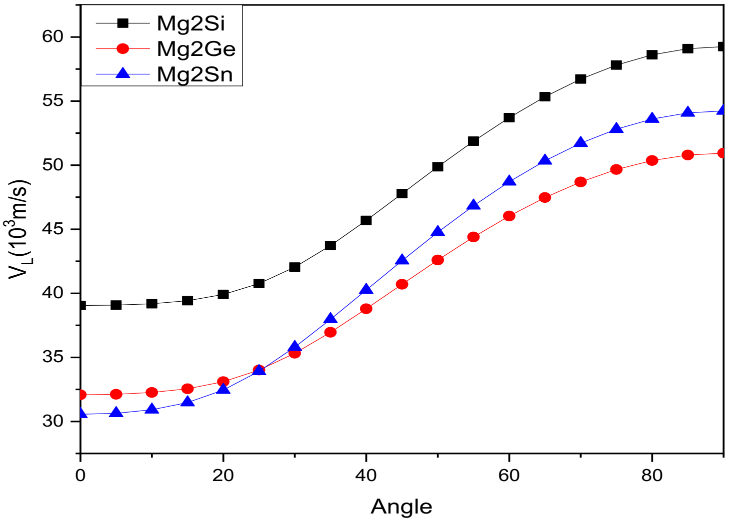


Figure 10. V_L vs. "ϴ"   for different nanomaterial compositions Mg_2X (X=Si, Ge, and Sn)

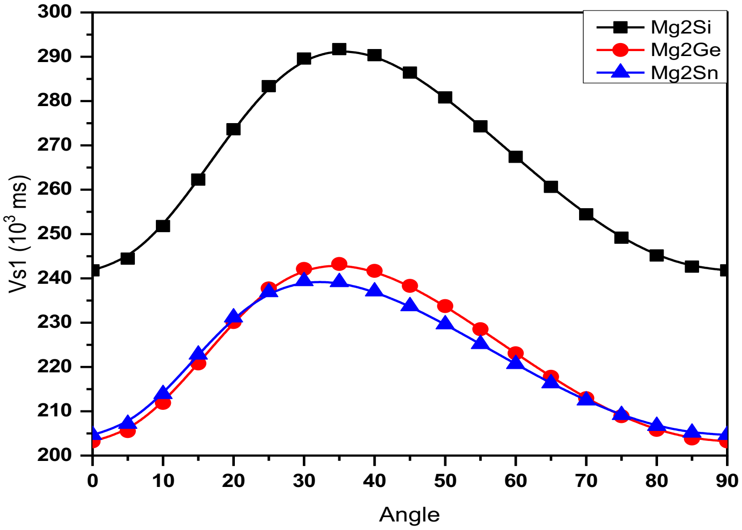


Figure 11. Vs_1 vs "ϴ" for different nanomaterial compositions Mg_2X (X=Si, Ge, and Sn)

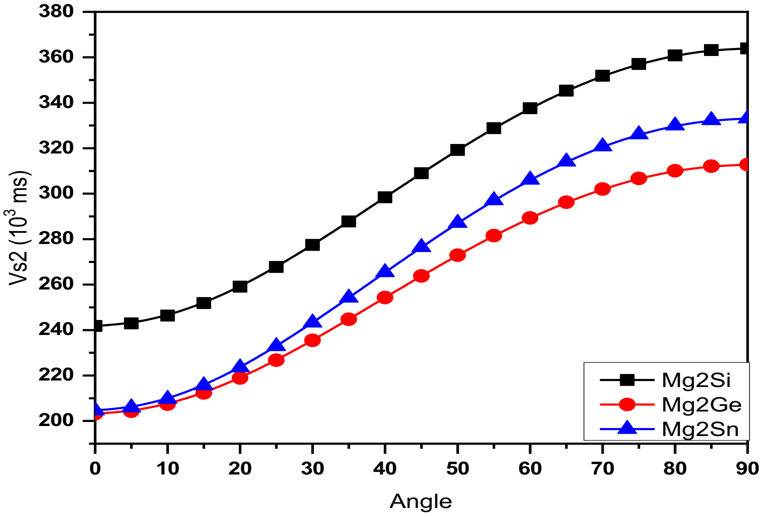


Figure 12. Vs_2 vs "ϴ" for different nanomaterial compositions Mg_2X (X=Si, Ge, and Sn)

in Fig. 12, and it is clear from the see-through that its values rise with the angular spread along the z-direction. The combined effects of SOECs and ρ (density) are held responsible in Favor of the uncharacteristic behavior of angle-dependent velocity. This effort's discovery of the angle (ϴ)-dependent velocity (V) curves is similar to those found for other hcp nanomaterials. This justifies the angle dependence of metal velocities [40]. Here, the V_D value rises with increasing angle and reaches its maximum at 55.0 degrees in figure 13. The relative value of the Debye average velocity is also influenced by the variation in the fundamental ultrasonic velocities, V_L, V_{S1}, and V_{S2}, which are used to evaluate V_D. Additionally, here When the z-axis of this crystal is at 55°, the average sound wave velocity shows a maximum.

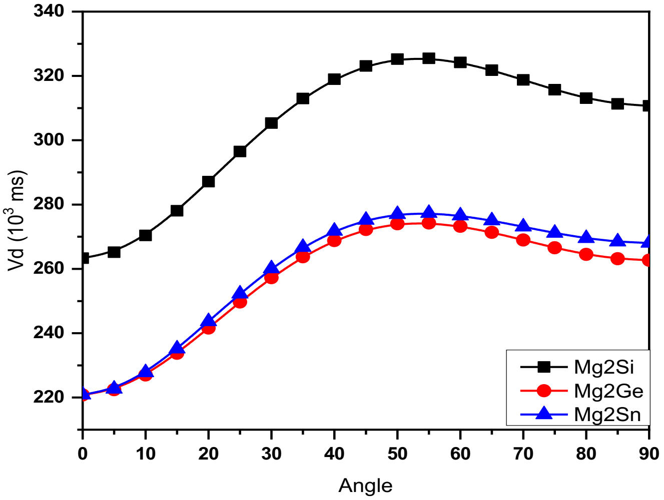


Figure 13. V_D vs "ϴ" for different nanomaterial compositions Mg_2X (X=Si, Ge, and Sn)

In this section, Figure 14 presents a graphical depiction of the calculated thermal relaxation time, illustrating its variation with composition across different angular orientations. Here, the reciprocal characteristics of V_D, which are relative to $\tau \propto 3k/C_V V_D^2$ is evaluated, and here we can conclude that the value of thermal relaxation time for the Mg_2X (X=Si, Ge, and Sn) is highly dependent on the value of the Fermi radius and can be affected by 'k'. The hexagonal structure of the Mg_2X (X=Si, Ge, and Sn) is explained by the computation. The minimum value of the thermal relaxation time (τ\tauτ) occurs along the wave propagation direction at an angle of 55°, indicating the shortest re-establishment time for thermally generated phonons and their equilibrium distribution. Consequently, the presence of phonon–phonon (p–p) interactions and the associated thermal relaxation give rise to ultrasonic attenuation observed in the material along this direction.

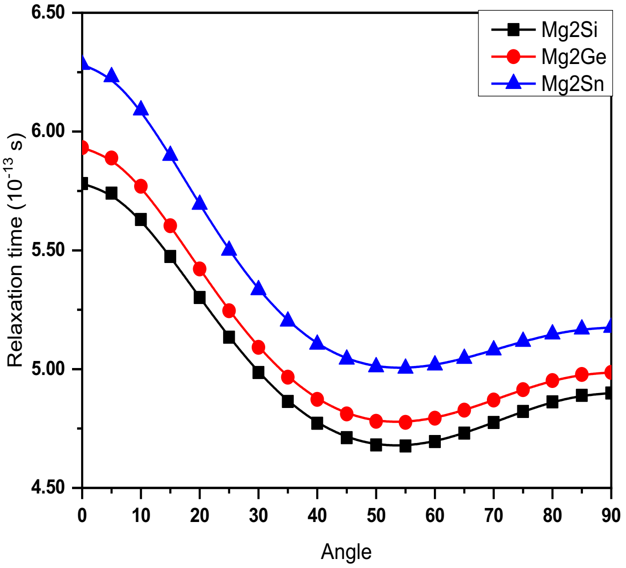


Figure 14. Relaxation time vs "ϴ" with z- axis of crystal for different nanomaterial compositions Mg_2X (X=Si, Ge, and Sn)

3.4. Ultrasonic attenuation due to phonon-phonon interaction and thermal relaxation phenomena

In analyzing the ultrasonic attenuation, it is assumed that the wave flows along the z-axis of nanomaterial composition Mg_2X (X=Si, Ge, and Sn). The Attenuation coefficient is alienated over frequency squared for longitudinal wave $(\alpha/f^2)_L$and for shear wave $(\alpha/f^2)_s$ under the condition ωτ<<1 at various compositions is calculated using Eq. 11. The thermo-elastic loss by frequency squared $(\alpha/f^2)_{th}$ is calculated using Eq. 11, and Table 2 illustrates the ultrasonic attenuation value resulting from the thermoelastic relaxation mechanism **(**$\alpha/f^2)_{th}$ **.**

It is assumed that the wave propagates down the z-axis of the Mg_2X (X=Si, Ge, and Sn) compound when computing ultrasonic attenuation (UA). Eqn. 14,15, and 16 is used for calculating the attenuation coefficient divided by frequency squared (α/f^2) Akh for the wave's longitudinal direction (α/f^2) L and for the shear wave (α/f^2). Table 2 presented values of the composition-dependent longitudinal, shear and total attenuation of the Mg_2X (X=Si, Ge, and Sn**).** Energy losses$(\alpha/f^2)_{Akh}$ are substantially proportional to D, E_0, and V^{-3} (Eq. 11). Thus, in the nanomaterial composition Mg_2X (X=Si, Ge, and Sn), E_0 and K have a significant impact on Akhieser losses. The thermo-elastic loss for the nanomaterial composition Mg_2X (X=Si, Ge, and Sn) is to be substantially less than the Akhieser loss, indicating that the p-p interaction mechanism-induced ultrasonic attenuation dominates over the thermoelastic loss. Total attenuation is mainly driven by two variables, namely thermal energy density and thermal conductivity. Thus, although nano material compositions of Mg_2X (X=Si, Ge, and Sn) are the least ductile, they show their purest form at for different nanomaterial compositions and show additional ductility, as shown by the minimal attenuation. Hence, the nanomaterial composition Mg_2X (X=Si, Ge, and Sn) will have the lowest impurity content at ambient compositions.

Table 2. Long. & Shear attenuation for different nanomaterial composition Mg_2X (X=Si, Ge, and Sn)

Composition	$\left\{\frac{\alpha}{f^2}\right\}_{shear}$ (10^{-18}) Nps^2m^{-1}	$\left\{\frac{\alpha}{f^2}\right\}_{Th}$ (10^{-21})Nps^2m^{-1}	$\left\{\frac{\alpha}{f^2}\right\}_{long}$ (10^{-18}) Nps^2m^{-1}	$\left\{\frac{\alpha}{f^2}\right\}_{Total}$ (10^{-18}) Nps^2m^{-1}
Mg_2Si	3.155	9.598	2.130	5.295
Mg_2Ge	14.361	14.061	5.071	19.440
Mg_2Sn	101.430	23.521	17.634	119.086

In the present study, it is observed that ultrasonic waves predominantly propagate along the z-axis of the Mg_2X (X = Si, Ge, and Sn) compositions. Equations.17 clearly show that the Akh. Types of energy loses (α/f^2) Akh proportionate to D, E_0, τ, and V^{-3}. E_0' and V' are rising with composition Mg_2X (X=Si, Ge, and Sn), as seen in Figure 9. Therefore, in Mg_2X (X=Si, Ge, and Sn), Consequently, the Akhieser losses in Mg_2X nanomaterials are strongly influenced by the values of E_0 and the thermal conductivity k. A reduction in thermal conductivity leads to a notable increase in ultrasonic attenuation (UA). This suggests that UA is significantly governed by phonon–phonon (p–p) interaction mechanisms. However, due to the lack of sufficient experimental and theoretical data on ultrasonic attenuation in these specific materials, direct comparison and validation remain limited.

As shown in Table 2, the thermo-elastic loss in the nanomaterial composition Mg_2X (X = Si, Ge, and Sn) is significantly lower than the Akhieser loss. According to Equation (18), the primary contributor to total ultrasonic attenuation is Akhieser-type attenuation (UA), which is predominantly governed by phonon–phonon (p–p) interactions for both longitudinal and shear waves. The two key factors influencing total attenuation are thermal energy density and thermal conductivity.

The observed minimum attenuation values suggest that, although Mg_2X compositions are relatively less ductile, they exhibit characteristics of high purity. Additionally, ductility improves progressively across the compositions from Si to Sn. As a result, the compound with the lowest impurity level at room temperature is inferred to be Mg_2Si.

Furthermore, Mg_2Si demonstrates the highest ultrasonic velocities among the compositions, indicating superior ductility. Since attenuation $\alpha \propto V^3$, it is expected to reach its maximum in Mg_2Si. The minimal ultrasonic attenuation (UA) values for Mg_2Si further confirm the structural stability of its hexagonal phase.

4. CONCLUSIONS

Based on the comprehensive analysis of the Mg_2X (X = Si, Ge, and Sn) nanomaterial system, the following conclusions have been drawn:

- The standard computational approach used to determine higher-order elastic coefficients for hexagonal nanomaterials, based on the Lennard-Jones potential model, has proven effective and applicable to the Mg_2X (X = Si, Ge, and Sn) compositions.
- The evaluated elastic properties confirm that the Mg_2X nanomaterials are mechanically stable across all compositions.
- Young's modulus for these nanomaterials demonstrates greater anisotropy compared to the bulk modulus, as indicated by various anisotropy indices, suggesting direction-dependent mechanical behavior.

- The thermal relaxation time (τ) for all Mg_2X compositions is found to be on the order of 10^{-13} seconds, supporting their classification as intermetallic nanomaterials. Notably, the minimum value of τ along the wave propagation direction at 550 K implies the fastest phonon redistribution in this direction.
- The G/B ratios for Mg_2Si, Mg_2Ge, and Mg_2Sn were 0.97, 0.98, and 0.99, respectively, confirming their brittle nature.
- The dominant contribution of p–p interactions to ultrasonic attenuation (U.A.) plays a key role in influencing the thermal conductivity of these nanomaterials.
- Among all compositions, Mg_2Si exhibits superior mechanical properties, with stronger ionic bonding characteristics compared to Mg_2Ge and Mg_2Sn.
- The Mg_2X nanomaterials exhibit their highest purity at the initial Si-based composition and demonstrate enhanced ductility, as evidenced by the lowest ultrasonic attenuation values.

This research could help in computational as well as non-destructive classification of nanomaterials. These findings serve as the basis for further research into the major thermo-physical properties of other nanomaterials.

Acknowledgments

Authors are thankful to Center of Excellence grant from Department of Higher Education, Uttar Pradesh for financial support.

Authors' Contributions

Adwitiya Yadav: Conceptualization, Writing-original draft, Analyzing the results. Prashant Srivastav: Conceptualization, Writing-original draft, Editing, Analyzing the results. Pramod Kumar Yadawa: Review, Visualization, Supervision.

Competing Interests

The authors declare that there are no conflicts of interest related to this work. There are no known competing financial interests or personal relationships that could have influenced the findings presented in this study.

References

[1] V. K. Zaitsev, M.I. Fedorov, I. S. Eremin, and E. A. Gurieva, Thermoelectrics Handbook MacrotoNano, CRCPress, NewYork 2006.

[2] J. Tani, and H. Kido, "Impurity doping into Mg2Sn: A first-principles study," *PhysicaB Condensed Matter*, vol. 407, no. 17, pp. 3493-3498, 2012 https://doi.org/10.1016/j.physb.2012.05.008.

[3] M. R. Kavuncu, M. Ekrem, and N. Yazıcı, "Mechanical Properties and Damage Behavior of MWCNT Reinforced Polyurethane Nanocomposites", El-Cezeri Journal of Science and Engineering, vol. 9, no. 3, pp. 988–995, 2022, doi: 10.31202/ecjse.1018789.

[4] Y. Noda, H. Kon, Y. Furukawa, N. Otsuka, I. A. Nishida, and K. Masumoto, "Temperature dependence of thermoelectric properties of Mg2Si0.6Ge0.4," *Mater. Trans. JIM*, vol. 33, no. 9 pp. 851–855, 1992 https://doi.org/10.2320/matertrans1989.33.851

[5] F. Yu, J. X. Sun, and T. H. Chen, "High-pressure phase transitions of Mg2Ge and Mg2Sn: first-principles calculations," *PhysicaB*, vol 406, no. 9, pp. 1789–1794, 2011 https://doi.org/10.1016/j.physb.2011.02.029

[6] R. Janot, F. Cuevas, M. Latroche, and A. P. Guegan, "Influence of crystallinity on the structural and hydrogenation properties of Mg₂Xphases(X=Ni,Si,Ge,Sn)," *Intermetallics*, vol. 14, no. 2 pp. 163–169, 2006 https://doi.org/10.1016/j.intermet.2005.05.003

[7] J. Prigent, and M. Gupta, "Abinitio study of the hydrogenation properties of Mg-based binary and ternary compounds Mg2X(X=Ni,Si)and YMgNi4," *J.Alloy.Compd.*, vol. 90, pp. 446–447, 2007 https://doi.org/10.1016/j.jallcom.2006.11.104

[8] W. B. Whitten, P. L. Chung, and G.C. Danielson, "Elastic constants and lattice vibration frequencies of Mg2Si," *J.Phys.Chem.Solids*, vol 26, no. 1 pp. 49–56, 1965 https://doi.org/10.1016/0022-3697(65)90071-5

[9] P. L. Chung, W. B. Whitten, G. C. Danielson, "Lattice dynamics of Mg2Ge", *J.Phys. Chem.Solids,* vol. 26, no. 12 pp.1753–1760, *1965* https://doi.org/10.1016/0022-3697(65)90206-4

[10] G. Murtaza, A. Sajidb, M. Rizwan, Y. Makagiwa, H. Khachai, M. Jibran, R. Khenata, and S. Bin Omran, "First principles study of Mg2X(X=Si,Ge,Sn,Pb): elastic, optoelectronic and thermoelectric properties", *Materials Science in Semiconductor Processing* , vol. 40, pp. 429–435, 2015 https://doi.org/10.1016/j.mssp.2015.06.075

[11] J. I. Tani, and H. Kido, "Lattice dynamics of Mg2Si and Mg2Ge compounds from first principles calculations," *Comp.Mater.Sci*, vol. 42, no. 3, pp. 531–536, 2008 https://doi.org/10.1016/j.commatsci.2007.08.018

[12] D. Zhou, J. Liu, S. Xu, and P. Peng, "Thermal stability and elastic properties of Mg2X(X =Si,Ge,Sn,Pb) phases from first-principle calculations", *Comp.Mater.Sci*, vol. 51, no. 1, pp. 409–414, 2012 https://doi.org/10.1016/j.commatsci.2011.07.012

[13] S. Ganeshan, S. L. Shang., Y. Wang, and Z. K. Liu, "Temperature dependent elastic coefficients of Mg2X (X=Si,Ge,Sn,Pb) compounds from first-principles calculations", *J. Alloy.Compd*, vol 498, no. 2, pp. 191–198, 2010 https://doi.org/10.1016/j.jallcom.2010.03.153

[14] S. Ganeshan, S. L. Shang, H. Zhang, Y. Wang, M. Mantina, and Z. K. Liu, "Elastic constants of binary Mg compounds from first-principles calculations", *Intermetallics*, vol. 17, no. 5 pp. 313–318, 2009 https://doi.org/10.1016/j.intermet.2008.11.005

[15] Z. W. Huang, Y. H, Zhao, H. Hou, and P. D, Han, "Electronic structural, elastic properties and thermodynamics of Mg17Al12, Mg2Si and Al2Y phases from first-principles calculations", *PhysicaB,* vol. 407, no. 7 pp. 1075–1081, 2012 https://doi.org/10.1016/j.physb.2011.12.132

[16] M. D. Segall, P. J. D. Lindan, M. J. Probert, C. J. Pickard, P. J. Hasnip, S. J. Clark, and M. C. Payne, "First-principles simulation: ideas, illustrations and the CASTEP code", *Journal of Physics: Condensed Matter*, vol. 14, no. 11 pp. 27-17, 2002 DOI 10.1088/0953-8984/14/11/301

[17] M. Akasaka, T. Iida, T. Nemoto, J. Soga, J. Sato, K. Makino, M. Fukano, and Y. Takanashi, "Non-wetting crystal growth of Mg2Si by vertical Bridgman method and thermoelectric characteristics", *Journal of Crystal Growth*, vol**.** 304, no. 1 pp. 196-201, 2007 https://doi.org/10.1016/j.jcrysgro.2006.10.270.

[18] E. N. Nikitin, V. G. Bazanov, and V. I.Tarasov, *Sov. Phys. Solid State*

[19] A.V. Chernatynskiy, and S. R. Phillpot, "Anharmonic properties in Mg2X (x= C, Si, Ge, Sn, Pb) from first-principles calculations", *Phys. Rev. B*, vol. 92, no. 6, pp. 064303–064309, 2015 https://doi.org/10.1103/PhysRevB.92.064303

[20] T. Caillat, A. Borshchevsky, and J. P. Fleurial, "Properties of single crystalline semiconducting CoSb3", *J. Appl. Phys.*, vol. 80, no. 8, pp. 4442, 1996. https://doi.org/10.1063/1.363405

[21] M. Umemoto, Y. Shirai, and K. Tsuchiya, "Proc. of the 4th Pacific Rim International Conference on Advanced Materials and Processing (PRICM4)", *The Japan Institute of Metals,* pp. 2145, 2001.

[22] J. Tani, H. Kido, "Thermoelectric properties of Bi-doped Mg_2Si semiconductors", *Physica. B*, vol. 364, no. 1-4, pp. 218, 2005. https://doi.org/10.1016/j.physb.2005.04.017

[23] M. Iida, T. Nakamura, K. Fujimoto, Y. Yamaguchi, R. Tamura, T. Iida, and K. Nishio, "Thermoelectric Properties of Mg2Si1-x-yGexSby Prepared by Spark Plasma Sintering", *MRS Advances*, no. 1, pp. 3971–3976, 2016. DOI: https://doi.org/10.1557/adv.2016.332

[24] M. Akasaka, T. Lida, A. Matsumoto, K. Yamanaka, Y. Takanashi, Imai, and N. Hamada, "The Thermoelectric properties of bulk crystalline n- and p-type Mg2Si prepared by the vertical Bridgman method", *Journal of Applied Physics*., vol. 104, no. 1, pp. 13703-13708, 2008, DOI:10.1063/1.2946722

[25] M. Ekrem, "Mechanical Properties of MWCNT Reinforced Polyvinyl Alcohol Nanofiber Mats by Electrospinnig Method", El-Cezeri Journal of Science and Engineering, vol. 4, no. 2, pp. 190–200, 2017, doi: 10.31202/ecjse.305851.

[26] S. Rai, A. K. Prajapati, and P. K Yadawa, "Effect of Pressure on Elastic Constants and Related Properties of Rare-Earth Intermetallic Compound TbNiAl", *Phys Mesomech*., vol. 26, no. 5, pp 495-504, 2023 https://doi.org/10.1134/S1029959923050028.

[27] I. Oral, and M. Ekrem, "Measurement of the elastic properties of epoxy resin/polyvinyl alcohol nanocomposites by ultrasonic wave velocities", *EXPRESS Polymer Letters,* vol. 16, no. 6, pp. 591-606, 2022. DOI: 10.3144/expresspolymlett.2022.44

[28] C. P. Yadav, and D.K. Pandey, "Pressure dependent ultrasonic characterization of nano-structured w-BN", Ultrasonic, vol. 96, pp 181-184, 2019 https://doi.org/10.1016/j.ultras.2019.01.008

[29] D. R. Clarke. "Materials selection guidelines for low thermal conductivity thermal barrier coatings", *Surf. Coat. Technol*., vol. 163, pp. 67-74, 2003 https://doi.org/10.1016/S0257-8972(02)00593-5/

[30] A. K. Prajapati, S. Rai, and P. K. Yadawa, "Theoretical Investigations on Mechanical and Ultrasonic Characteristics of Gallium Nitride Semiconductor under High Pressure", *Emergent mater*, vol. 5, pp 1985-1993, 2022 https://doi.org/10.1007/s42247-022-00419-2

[31] W. Voigt, Lehrbuch der kristallphysik (mitausschluss der kristalloptik), (Leipzig Berlin, B.G. Teubner), 1928.

[32] T. Morelli, and A. G. Slack, "High lattice thermal conductivity solids in high thermal conductivity of materials"*, Goela XVIII*, *Springer Publishe*r., vol. 2, pp. 37–68, 2006

[33] P. K. Yadawa, D. Singh, D. K. Pandey, and R. R. Yadav, "Elastic and Acoustic Properties of Heavy Rare-Earth Metals", *The Open Acoustics Journal*, vol. 2, pp. 61-67, 2009 https://doi.org/10.2174/1874837600902010061.

[34] D. Singh, D. K. Pandey, and P. K. Yadawa, "Ultrasonic wave propagation in rare-earth_monochalcogenides", *centr.eur.j. phys.* , vol. 7, pp. 198-205, 2009 https://doi.org/10.2478/s11534-008-0130-1

[35] R. P. Singh, S. Yadav, D. Singh, G. Mishra, "Theoretical investigation of temperature dependent elastic, thermophysical and ultrasonic properties of Sc-Ti-Zr-Hf quaternary alloy", 2012, vol. 4, no. 1, pp. 33-40. ISSN 2581-8198.

[36] L. Bao, Z. Kong, D. Qu. and Y. Duan, “Revealing the elastic properties and anisotropies of Mg2X (X =Si, Ge and Sn) with different structures from a first-principles calculation”, *Materials Today Communications*, vol. 24, pp. 2352-4928, 2020 https://doi.org/10.1016/j.mtcomm.2020.101337

[37] A. Guechi, A. Merabet, M. Chegaar, A. Bouhemadou, and N. Guechi, “Pressure effect on the structural, elastic, electronic and optical properties of the Zintl phase KAsSn, first principles study”, *J. Alloys Compd.*, vol. 623, pp. 219-228, 2015. https://doi.org/10.1016/j.jallcom.2014.10.114.

[38] S. I. Ranganathan, and M. Ostoja-Starzewski, “Universal Elastic Anisotropy Index.” *Phys. Rev. Lett*., vol. 101, pp. 9007-9008, 2008.https://doi.org/10.1103/PhysRevLett.101.055504.

[39] S. P. Singh, G. Singh, A. K. Verma, P. K. Yadawa, and R. R. Yadav, “Ultrasonic wave propagation in thermoelectric ZrX2 (X= S, Se) compounds”, *Pramana-J. Phys*, vol. 93, pp. 83, 2019. https://doi.org/10.1007/s12043-019-1846-8.

[40] A. K. Jaiswal, P. K. Yadawa, and R. R. Yadav, “Ultrasonic wave propagation in ternary intermetallic CeCuGe compound”, *Ultrasonics*., vol. 89, pp 22-25, 2018. https://doi.org/10.1016/j.ultras.2018.04.009

El Cezeri
Journal of Science and Engineering
ISSN: 2148-3736

Research Article

Investigation Of Heat Transfer and Joule-Thomson Effect in Wells of Depleted Oil and Gas Reservoirs Used for Carbon Dioxide (CO_2) Storage

Okan Kon[1a], İsmail Caner[1b]
[1]Mechanical Engineering Department, Balıkesir University, 10145, Türkiye

ismail@balikesir.edu.tr.

DOI : 10.31202/ecjse.1642591
Received: 18.02.2025 Accepted: 03.08.2025
How to cite this article:
Okan Kon and İsmail Caner, " Investigation Of Heat Transfer and Joule-Thomson Effect in Wells of Depleted Oil and Gas Reservoirs Used for Carbon Dioxide (CO_2) Storage", El-Cezeri Journal of Science and Engineering, Vol: 12, Iss: 3, (2025), pp.(298-310).
ORCID: [a]0000-0002-5166-0258; [b]0000-0003-1232-649X

Abstract This study investigated the heat transfer mechanisms and the Joule-Thomson effect at the wellhead while storing carbon dioxide (CO_2) in depleted oil, gas, and coal reservoirs. It was assumed that the injected CO_2 for storage is in a single-phase pure state. In the reservoir well, convection heat transfer along the wellbore and conduction heat transfer with the surrounding rock soil were analysed during the production of CO_2 to the surface. Additionally, the cooling effect at the wellhead caused by the Joule-Thomson effect was examined. A positive value of the Joule-Thomson coefficient indicated the presence of a cooling effect. For the production well, the study considered temperatures of 30, 51, and 78 °C, pressures of 3.8, 4.3, and 6.1 MPa, and well depths of 1000, 1700, and 2600 meters. Six different rock-soil types surrounding the production well at the reservoir head were included, with a thermal gradient of 25 °C/km and a CO_2 flow velocity of 1 m/s. The calculated difference in conduction and convection heat loss between the wellhead entry and exit ranged from 23.918 to 481.980 W. The Joule-Thomson coefficient was found to vary between 6.797 and 17.91 0C/MPa, depending on the depth, temperature and pressure of the well. The change in exergy efficiency due to the Joule-Thomson effect (throttling exergy) was calculated to vary between 3.042 and 10.766.

Keywords: Carbon dioxide storage, Exergy analysis, Heat transfer, Joule-Thomson effect, Reservoir.

I. INTRODUCTION

The geological storage of CO_2 in reservoirs helps reduce carbon dioxide emissions and is recognized as a feasible approach in the literature. Numerous studies have demonstrated that underground oil and gas reservoirs and coal seams are suitable for the long-term storage of CO_2. Although CO_2 storage capacity is unevenly distributed worldwide, it offers the potential for storage over thousands of years [1]-[2].

The physical properties of CO_2 are crucial in determining the depth and conditions under which storage should occur. Therefore, the different phase states of CO_2 and their effects on underground storage conditions must be studied. At temperatures below the critical temperature and pressures above the vaporization curve, CO_2 exists as a liquid. However, at temperatures and pressures above the critical point, CO_2 is considered a supercritical fluid. Under supercritical conditions, a fluid exhibits properties distinct from those of either the liquid or gas phases. In underground reservoirs, both temperature and pressure increase with depth. While an increase in pressure leads to higher CO_2 density, an increase in temperature has a counteracting effect, reducing the density of CO_2 [1]-[2].

Reservoirs suitable for CO_2 storage must possess specific characteristics. They should be capable of accommodating the desired volume of CO_2, allowing for its injection at a certain rate, and preventing the leakage of light and mobile CO_2 into other subsurface regions. The reservoir must be deeper than 800 meters to achieve sufficient density through pressurisation. CO_2 injected into depths of 800 meters or more becomes a high-density fluid but remains lighter than water. Above this depth, CO_2 exists as a gas with a low density for economically viable storage [1]-[2].

Currently, 18 large-scale commercial carbon capture and storage (CCS) projects are operational, with a total annual CO_2 capture capacity of approximately 40 million tons. To date, 230 million tons of CO_2 have been safely injected underground. This practice is predominantly concentrated in countries like the United States, China, Canada, Australia, and Norway, where oil recovery and coal production are widespread [1]-[2].

A comprehensive review of the literature reveals key insights from various researchers. Abass A. Olajire [3] examined predictive models for gas hydrate formation in gas flow cycles, focusing on management strategies such as chemical injection and pressure reduction to mitigate hydrate risks. He emphasized that these methods help ensure safe production and prevent economic losses. Li et al. [4] investigated CO_2 injection in the Weyburn Midale cap rock, concluding that CO_2 leakage occurs through volume flow and emphasizing the importance of sealing pressure for preventing leakage. They also found that increasing reservoir pressure does not significantly boost storage capacity, but removing water from the reservoir could enhance it. Manovic et al. [5] explored CO_2 storage concepts, outlining criteria for site selection and long-term monitoring as key challenges. Liu et al. [6] discussed waste heat recovery from mature oil and gas reservoirs, proposing a framework for assessing low-temperature recovery potential. Their study on the Villafortuna-Trecate oil field demonstrated that a single well could recover about 25 GWh of electricity over 10 years. Yan et al. [7] analysed deep-borehole heat exchangers for geothermal systems, showing that irregular insulation improves cooling efficiency, achieving up to 21.8% energy savings. Wang et al. [8] proposed a hydro-thermal model for shale oil production, showing that increasing temperature by 60 K could boost production by 31.16% and emphasizing the importance of formation factors. Guo et al. [9] proposed a semi-analytic solution for groundwater flow in geothermal heat pumps, achieving high computational efficiency while maintaining accuracy. Bai et al. [10] addressed well integrity in CO_2 sequestration, focusing on corrosion issues and proposing new risk-based evaluation methods. Amar et al. [11] developed AI models for predicting CO_2 thermal conductivity, outperforming traditional methods. Liu et al. [12] highlighted the potential of closed-loop heat extraction systems in abandoned wells and stressed the need for further optimization. Jia et al. [13] demonstrated that a smooth velocity profile reduces friction in CO_2 flow, suggesting adjustments to minimize friction losses. Zhou et al. [14] examined advanced thermal techniques in abandoned wells, emphasizing renewable integration and carbon neutrality strategies. Kengerli and Agayeva [15] modelled the impact of integrating a new gas pipeline, observing initial instability before system equilibrium. Apak [16] developed an analytical model for drilling fluid temperatures, validated with simulations. Sorgun [17] explored drilling fluids and cutting transport in inclined wells, finding that fluid velocity and pipe rotation significantly affect performance. Ettehadi [18] assessed sepiolite-based mud under challenging conditions, noting temperature effects on viscosity. Ahmed et al. [19] analysed heat exchangers in geothermal systems, showing that well length impacts heat transfer rates and pressure drop.

Holloway [20] explored CO_2 capture and geological storage technologies, focusing on the processes involved in capturing CO_2 at industrial sites, transporting it to storage locations, and ensuring its long-term containment. He identified cost and financing issues related to underground storage as key challenges. This technology is gaining traction due to the continued dependence on fossil fuels for energy production. Bachu [21] proposed 15 criteria for evaluating sedimentary basins for CO_2 sequestration, stressing that warm basins are less suitable due to higher CO_2 rise tendencies and lower storage capacities. His method calculates suitability scores, offering flexibility by adjusting the importance of each criterion, and it was applied to Canadian and Alberta basins. Bahadori and Vuthaluru [22] created a tool to predict CO_2's transport properties, such as viscosity and thermal conductivity, with fewer parameters than existing models. Their tool showed an average deviation of 1.1% for viscosity and 1.3% for thermal conductivity, providing a practical approach for engineers and regulators. Vesovic et al. [23] introduced new CO_2 viscosity and thermal conductivity equations, based on experimental data and theoretical models. These equations are valid for temperatures between 200 K and 1500 K and pressures up to 100 MPa, offering accurate predictions with small uncertainties.

Pipelines, initially constructed in the late 1800s for transporting low-calorific coal gas, are now the primary means of transporting gas. The capacity of pipelines is mainly determined by operating pressure and diameter, with increasing diameters up to 48 inches, which may mark the practical limit for land-based pipelines. Increasing operating pressures is a more common solution to meet growing demand than expanding diameters. Most transmission pipelines operate at pressures above 60 bar, with some reaching 125 bar, requiring larger pipes, bigger compressors, and more stringent safety standards, leading to higher costs [24]. CO_2 is widely used in Enhanced Oil Recovery (EOR), which helps release trapped oil, and CO_2 injection projects produce thousands of barrels daily. CO_2 is also an effective long-term storage method for CO_2 sequestration when available at low costs [25].

A model for high-temperature, high-pressure wells was validated using data from a gas well in China, showing accurate gas pressure and temperature predictions along the well's depth, with implications for dynamic production analysis [26]. Extracting geothermal energy from abandoned oil wells offers an opportunity to extend their life, but technical, economic, and regulatory factors must be addressed. A comprehensive review provides insights into these challenges and the policies for converting these wells into geothermal energy sources [27]. A roadmap for evaluating low-temperature waste heat recovery from mature oil and gas reservoirs was tested in Italy, showing that a single well could generate 25 GWh of electrical energy over a decade [6]. The use of computational fluid dynamics and equations of state to predict the Joule-Thomson (JT) effect in gases was validated with industrial natural gas data, highlighting its importance in industrial applications [28]. Modern drilling often shows variations between well fluid temperatures and the original reservoir temperature, significantly affecting low-permeability reservoirs and high-pressure conditions. An analytical model was developed to predict fluid temperature distribution, accounting for the Joule-Thomson effect, and validated with field data, enabling more accurate predictions of parameters like the productivity index [29]. A study on single-phase oil reservoirs developed a 1D radial flow model to predict fluid temperature changes, showing how production rate impacts reservoir heating and fluid properties. The model is also applicable to single-phase gas reservoirs [30]. Micro-scale thermal effects during ultra-deep gas well production were accounted for in a model predicting bottom-hole temperature changes, showing that factors like gas production rate and initial formation conditions affect flow temperature [31]. CO_2 Plume Geothermal systems outperform traditional geothermal systems at low temperatures. This study explores heat transfer within the wellbore in CO_2-based geothermal operations, suggesting that the wellbore flow can remain adiabatic during much of the system's lifespan [32]. CO_2 storage in depleted gas reservoirs is feasible, providing secure long-term storage, with capacity dependent on various factors [33]. The Joule-Thomson effect's importance in mechanical engineering and its thermodynamic properties were analysed, alongside the throttling effect of gases on pressure, temperature, and flow rate [34]-[35]. Using CO_2 as a heat transfer fluid in geothermal systems shows higher heat extraction rates than water-based systems, allowing more regions to generate geothermal electricity while ensuring CO_2 storage for negative carbon emissions [36].

The primary objective of this study is to analyse the heat transfer mechanisms involved in the storage of CO_2 in depleted hydrocarbon reservoirs at power plants, specifically focusing on the calculation of the Joule-Thomson (JT) coefficient at the wellhead ve well bottom. The research aims to explore various operational conditions, considering three distinct temperatures (30°C, 51°C, and 78°C), three different pressures (3.8 MPa, 4.3 MPa, and 6.1 MPa), and three well depths (1000 m, 1700 m, and 2600 m). Furthermore, the study examines various rock types, including sandy subclay, sand, sandstone, conglomerate, andesite, and gneiss, considering their respective properties, such as specific heat capacity, thermal conductivity, and thermal diffusivity. The physical exergy and exergy efficiency changes of the Joule-Thomson effect (throttling exergy) at the atmospheric outlet were calculated based on the pressure and temperature of CO2 at the wellhead and well bottom. The exergy heat loss was determined for both conduction and convection heat losses at the reservoir outlet.

Joule-Thomson effect (throttling process): It refers to the process in which a gaseous fluid flows from a high-pressure environment to a lower-pressure environment through a narrow opening. This narrow opening can be equipment such as a nozzle, valve, or orifice. During this process, no work is done externally, the pressure of the gas decreases, and its temperature changes. The temperature change may increase or decrease depending on the type of gas. For carbon dioxide in this study, the pressure decreases [35].

In the study, a well thermal gradient of 25°C/km, a CO_2 flow velocity of 1 m/s, and an ambient temperature of 5°C were selected. The rock and soil porosity were set at 0.1, and the well diameter was fixed at 0.100 m. By examining these parameters, the research aims to provide a comprehensive understanding of the thermal behaviour of CO_2 under varying conditions, contributing to the optimization of CO_2 storage processes in depleted reservoirs. The well depths at the reservoir outlet were taken as 1000, 1700, and 2600 meters. These depths were chosen because the pressure and temperature at the well outlet correspond to conditions where carbon dioxide remains in the gas phase. The study investigates the heat transfer mechanisms in the well, energy loss due to heat, and the Joule-Thomson effect (throttling) at the outlet following the storage of large amounts of carbon dioxide in depleted oil, gas, and coal reservoirs. In the existing literature, each of these aspects has been examined separately. However, the distinctive feature of this study is that all of these investigations are conducted together. By performing these combined analyses, the study aims to contribute to the literature.

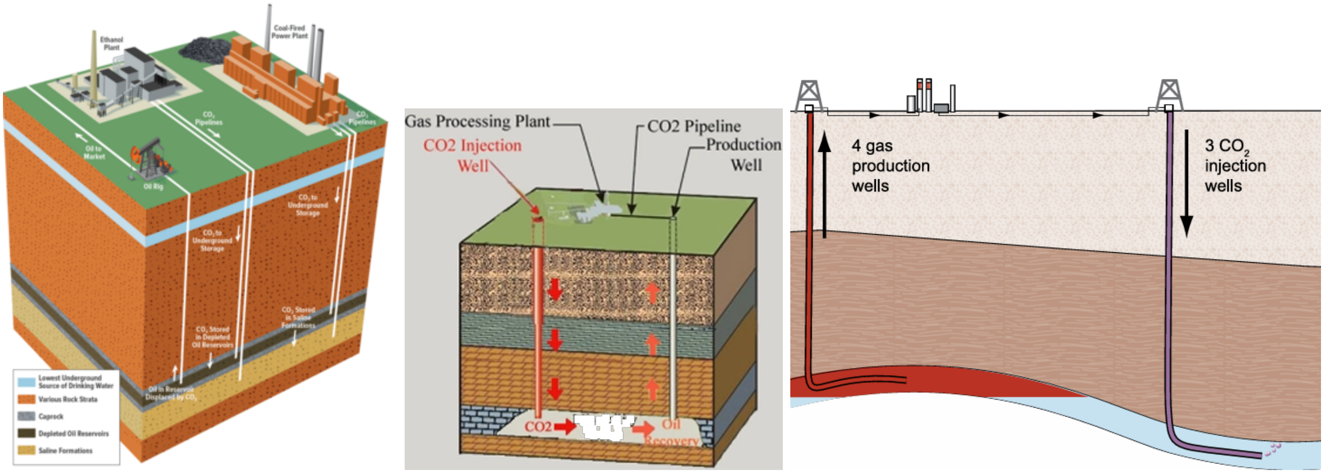


Figure 1. Transportation of carbon dioxide (CO_2) and its storage in reservoirs through wells [1]-[37]-[38].

Table 1. Storage capacity for geological storage [1]

Reservoir Type	Lower estimate of storage capacity ($GtCO_2$)	Upper estimate of storage capacity ($GtCO_2$)
Oil and Gas fields	675	900
Unmineable coal seams	3-15	200
Deep saline formations	1000	possibly 10^4

The mechanisms and equipment for transporting and storing CO_2 from power plants into reservoirs via wells are shown in Figure 1. Figure 2 illustrates suitable saline formations, oil, gas fields, and coal or coal beds where CO_2 can be stored. Table 1 presents the lower and upper storage capacities for geological storage, depending on the reservoir type worldwide.

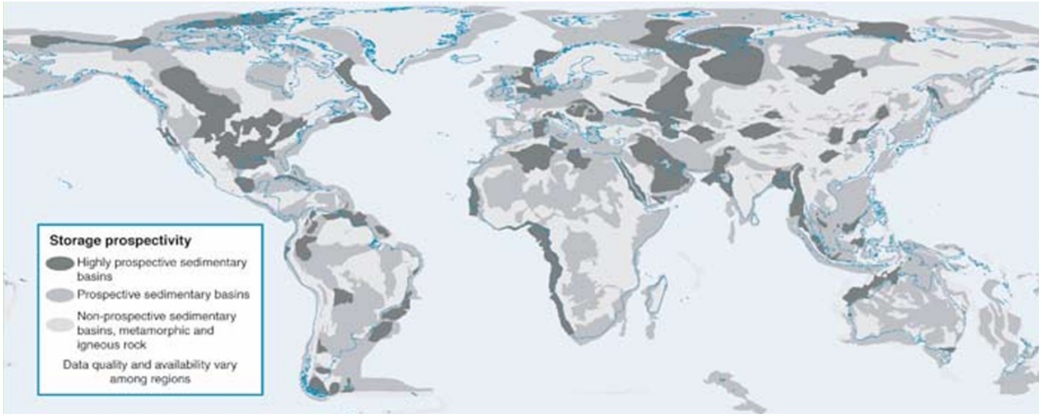


Figure 2. Worldwide Suitable saline formations, oil and gas fields, coal or coal seams where carbon dioxide (CO_2) can be stored [1].

Figure 3 presents the injection and production mechanisms in wells where CO_2 is stored in reservoirs. Figure 4 shows the heat transfer mechanisms in wells where CO_2 is stored. Figure 5 provides the wellhead equipment used for CO_2 storage. Figure 6 depicts the Pressure-Temperature diagram and phase properties for CO_2 storage.

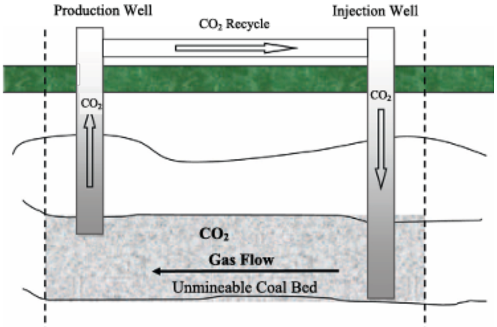


Figure 3. Injection and production well mechanisms for the storage of carbon dioxide (CO_2) [39].

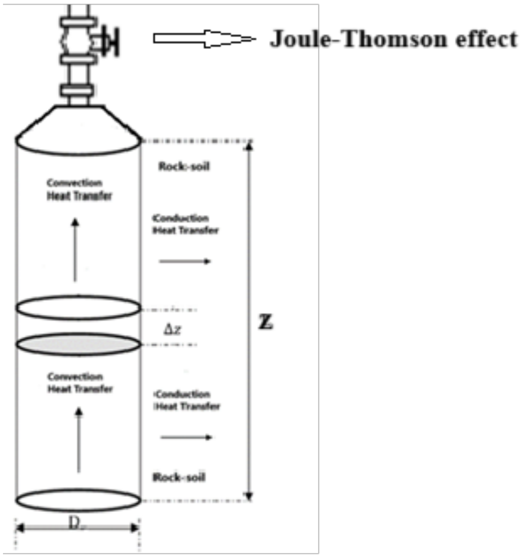


Figure 4. Heat transfer mechanisms in wells where carbon dioxide (CO_2) is stored [32]-[40].

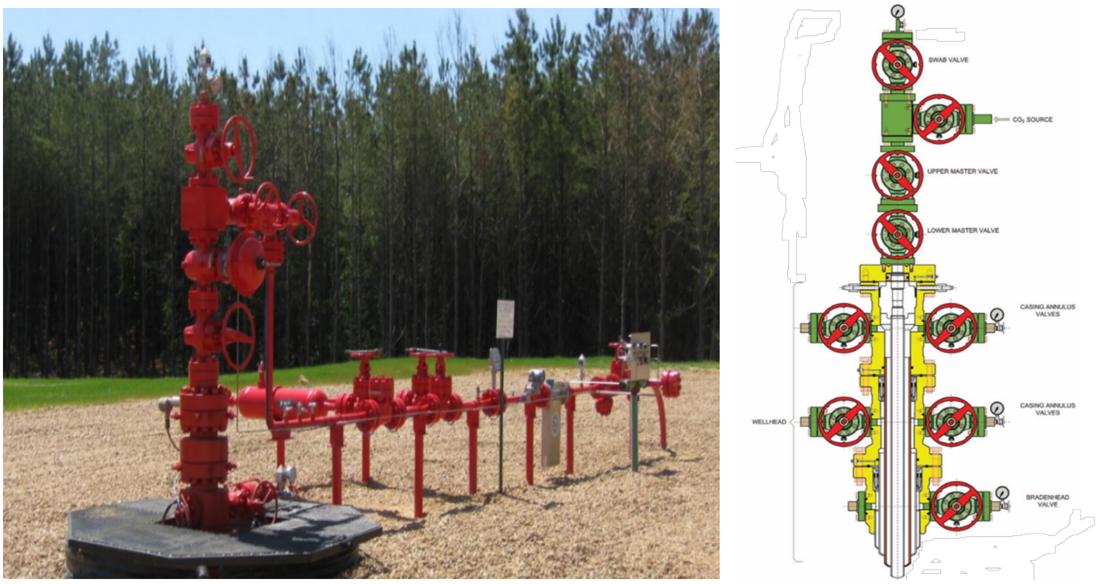


Figure 5. Wellhead equipment for storing carbon dioxide (CO_2) [41]-[42].

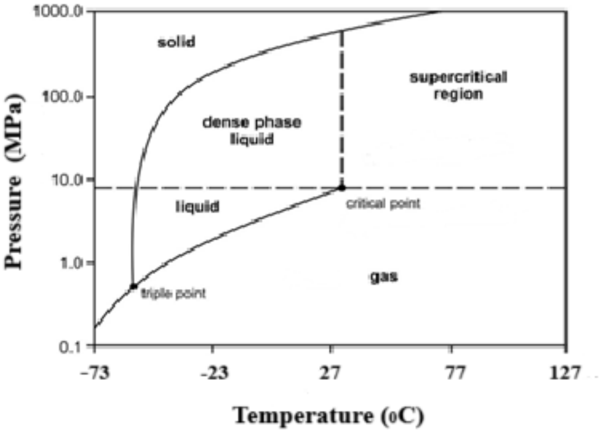


Figure 6. Pressure-Temperature diagram and phase properties for carbon dioxide (CO_2) storage [1]-[43]

II. MATERIALS AND METHOD

The values for Reservoir Temperature, Reservoir Pressure, Ambient Temperature, Thermal Gradient, Flow Velocity, Well Depth, and Well Diameter are provided in Table 2. Table 3 presents the specific heat capacity, thermal conductivity, and thermal diffusivity values for six rock-soil types: Sandy Subclay, Sand, Sandstone, Conglomerate, Andesite, and Gneiss. The critical properties of carbon dioxide, including its critical temperature, critical pressure, and critical density, are shown in Table 4. Equations (1–6) are used to calculate the thermal conductivity and viscosity of carbon dioxide (CO_2). The correlation used to determine the thermal conductivity and viscosity of carbon dioxide is a function of pressure. The four coefficients in the equation (a, b, c, d) depend on the properties of CO_2 within the temperature range of -13 to 177 °C and the pressure range of 10–70 MPa [22]. Equations (7) and (8) are used to determine the Joule-Thomson coefficient at the wellhead based on temperature and pressure differences. The Joule-Thomson coefficient at the reservoir well outlet represents the temperature change per unit pressure change.

$$\ln\lambda = a + \frac{b}{P} + \frac{c}{P^2} + \frac{d}{P^3} \tag{1}$$

$$\ln\mu = a + \frac{b}{P} + \frac{c}{P^2} + \frac{d}{P^3} \tag{2}$$

$$a = A_1 + \frac{B_1}{T} + \frac{C_1}{T^2} + \frac{D_1}{T^3} \tag{3}$$

$$b=A_2+\frac{B_2}{T}+\frac{C_2}{T^2}+\frac{D_2}{T^3} \tag{4}$$

$$c=A_3+\frac{B_3}{T}+\frac{C_3}{T^2}+\frac{D_3}{T^3} \tag{5}$$

$$d=A_4+\frac{B_4}{T}+\frac{C_4}{T^2}+\frac{D_4}{T^3} \tag{6}$$

Table 2. Parameters and Values Used for Calculations in the Study [32]-[44]-[45]-[46].

Parameter	Value
Reservoir Temperature	30, 51, 78 ^{0}C
Reservoir Pressure	3.8, 4.3, 6.1 MPa
Ambient Temperature	5 ^{0}C
Thermal Gradient	25 ^{0}C/km
CO_2 flow velocity	1 m/s
Rock and Soil Porosity	0.1
Well Depth	1000, 1700, 2600 m
Well Diameter	0.100 m
Well Roughness	0.000045 m

Table 3. Rock-Soil Properties and their [47]-[48].

Rock-Soil Thermal Conductivity	Specific Heat Capacity (kJ/kg.K)	Thermal Conductivity (W/m.K)	Thermal Diffusivity (mm2/s)
Sandy Subclay	1.289	1.738	0.681
Sand	1.079	1.989	0.955
Sandstone	0.789	1.967	1.035
Conglomerate	0.795	1.726	0.916
Andesite	0.785	2.084	1.062
Gneiss	0.782	2.553	1.231

Table 4. Critical Properties of Carbon Dioxide (CO2) [1].

Parameter	Value
Molecular Weight	44.01 kg/kmol
Critical Temperature	31 ^{0}C
Critical Pressure	7.39 MPa
Critical Density	467.7 kg/m^3

Table 5. Coefficients for the Calculation of Thermal Conductivity and Viscosity of Carbon Dioxide (CO2) [22]

	Thermal Conductivity (λ_f)	Dynamic Viscosity (μ_f)	
Symbol	Coefficients for thermal conductivity of carbon dioxide	Coefficients for viscosity of carbon dioxide at temperatures less than 67 ^{0}C	Coefficients for viscosity of carbon dioxide, at temperature more than 67 ^{0}C
A_1	2.511772164091	$-8.381727231932328 \times 10^{1}$	$-6.304360942940384 \times 10^{1}$
B_1	$-4.61299395833 \times 10^{3}$	$7.170262916398216 \times 10^{4}$	$7.089412819202834 \times 10^{4}$
C_1	$1.56039978824 \times 10^{6}$	$-2.077352606491789 \times 10^{7}$	$-2.729618206187531 \times 10^{7}$
D_1	$-1.6486817476956 \times 10^{8}$	$2.035238087953347 \times 10^{9}$	$3.491954145885637 \times 10^{9}$
A_2	$-6.7843586693162 \times 10^{2}$	$7.688274861237018 \times 10^{3}$	$5.392507286567643 \times 10^{3}$
B_2	$5.9472900533367 \times 10^{5}$	$-6.832908603727831 \times 10^{6}$	$-6.48675327864201 \times 10^{6}$
C_2	$-1.8136903173662 \times 10^{8}$	$2.00319868619153 \times 10^{9}$	$2.543938513422521 \times 10^{9}$
D_2	$1.8606361834906 \times 10^{10}$	$-1.94536522596535 \times 10^{11}$	$-3.281228975928387 \times 10^{11}$
A_3	$2.0648978571659 \times 10^{4}$	$-1.967260059076993 \times 10^{5}$	$-1.182481836340281 \times 10^{5}$
B_3	$-1.9966713570538 \times 10^{7}$	$1.732142393454871 \times 10^{8}$	$1.438608962538429 \times 10^{8}$
C_3	$6.4236725264301 \times 10^{9}$	$-5.049067845006425 \times 10^{10}$	$-5.738803284656972 \times 10^{10}$
D_3	$-6.8021988393284 \times 10^{11}$	$4.882358762211981 \times 10^{12}$	$7.535042772730154 \times 10^{12}$
A_4	$-1.095035226623 \times 10^{5}$	$1.3529778432466 \times 10^{6}$	$6.947087585578619 \times 10^{6}$
B_4	$1.0878297025052 \times 10^{8}$	$-1.19567721576674 \times 10^{9}$	$-8.506349304338924 \times 10^{9}$
C_4	$-3.575489373317 \times 10^{10}$	$3.498814034450212 \times 10^{11}$	$3.424312685872325 \times 10^{11}$
D_4	$3.8549993712053 \times 10^{12}$	$-3.395109635057981 \times 10^{13}$	$-4.542379235870166 \times 10^{13}$

Equation of Joule-Thomson coefficient [28, 29]-[34]-[35]-[37]-[38]-[55];

$$K_{JT}=\left(\frac{\Delta T}{\Delta P}\right)_{h_{en}}=\left(\frac{T_{in}-T_0}{P_{in}-P_0}\right)_{h_{en}} \tag{7}$$

$$\mu_{JT}=\left(\frac{\partial T}{\partial P}\right)_{h_{en}} \tag{8}$$

The pressure of well Van Der Waals equation can be calculated by using below equations [28, 29, 55, 57-59];

$$P_{in}=\frac{RT}{V-x}-\frac{y}{V^2} \tag{9}$$

$$x=\frac{27}{64}\frac{R^2T_c^2}{P_c^2} \tag{10}$$

$$y=\frac{RT_c}{8P_c} \tag{11}$$

The energy balance related to the heat loss of carbon dioxide (CO_2) inside the production well of a reservoir where CO_2 is geologically stored includes heat loss through conduction from the CO_2 within the well and conductive heat loss between the well and the surrounding rock-soil formations [32], [49], [50], [51], [52], [53]. The coefficients A, B, C, and D used in Equations (3–6) for the calculation of the thermal conductivity and viscosity of carbon dioxide are provided in Table 5.

Heat loss in the well during the discharge of carbon dioxide from the reservoir [33,54-55],

$$q_{storage}=q_{in\text{-}convection}-q_{out\text{-}conduction} \tag{12}$$

The equation for convection heat loss is;

$$q_{convection}=h\,.\,A\,.\,\Delta T \tag{13}$$

Conduction heat transfer equation is,

$$q_{conduction}=\frac{2\pi\,.\,\lambda_{rock\text{-}soil}.z.\Delta T}{\ln\left(\frac{4.Z}{D}\right)} \tag{14}$$

A is the contact area for heat transfer between carbon dioxide (CO_2) and rock-soil. h is the convective heat transfer coefficient, $\lambda_{rock\text{-}soil}$ is the heat conduction coefficient of rock-soil, z is the dept of well, and in (4.Z/D) represents the shape factor for the well. The Nusselt number is used for calculating the convective heat transfer coefficient, where D is the well diameter and λ_f is the thermal conductivity of the fluid (CO_2).

$$h=\frac{Nu\,.\,\lambda_f}{D} \tag{15}$$

For turbulent pipe, flow is used to calculate the Nusselt number (Nu) based on the Reynold number (Re) and Prandtl number (Pr).

$$Nu=0.023\,.\,Re^{0.8}.Pr^{0.4} \tag{16}$$

The Reynold number are,

$$Re=\frac{\rho.V.D}{\mu} \tag{17}$$

where ρ is density, μ is dynamic viscosity and V is the speed of carbon dioxide (CO2) [33,54-55].

General exergy equations is given below. Here T_0 is environmental temperature. T_{in} is inner side of well temperature, h_{en} is enthalpy (kJ/kg) and s_{en} is entropy (kJ/kg.K) [43]-[44]-[55];

$$E_x=h_{en}-h_{en0}-T_0(s_{en}-s_{en0}) \tag{18}$$

The exergy loss equation based on the temperature difference is given below.

$$E_x=q\ .\ \left(1-\frac{T_0}{T_{in}}\right) \tag{19}$$

Joule-Thompson effect on exergy ratio equation [43]-[44]-[55];

$$E_x=T_0\ .\ (s_{en1}-s_{en2}) \tag{20}$$

$$\varepsilon=\frac{\text{Exergy ratio before the Joule-Thompson effect}}{\text{Exergy ratio after the Joule-Thompson effect}} \tag{21}$$

III. RESULTS AND DISCUSSIONS

A Carbon dioxide (CO_2) values for different rock-soil types such as Sandy Subclay, Sand, Sandstone, Conglomerate, Andesite, and Gneiss, the heat loss by conduction for well depths of 1000 m and a well-bottom temperature of 30°C ranges from 73,460 to 108,660, and for 78°C, it ranges from 214,500 to 317,280. For a well depth of 1700 m and a well-bottom temperature of 30°C, the heat loss ranges from 124,880 to 184,720, and for 78°C, it ranges from 364,650 to 539,370. For a well depth of 2600 m and a well-bottom temperature of 30°C, the heat loss ranges from 191,000 to 282,510, and for 78°C, it ranges from 557,700 to 824,920.

The convection heat loss for Carbon dioxide (CO_2) in a well depth of 1000 m and well-bottom temperatures of 30°C and 78°C ranges from 49,542 to 132,140. A well depth of 1700 m and the same well-bottom temperatures ranges from 83,670 to 223,160; for a well depth of 2600 m, it ranges from 128,570 to 342,940.

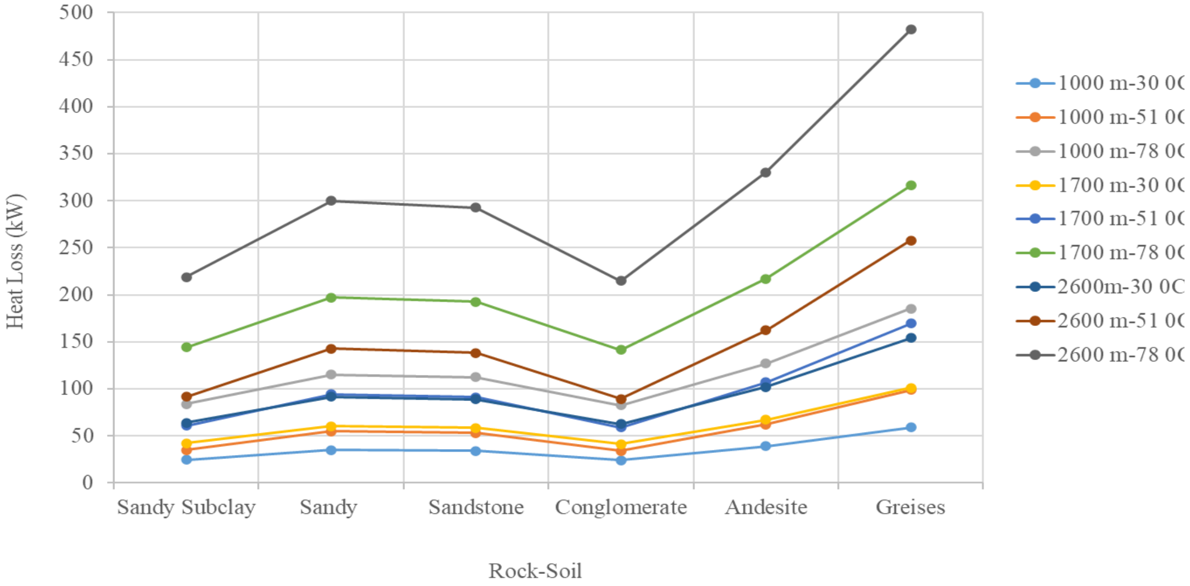


Figure 7. Heat Loss depending on rock-soil type and temperature.

Carbon dioxide (CO_2) values for a well depth of 1000 m and a well-bottom temperature of 30°C, the conduction heat loss ranges from 23,918 to 59,118, and for 78°C, it ranges from 82,360 to 185,140. For a well depth of 1700 m and a well-bottom temperature of 30°C, it ranges from 41,210 to 101,050, and for 78°C, it ranges from 141,490 to 316,210. For a well depth of 2600 m and a well-bottom temperature of 30°C, the conduction heat loss ranges from 62,430 to 153,940, and for 78°C, it ranges from 214,760 to 481,980. A difference between conduction and convection heat loss was detected. For a well-bottom temperature of 51°C, different the heat loss for a well depth of 1000 m ranges from 34,159 to 98,909; for 1700 m, it ranges from 59,190 to 169,290; and for 2600 m, it ranges from 89,280 to 257,660. These and all other values are shown in Figure 7, with heat loss values dependent on rock-soil type, well depth, and well-bottom temperature.

For different rock-soil types such as Sandy Subclay, Sand, Sandstone, Conglomerate, Andesite, and Gneiss, the exergy heat loss for well depths of 1000 m and a well-bottom temperature of 30°C ranges from 1.972 to 4.875, temperature of 51°C ranges from 4.847 to 14.036 and for 78°C, it ranges from 17.122 to 38.488. For a well depth of 1700 m and a well-bottom temperature of 30°C, the exergy heat loss ranges from 3.398 to 8.333, for a temperature of 51°C ranges from 8.400 to 24.024 and for 78°C, it ranges from 29.414 to 65.736. For a well depth of 2600 m and a well-bottom temperature of 30°C, the exergy heat loss ranges from 5.148 to 12.695, and temperature of 51°C ranges from 12.670 to 36.564 for 78°C, it ranges from 44.646 to 100.198. These values are given in Figure 8.

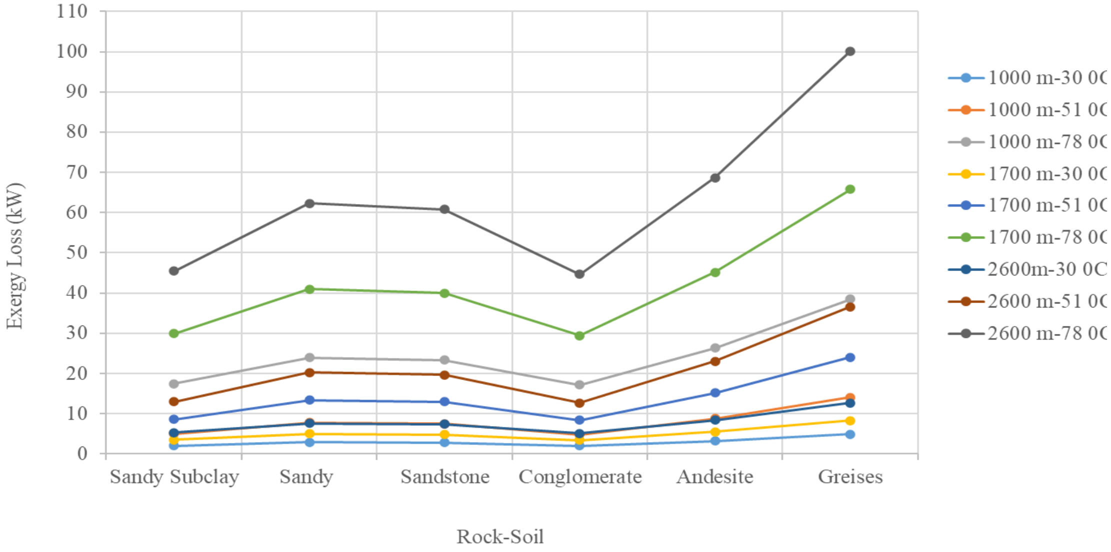


Figure 8. Exergy Loss values depending on rock-soil type and temperature.

The Joule-Thomson coefficient values for carbon dioxide (CO_2) were calculated as 7.006, 7.922, and 17.91^{0}C/MPa for 1000 m at 30 °C, 51 °C, and 78 °C, respectively; 6.866, 7.817, and 17.541 ^{o}C/MPa for 1700 m at 30 °C, 51 °C, and 78 °C, respectively; and 6.797, 7.766, and 17.365 ^{0}C/MPa for 2600 m at 30 °C, 51 °C, and 78 °C, respectively. These values are illustrated in Figure 9. These Joule-Thomson coefficient values are consistent with the literature.

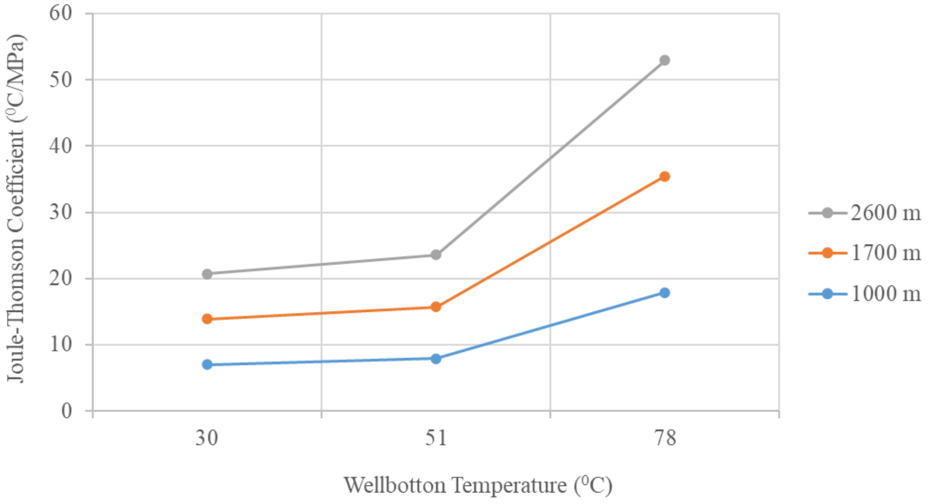


Figure 9. Change in Joule-Thomson Coefficient Depending on Well bottom Temperature.

Carbon dioxide (CO2), initially at atmospheric pressure and a temperature of 5°C at the well bottom, was analysed at various depths. The physical exergy value due to the Joule-Thomson effect (throttling exergy) was calculated as 10.49 at 1000 m depth (3.8 MPa, 30°C), 16.29 at 1700 m depth (4.3 MPa, 51°C), and 25.26 at 2600 m depth (6.1 MPa, 78°C). The throttling exergy associated with CO2 entering the well bottom and exiting the atmosphere was determined to be 581.4 at 1000 m depth (3.8 MPa, 30°C), 527.86 at 1700 m depth (4.3 MPa, 51°C), and 557.5 at 2600 m depth (6.1 MPa, 78°C). The ratio in throttling exergy efficiency, based on the wellhead and well bottom, was calculated as 10.766 at 1000 m depth (3.8 MPa, 30°C), 4.905 at 1700 m depth (4.3 MPa, 51°C), and 3.042 at 2600 m depth (6.1 MPa, 78°C).

IV. CONCLUSIONS

As the well depth and well bottom temperature increase, heat loss through conduction and convection also increases. Depending on the type of rock/soil, well depth, and well bottom temperature, heat loss through conduction was calculated to range from 73.460 to 824.920. Heat loss through convection ranged from 49.542 to 342.940. The total heat loss was determined to be between 123.002 and 1167.860. The difference in heat loss between the well inlet and outlet ranged from 23.918 to 481.980 for conduction and convection.

The Joule-Thomson coefficient decreases as the well depth increases but increases with rising well bottom temperature. Depending on the well depth and temperature, the Joule-Thomson coefficient was found to vary between 6.797 and 17.91.

The change in exergy efficiency due to the Joule-Thomson effect (throttling exergy) was calculated to vary between 3.042 and 10.766, depending on the well depth, pressure, and temperature. Similarly, the throttling exergy was determined to range from 557.5 to 581.4 under varying depth, pressure, and temperature conditions.

Future studies in this field could focus on optimizing heat loss management techniques by incorporating advanced materials and innovative insulation technologies tailored to various rock and soil types. Additionally, research could explore the relationship between well depth (3000 and 4000 m), temperature (80-100 ^{0}C) pressure (7-10 MPa), and the Joule-Thomson coefficient to develop predictive models that enhance geothermal and oil extraction efficiency. As the industry moves toward sustainable practices, integrating renewable energy solutions and minimizing environmental impact will likely become key focus areas.

Authors' Contributions

All authors contributed to the conception and design. Ismail Caner performed conceptualization, funding acquisition, writing review, and editing. Okan Kon conducted conceptualization, methodology, supervision, and writing-original draft. All authors read and approved the final manuscript.

Competing Interests

The authors declare that they have no relevant financial or other interests.

References

[1] Metz, B., Davidson, O., De Coninck, H. C., Loos, M., & Meyer, L. (2005). IPCC special report on carbon dioxide capture and storage. Cambridge: Cambridge University Press.

[2] Rifat US, Çağlar S, Elif K, Türkiye'nin Karbon Yakalama, Kullanma ve Depolama Potansiyeli, Kaynak, Çevre ve İklim Derneği – REC, Mart 2024, Ankara

[3] Olajire, A. A. (2020). Flow assurance issues in deep-water gas well testing and mitigation strategies with respect to gas hydrates deposition in flowlines—A review. Journal of molecular liquids, 318, 114203. https://doi.org/10.1016/j.molliq.2020.114203

[4] Li, Z., Dong, M., Li, S., & Huang, S. (2006). CO2 sequestration in depleted oil and gas reservoirs—caprock characterization and storage capacity. Energy conversion and management, 47(11-12), 1372-1382. https://doi.org/10.1016/j.enconman.2005.08.023

[5] Aminu, M. D., Nabavi, S. A., Rochelle, C. A., & Manovic, V. (2017). A review of developments in carbon dioxide storage. Applied Energy, 208, 1389-1419. https://doi.org/10.1016/j.apenergy.2017.09.015

[6] Liu, X., Falcone, G., & Alimonti, C. (2018). A systematic study of harnessing low-temperature geothermal energy from oil and gas reservoirs. Energy, 142, 346-355. https://doi.org/10.1016/j.energy.2017.10.058

[7] Luo, Y., Xu, G., & Yan, T. (2020). Performance evaluation and optimization design of deep ground source heat pump with non-uniform internal insulation based on analytical solutions. Energy and Buildings, 229, 110495. https://doi.org/10.1016/j.enbuild.2020.110495

[8] Wang, Z., Fan, W., Sun, H., Yao, J., Zhu, G., Zhang, L., & Yang, Y. (2020). Multiscale flow simulation of shale oil considering hydro-thermal process. Applied Thermal Engineering, 177, 115428. https://doi.org/10.1016/j.applthermaleng.2020.115428

[9] Guo, Y., Zhao, J., & Liu, W. V. (2024). Effects of varying heat transfer rates for borehole heat exchangers in layered subsurface with groundwater flow. Applied Thermal Engineering, 247, 123007. https://doi.org/10.1016/j.applthermaleng.2024.123007

[10] Bai, M., Zhang, Z., & Fu, X. (2016). A review on well integrity issues for CO2 geological storage and enhanced gas recovery. Renewable and Sustainable Energy Reviews, 59, 920-926. https://doi.org/10.1016/j.rser.2016.01.043

[11] Amar, M. N., Ghahfarokhi, A. J., & Zeraibi, N. (2020). Predicting thermal conductivity of carbon dioxide using group of data-driven models. Journal of the Taiwan Institute of Chemical Engineers, 113, 165-177. https://doi.org/10.1016/j.jtice.2020.08.001

[12] Liu, Z., Yang, W., Xu, K., Zhang, Q., Yan, L., Li, B., ... & Yang, M. (2023). Research progress of technologies and numerical simulations in exploiting geothermal energy from abandoned wells: a review. Geoenergy Science and Engineering, 224, 211624. https://doi.org/10.1016/j.geoen.2023.211624

[13] Jia, M., Deng, S., Li, X., Jin, W., Yang, Z., & Rao, D. (2023). A numerical simulation study of the micro-mechanism of CO2 flow friction in fracturing pipe string. Gas Science and Engineering, 112, 204941. https://doi.org/10.1016/j.jgsce.2023.204941

[14] Zhou, Y., Liu, Z., & Xing, C. (2022). Application of abandoned wells integrated with renewables. In Utilization of Thermal Potential of Abandoned Wells (pp. 255-273). Academic Press. https://doi.org/10.1016/B978-0-323-90616-6.00013-0

[15] Kengerli, T. S., & Agayeva, N. A. (2024). Influence of connecting a new gas pipeline to the operating gas pipeline on the flow rate of production wells. https://doi.org/10.53404/Sci.Petro.20240100053

[16] Apak, E. C. (2006). A study on heat transfer inside the wellbore during drilling operations. Master's thesis, Middle East Technical University.

[17] Sorgun, M. (2010). Modeling of Newtonian fluids and cuttings transport analysis in high inclination wellbores with pipe rotation.

[18] Ettehadi, A. (2016). Modelling Wellbore Hydraulics through Thermal Rheological Sepiolite Mud Properties. PhD thesis, Graduate School of Science, Engineering and Technology, Istanbul Technical University.

[19] Ahmed, N., Kabir, E., & Islam, M. A. (2024, April). The influence of borehole lengths on a numerical model of a double-tube vertical ground heat exchanger. In IOP Conference Series: Materials Science and Engineering (Vol. 1305, No. 1, p. 012002). IOP Publishing. https://doi.org/10.1088/1757-899X/1305/1/012002

[20] Holloway, S. (2007). Carbon dioxide capture and geological storage. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365(1853), 1095-1107. https://doi.org/10.1098/rsta.2006.1953

[21] Bachu, S. (2003). Screening and ranking of sedimentary basins for sequestration of CO2 in geological media in response to climate change. Environmental Geology, 44(3), 277-289. https://doi.org/10.1007/s00254-003-0762-9

[22] Bahadori, A., & Vuthaluru, H. B. (2010). Predictive tool for an accurate estimation of carbon dioxide transport properties. International Journal of Greenhouse Gas Control, 4(3), 532-536. https://doi.org/10.1016/j.ijggc.2009.12.007

[23] Vesovic, V., Wakeham, W. A., Olchowy, G. A., Sengers, J. V., Watson, J. T. R., & Millat, J. (1990). The transport properties of carbon dioxide. Journal of physical and chemical reference data, 19(3), 763-808. https://doi.org/10.1063/1.555875

[24] Lim, O.S. Optimization of Gas Transmission Design. 2011. Available online: http://utpedia.utp.edu.my/id/eprint/10565/1/20 11%20-%20Optimizing%20of%20gas%20transmission%20design.pdf (accessed on 22 Jan 2025).

[25] Sagir, M., Mushtaq, M., Tahir, M. S., Tahir, M. B., Ullah, S., Abbas, N., & Pervaiz, M. (2018). CO2 capture, storage, and enhanced oil recovery applications. Encyclopedia of Renewable and Sustainable Materials. Elsevier, 52-58. https://doi.org/10.1016/B978-0-12-803581-8.10360-1

[26] Wu, Z., Xu, J., Wang, X., Chen, K., Li, X., & Zhao, X. (2011). Predicting temperature and pressure in high-temperature–high-pressure gas wells. Petroleum Science and Technology, 29(2), 132-148. https://doi.org/10.1080/10916460903330213

[27] Kurnia, J. C., Shatri, M. S., Putra, Z. A., Zaini, J., Caesarendra, W., & Sasmito, A. P. (2022). Geothermal energy extraction using abandoned oil and gas wells: Techno-economic and policy review. International Journal of Energy Research, 46(1), 28-60. https://doi.org/10.1002/er.6386

[28] Shoghl, S. N., Naderifar, A., Farhadi, F., & Pazuki, G. (2020). Prediction of Joule-Thomson coefficient and inversion curve for natural gas and its components using CFD modeling. Journal of Natural Gas Science and Engineering, 83, 103570. https://doi.org/10.1016/j.jngse.2020.103570

[29] Chevarunotai, N. (2014). Analytical Models for Flowing-Fluid Temperature Distribution in Single-Phase Oil Reservoirs Accounting for Joule-Thomson Effect. PhD thesis.

[30] Islam, R. (2017). Analytical Model for Fluid Temperature Change During Expansion in the Reservoir. PhD thesis.

[31] Ding, L., Yang, Z., Chen, W., & Zhang, Q. (2023). Transient prediction method for flow temperature at wellbore bottom. Applied Thermal Engineering, 234, 121208. https://doi.org/10.1016/j.applthermaleng.2023.121208

[32] Randolph, J. B., Adams, B., Kuehn, T. H., & Saar, M. O. (2012). Wellbore heat transfer in CO2-based geothermal systems. Geothermal Resources Council Transactions, 36, 549-554.
[33] Van Der Meer, B. (2005). Carbon dioxide storage in natural gas reservoir. Oil & gas science and technology, 60(3), 527-536. https://doi.org/10.2516/ogst:2005035
[34] Kon, O. (2004). Termodinamik kısılma ve madde özelliklerine etkisi. Master's thesis, Balıkesir Üniversitesi Fen Bilimleri Enstitüsü.
[35] Kon, O. (2009). Termodinamik kısılma olayında Joule-Thomson katsayısı ve inversiyon eğrileri. Balıkesir Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 11(2), 81-93.
[36] Randolph, J. B., & Saar, M. O. (2011). Combining geothermal energy capture with geologic carbon dioxide sequestration. Geophysical Research Letters, 38(10). https://doi.org/10.1029/2011GL047265
[37] Phillip L. Swagel, Carbon Capture and Storage in the United State, December 2023.
[38] Energy, Technology, & Policy, Carbon Dioxide Enhanced Oil Recovery: A Great Environmental Choice, https://webberenergyblog.wordpress.com/2012/04/01/carbon-dioxide-enhanced-oil-recovery-a-great-environmental-choice/ (access: 12.27.2024)
[39] Huang, Y., Zheng, Q. P., Fan, N., & Aminian, K. (2014). Optimal scheduling for enhanced coal bed methane production through CO2 injection. Applied energy, 113, 1475-1483. http://dx.doi.org/10.1016/j.apenergy.2013.08.074
[40] [40] Samuel, R. J. (2019). Transient flow modelling of carbon dioxide (CO2) injection into depleted gas fields. PhD thesis. University College London.
[41] Gaurina-Međimurec, N., & Pašić, B. (2011). Design and mechanical integrity of CO2 injection wells. Rudarsko-Geolosko-Naftni Zbornik, 23.
[42] CO2 injection well in Mississippi, https://www.usgs.gov/media/images/co2-injection-well-mississippi, April 9, 2010 (access: 12.26.2024)
[43] Witkowski, A., Majkut, M., & Rulik, S. (2014). Analysis of pipeline transportation systems for carbon dioxide sequestration. Archives of thermodynamics, 35(1), 117-140. http://dx.doi.org/10.2478/aoter-2014-0008
[44] Lu, T., Li, Z., Fan, W., & Li, S. (2016). CO2 huff and puff for heavy oil recovery after primary production. Greenhouse Gases: Science and Technology, 6(2), 288-301. https://doi.org/10.1002/ghg.1566
[45] Longe, P. O., Danso, D. K., Gyamfi, G., Tsau, J. S., Alhajeri, M. M., Rasoulzadeh, M., ... & Barati, R. G. (2024). Predicting CO2 and H2 Solubility in Pure Water and Various Aqueous Systems: Implication for CO2–EOR, Carbon Capture and Sequestration, Natural Hydrogen Production and Underground Hydrogen Storage. Energies, 17(22), 5723. https:// doi.org/10.3390/en17225723
[46] Chow, Y. F., Maitland, G. C., & Trusler, J. M. (2016). Interfacial tensions of the (CO2+ N2+ H2O) system at temperatures of (298 to 448) K and pressures up to 40 MPa. The Journal of Chemical Thermodynamics, 93, 392-403.http://dx.doi.org/10.1016/j.jct.2015.08.006
[47] Wang, R., Shi, M., Zhu, K., Yu, J., Ren, W., Yan, G., ... & Gao, S. (2024). Research on the heat transfer model of double U-pipe ground heat exchanger based on in-situ testing. Frontiers in Energy Research, 12, 1442185. https://doi.org/10.3389/fenrg.2024.1442185
[48] Fuchs, S., Balling, N., & Förster, A. (2015). Calculation of thermal conductivity, thermal diffusivity and specific heat capacity of sedimentary rocks using petrophysical well logs. Geophysical Journal International, 203(3), 1977-2000. https://doi.org/10.1093/gji/ggv403
[49] Mobaraki, H. (2024). The Cooling Effect of Joule Thomson on CO2 Storage during Start-Up Period: Integration of Reservoir and Wellbore. PhD thesis, Politecnico di Torino.
[50] Shoghl, S. N., Naderifar, A., Farhadi, F., & Pazuki, G. (2021). Thermodynamic analysis and process optimization of a natural gas liquid recovery unit based on the Joule–Thomson process. Journal of Natural Gas Science and Engineering, 96, 104265. https://doi.org/10.1016/j.jngse.2021.104265
[51] Çengel, Y. A., & Ghajar A.F. (2015). Isı ve Kütle Transferi, Palme Yayınevi.
[52] Kayansayan N., Thermodynamics Principle & Applications, Nobel Akademik Yayıncılık, 2013.
[53] Span, R., & Wagner, W. (1996). A new equation of state for carbon dioxide covering the fluid region from the triple-point temperature to 1100 K at pressures up to 800 MPa. Journal of physical and chemical reference data, 25(6), 1509-1596. https://doi.org/10.1063/1.555991
[54] Menon, E.S. (2005). Gas Pipeline Hydraulics (1st ed.). CRC Press. https://doi.org/10.1201/9781420038224
[55] Luo, Y., & Wang, X. (2010). Exergy analysis on throttle reduction efficiency based on real gas equations. Energy, 35(1), 181-187. https://doi.org/10.1016/j.energy.2009.09.008
[56] Kotas, T. J. (2012). The Exergy Method of Thermal Plant Analysis, 1985. Great Britain by Anchor Brendon Ltd, Tiptree, Essex.
[57] Li, J., Chen, Y., Ma, Y. B., Kwon, J., & Xu, H. (2023). Ji-Chao Li A study on the Joule-Thomson effect of during filling hydrogen in high pressure tank, Case Studies in Thermal Engineering, 41, 102678. https://doi.org/10.1016/j.csite.2022.102678

[58] Erdoğan, M., & Acar, M. Ş. (2024). Thermodynamic analysis of a tunnel biscuit oven and heat recovery system. *WAPRIME*, *1*(1), 1-15.
[59] Ünal, E. K. (2024). Analytical and numerical investigation of viscous heating in parallel-plate Couette flow. WAPRIME, 1(1), 57-69.

El Cezeri
Journal of Science and Engineering
ISSN: 2148-3736

Research Article

Accelerating Convergence in LMS Adaptive Filters Using Particle Swarm Optimization: A Hybrid Approach for Real-Time Signal Processing

Muhammed Davud[1a]
[1]Department of Software Engineering, Istinye University, 34010, Istanbul, Türkiye

muhammed.davud@istinye.edu.tr

DOI : 10.31202/ecjse.1598491
Received: 09.12.2024 Accepted: 30.05.2025
How to cite this article:
Muhammed Davud, " Accelerating Convergence in LMS Adaptive Filters Using Particle Swarm Optimization: A Hybrid Approach for Real-Time Signal Processing", El-Cezeri Journal of Science and Engineering, Vol: 12, Iss: 3, (2025), pp.(311-328).
ORCID: [a]0000-0002-6864-2339.

Abstract Adaptive filtering is essential for control systems, system identification, and noise cancellation, especially when using Least Mean Square (LMS) algorithms. Although LMS-based approaches are popular because they are straightforward and efficient, they frequently have sluggish convergence and numerical instability. This study offers a hybrid framework that integrates Particle Swarm Optimization (PSO) with LMS variations, such as ZA-LLMS, RZA-LLMS, ZA-VSS-LMS, and RZA-VSS-LMS, in order to overcome these issues. PSO outperforms conventional methods in terms of mean square error (MSE) performance and convergence speed by dynamically modifying weight coefficients. The proposed system was evaluated utilizing synthetic noise models, such as Additive White Gaussian Noise (AWGN) and Colored Gaussian Sequence (CGS), in addition to MRI scan restoration. The results show that the PSO-enhanced LMS versions reduce the number of required iterations by up to 67% while improving filtering accuracy. Under Gaussian noise, the PSO-RZA-VSS-LMS approach remarkably obtained a Peak Signal-to-Noise Ratio (PSNR) of 24.58 dB while its non-PSO equivalent only obtained 20.51 dB. In a similar line, PSO-RZA-LLMS attained 17.12 dB for Salt & Pepper noise, at 15.32 dB, surpassing the baseline RZA-LLMS. These results show the flexibility of the PSO-driven approach toward different noise distributions. Apart from raising filtering accuracy, PSO integration significantly accelerates convergence without sacrificing signal integrity. The results show that the proposed method offers a practical and efficient replacement for real-time adaptive filtering systems including medical imaging, speech processing, and high-speed communication networks.

Keywords: Adaptive filtering, Hybrid Optimization, LMS, Particle Swarm, Image Denoising.

I. INTRODUCTION

With a wide range of applications, adaptive filtering (AF) is a well-established and extensively researched topic [1]–[4]. The fundamental component of an adaptive filter is a linear filter with movable parameters that are updated dynamically through optimization. Because of its adaptability, AF can be set up in a variety of ways to meet the needs and demands of diverse applications. The Finite Impulse Response (FIR) filter is a popular setup in which a set of tunable coefficients determines the output by means of a weighted sum of the input samples [5]. The Infinite Impulse Response (IIR) filter is another method that combines feedback into its construction. With fewer coefficients than FIR filters, IIR filters can computationally be more efficient while nevertheless attaining comparable filtering characteristics [6]. This efficiency makes them especially valuable in real-time applications and situations with limited memory capacity. Apart from these basic setups, various other adaptive filtering methods provide special signal-processing capacity. One type of block adaptive filter divides the input signal into blocks and independently updates the filter coefficients for each block. This method allows more exact adaptation and tracking of changes, hence it is particularly useful for managing signals whose attributes change with time [7]. Another fascinating configuration, whereby a filter bank divides the input signal into many frequency subbands, is the subband adaptive filter. When different frequency components require distinct filtering techniques in applications like audio coding and noise cancellation, this approach is very helpful since an adaptive filter handles each subband individually [8]. Still another sometimes used model is the adaptive Radial Basis Function (RBF). Particularly effective for nonlinear and non-Gaussian signals, it bases the output on an array of RBFs. Common uses of this approach in systems identification, pattern recognition, and time-series forecasting [9]. Designed particularly to remove undesirable noise from a target signal, the Adaptive Noise Canceller (ANC) is a specific use of adaptive filtering [10]. The ANC uses two inputs: the main input has the intended signal mixed with noise, while the reference input has a noise component linked with the noise in the main input. An adaptive filter then continuously adjusts to lower the correlation between the projected noise and the primary input. By eliminating this estimated noise from the main input, the ANC can dynamically adapt to match changing noise levels, thereby enhancing signal clarity and intelligibility. Usually, the convergence speed and the ability to obtain the lowest Mean Square Error (MSE) define the performance of an AF method. One of the fundamental strategies for optimizing AF performance is the Wiener filter, which assumes prior knowledge of the input signal [11]. Nonetheless, some adaptive filtering techniques, such as Least Mean Square (LMS), are often applied despite not requiring this prior knowledge because of their simplicity and ease of implementation [12]. The typical LMS technique has limits notwithstanding these benefits. When the input sequence lacks enough excitation, it can suffer from unbounded parameter estimations, which cause numerical instability and poor performance due to excessively high prediction errors [13]. Several variants of the LMS method have been proposed to handle these problems. The Leaky LMS (LLMS) model is one such development, since it reduces parameter drift, thereby stabilizing the system [14]. Proposed to improve sparsity control and reduce update complexity, the Zero-Attracting LLMS (ZA-LLMS) method is a further development [15]. Convergence speed and steady-state performance studies reveal that ZA-LLMS outperforms both conventional LLMS and ZA-LMS algorithms [16]. The Reweighted Zero-Attracting LLMS (RZA-LLMS) algorithm, introduced in [17], shows even greater performance in adaptive filtering applications, further enhancing the efficiency of ZA-LLMS.

In a different track, it is recognized that the Mean Square Error (MSE) is directly related to the adaptation step size in terms of performance, and that it determines the convergence rate to optimal weights. Usually, a small step size is recommended to attain a low final excess MSE, but this results in sluggish convergence, which is, in many circumstances, unsatisfactory. Variable Step Size (VSS) adaptation is one method for achieving greater performance than the LMS algorithm, which uses a constant fixed step size. In the Variable Step Size LMS (VSS-LMS) method, the step size is large when the LMS weights are suboptimal, and it decreases as the weights approach optimality. For changing the step size, several criteria are available, including the squared estimation error [18], the gradient of the squared estimation error with respect to the step [19], and the autocorrelation of subsequent estimation errors [20]. In [21], the authors introduced the Zero-Attracting VSS-LMS (ZA-VSS-LMS) algorithm, which outperforms the standard VSS-LMS algorithm. To minimize the update formula, they added the ℓ_1 norm of the filter coefficient vector in the ZA-VSS-LMS cost function, particularly when the majority of coefficients are zero. Nevertheless, the ZA-VSS-LMS algorithm's shrinkage does not distinguish between zero taps and non-zero taps. In [21], the Reweighted Zero-Attracting VSS-LMS (RZA-VSS-LMS) method is presented to address this issue.

Although the LMS method was utilized to generate a reference current because of its ease of implementation, there was a trade-off between its convergence rate and minimum MSE [22]. In the past decade, convex AF combinations have emerged as a way to avoid this trade-off [23]–[25]. The combination of convex AFs has improved both stationary and non-stationary performance [25]. Typically, two AFs are used to construct the convex combination: one provides rapid convergence with a high steady-state output error, while the other has slow convergence and a small MSE [26]. This strategy offers a solution to the trade-off problem by merging the faster-convergent and lower-MSE features of the two filtering techniques [25]. Particle Swarm Optimization (PSO) [27] is a swarm intelligence-based metaheuristic optimization method that can be applied in adaptive filters (AFs). To identify the best solution inside a search space, the PSO technique has been extensively used in several disciplines, including signal processing, control systems, and machine learning [28], [29]. In signal processing, PSO has been applied to create adaptive filters for addressing noise reduction concerns [30]. Furthermore, the integration of LMS with PSO has been suggested in [31] to improve its performance.

This work presents hybrid Adaptive Noise Canceller (ANC) algorithms using the Particle Swarm Optimization (PSO) approach to improve adaptive filtering performance. To enhance adaptive filtering performance, we propose the PSO-ZA-LLMS, PSO-RZA-LLMS, PSO-ZA-VSS-LMS, and PSO-RZA-VSS-LMS adaptive filtering techniques. While ZA-LLMS, RZA-LLMS, ZA-VSS-LMS, and RZA-VSS-LMS are used to address the local convergence issue that may result from employing PSO, these algorithms utilize PSO during the iterative process to seek the best solution. In terms of both convergence speed and performance, experimental results demonstrate that the proposed systems, PSO-RZA-LLMS and PSO-RZA-VSS-LMS, outperform the conventional RZA-LLMS and RZA-VSS-LMS filters.

The paper is organized as follows. Section II provides a review of different models of LMS adaptive filters, including LMS, LLMS, ZA-LLMS, RZA-LLMS, VSS-LMS, ZA-VSS-LMS, and RZA-VSS-LMS algorithms. In Section III, Particle Swarm Optimization (PSO) is reviewed as an optimization technique that can be used in conjunction with adaptive filters. Section IV presents the proposed method, which is a combination of PSO and the weighted zero-attracting VSS-LMS (RZA-VSS-LMS) algorithm. The simulation results and discussion are presented in Section V, where the performance of the proposed method is compared with other traditional adaptive filter algorithms. The application of the proposed algorithms to 2D images is explored in Section VI. Section VII provides an in-depth analysis of the proposed algorithms, evaluating their effectiveness and limitations while discussing potential future research directions. Finally, in Section VIII, the paper concludes by summarizing the main findings and contributions of the proposed method, and by highlighting potential directions for future research in this area.

II. VARIOUS MODELS OF LMS

This paper uses uppercase boldface letters to show matrices, while lowercase boldface letters are used to show vectors. The transpose operator is shown as $(\cdot)^T$, and the inverse operator is shown as $(\cdot)^{-1}$. The operator $\| . \|_1$ stands for the $\ell 1$ norm, and $tr(\cdot)$ stands for the trace operator. Also, $E[\cdot]$ is the symbol for the expectation operator. The basic ANC filter considered in this paper is depicted in Figure 1. The input signal that serves as the primary input to the ANC system is the noisy signal $d(\omega)$, which is depicted as:

$$d(\omega) = s(\omega) + v(\omega), \tag{1}$$

where the components of the input signal $d(\omega)$ are the noise-free signal $s(\omega)$ and the added noise signal $v(\omega)$, which are assumed to be uncorrelated in time. The reference input signal $x(\omega)$ provided to the adaptive filter is a noise signal that is correlated with $v(\omega)$. By processing $x(\omega)$, the adaptive filter calculates an estimation of the noise signal as:

$$y(\omega) = \theta^T(\omega)x(\omega), \tag{2}$$

where $\theta = [z_0, z_1, \cdots, z_{N-1}]^T$ is the filter coefficient vector with length N taps, $\mathbf{x}(\omega) = [x(\omega), x(\omega-1), \cdots, x(\omega-N+1)]^T$. Therefore, the error signal is calculated as:

$$e(\omega) = d(\omega) - y(\omega). \tag{3}$$

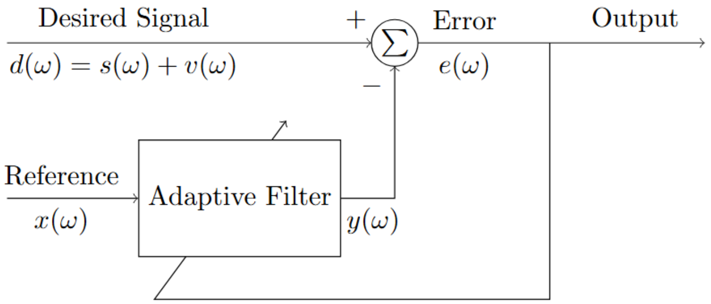


Figure 1. Block diagram of the basic Adaptive Noise Canceller (ANC).

A. LEAST MEAN SQUARE (LMS)

The objective of LMS-type filters is to gradually estimate the unknown coefficient vector using the reference signal $x(\omega)$ and the target signal $d(\omega)$. The estimated coefficient vector of the adaptive filter is indicated as $\theta(\omega)$ for each iteration. The typical LMS defines the cost function $L_1(\omega)$ as:

$$L_1(\omega) = 0.5\, e^2(\omega), \tag{4}$$

where

$$e(\omega) = d(\omega) - y(\omega) = d(\omega) - \theta^T(\omega)x(\omega), \tag{5}$$

is the instantaneous error. The filter coefficient vector is then updated by

$$\theta(\omega+1) = \theta(\omega) - \nabla\frac{\partial L_1(\omega)}{\partial\theta(\omega)} = \theta(\omega) + \nabla e(\omega)x(\omega), \tag{6}$$

where ∇ is the step size that controls how the LMS algorithm gets to the steady state. Considering that R stands for the covariance matrix of $\mathbf{x}(\omega)$, and λ_{max} stands for its largest eigenvalue of $\boldsymbol{R}$. The LMS algorithm's generally known convergence criterion is

$$0 < \nabla < \frac{1}{\lambda_{max}}, \tag{7}$$

and with the premise of independence, the steady-state excess MSE is

$$P_{\text{ex}}(\infty) = \lim_{n\to\infty} E\left[\left((\theta(\omega)-\theta)^T x(\omega)\right)^2\right] = \frac{\eta}{2-\eta}P_0, \tag{8}$$

where P_0 is the power of observation noise and calculated as:

$$P_0 = E[v^2(\omega)], \tag{9}$$

and

$$\eta = \mathrm{tr}(R(I - \nabla R)^{-1}). \tag{10}$$

B. LEAKY LMS (LLMS)

The LLMS algorithm redefine the cost function $L_1(\omega)$ shown in (4) as $L_2(\omega)$, and

$$L_2(\omega) = 0.5e^2(\omega) + \gamma\theta^T(\omega)\theta(\omega), \tag{11}$$

where γ is the leakage factor, which is a positive parameter. Therefore, the minimum of $L_2(\omega)$ is found using a recursive process as:

$$\theta(\omega + 1) = \theta(\omega) - \nabla\frac{\partial L_2(\omega)}{\partial\theta(\omega)} = (1 - \nabla\gamma)\theta(\omega) + \nabla e(\omega)x(\omega). \tag{12}$$

C. ZERO-ATTRACTING LLMS (ZA-LLMS)

The ZA-LLMS algorithm redefine the cost function $L_1(\omega)$ shown in (4) as $L_3(\omega)$, and

$$L_3(\omega) = 0.5e^2(\omega) + \gamma\theta^T(\omega)\theta(\omega) + \gamma_3'|\theta(\omega)|, \tag{13}$$

where γ_3' is a positive constant, and $|\theta(\omega)| = \sqrt{\sum_{i=1}^{N} z_i^2}$. Then, the ZA-LLMS algorithm's update equation becomes

$$\theta(\omega + 1) = \theta(\omega) - \nabla\frac{\partial L_3(\omega)}{\partial\theta(\omega)} = (1 - \nabla\gamma)\theta(\omega) + \nabla e(\omega)x(\omega) - \rho\,\mathrm{sgn}[\theta(\omega)], \tag{14}$$

where $\rho = \nabla\gamma_3'$, and $sgn(.)$ is a component-wise sign function defined by

$$\mathrm{sgn(x)} = \frac{\mathrm{x}}{|\mathrm{x}|},\ \mathrm{x} \neq 0;\ \ 0,\ \mathrm{x} = 0. \tag{15}$$

D. REWEIGHTED ZERO-ATTRACTING LLMS (RZA-LLMS)

The RZA-LLMS algorithm define the cost function $L_4(\omega)$ as:

$$L_4(\omega) = 0.5\ e^2(\omega) + \gamma\theta^T(\omega)\theta(\omega) + \gamma_4'\sum_{i=1}^{N}\log\left(1 + \frac{|z_i|}{\zeta_4'}\right), \tag{16}$$

where γ_4' and ζ_4' are positive constants. Then, the RZA-LLMS algorithm's update equation becomes

$$\theta(\omega + 1) = \theta(\omega) - \nabla\frac{\partial L_4(\omega)}{\partial\theta(\omega)} = (1 - \nabla\gamma)\theta(\omega) + \nabla e(\omega)x(\omega) - \rho_4\frac{\mathrm{sgn}[\theta(\omega)]}{1 + \zeta_4|\theta(\omega)|}, \tag{17}$$

where $\rho_4 = (\nabla\gamma_4'/\zeta_4')$ is the zero-attracting parameter $\zeta_4 = (1/\zeta_4')$, and $\mathrm{sgn}(.)$ is the same as in (15). Upon comparing (12) and (17), it is evident that the RZA-LLMS algorithm contains an additional term, i.e., $-\rho_4(sgn[\theta(\omega)])/(1 + \zeta_4|\theta(\omega)|)$, which always pulls the tap coefficients towards zero. This term, known as a zero-attractor, is governed by ρ_4 to accelerate convergence when the preponderance of system coefficients is zero.

E. VARIABLE STEP SIZE LMS (VSS-LMS)

A different algorithm that uses variable step-size has been proposed in [18] and called variable-step-size LMS (VSS-LMS). In this algorithm, the step size (i.e. ∇) is recalculated in each iteration as:

$$\nabla'(\omega + 1) = \alpha_5\nabla'(\omega) + \kappa_5 e^2(\omega), \tag{18}$$

where $0 < \alpha_5 < 1$, and $\kappa_5 > 0$. Therefore

$$\nabla(\omega) = \nabla_{max}\ (\nabla'(\omega + 1) > \nabla_{max}),\ \nabla_{min}\ (\nabla'(\omega + 1) < \nabla_{min}),\ \nabla'(\omega + 1)\ (\text{otherwise}), \tag{19}$$

where $0 < \nabla_{\mathrm{min}} < \nabla_{\mathrm{max}}$, and $\nabla'(0)$ has no restrictions (although ∇_{max} could be better choice, as indicated in [18]). Equation (18) makes it clear that the step-size is always positive and is defined by the prediction error $\mathrm{e}(\omega)$, α_5, and κ_5. A higher step-size and quicker tracking are often produced by a significant initial prediction error. Reduced misadjustment results from a reduction in step size as the prediction error becomes less. According to [18], ∇_{max} is chosen in a manner that assures a limited mean-square error (MSE) as:

$$\nabla_{max} \leq \frac{2}{3\,\mathrm{tr}(E[\boldsymbol{R}])}. \tag{20}$$

F. ZERO-ATTRACTING VSS-LMS (ZA-VSS-LMS)

In this algorithm, the $\ell 1$ norm of the filter coefficient vector is added to the square of e in (4) to be as follows:

$$L_6(\omega) = 0.5e^2(\omega) + \varphi_6\|\theta(\omega)\|_1, \tag{21}$$

where φ_6 is a positive constant. Therefore, θ is updated as follows:

$$\theta(\omega+1) = \theta(\omega) - \frac{\nabla(\omega)}{2}\frac{\partial L_6(\omega)}{\partial\theta(\omega)} = \theta(\omega) + \nabla(\omega)e(\omega)x(\omega) - \beta_6(\omega)f\big(\theta(\omega)\big) \tag{22}$$

where $\beta_6(\omega) = \varphi_6\nabla(\omega)$ and $f\big(\theta(\omega)\big)$ is the sign function given in (15). When comparing (6) and (22), it is evident that (22) has an additional term $\Big(-\beta_6(\omega)f\big(\theta(\omega)\big)\Big)$. The tap coefficients are always pushed towards zero by this additional term. When the majority of the coefficients in θ are zero, the zero-attractor in (22) accelerates the convergence process. The method is known as the zero-attracting VSS-LMS (ZA-VSS-LMS) because the intensity of the zero-attractor is adjusted by $\beta_6(\omega)$.

G. REWEIGHTED ZERO-ATTRACTING VSS-LMS (RZA-VSS-LMS)

The shrinkage mechanism of the ZA-VSS-LMS algorithm handles zero and non-zero taps equally, which has a negative impact on performance for less sparse systems. When working with less sparse systems, weighting the ZA term in (22) may enhance the algorithm's performance [19]. Equation (4)'s cost function L_1 is redefined as:

$$L_7(\omega) = 0.5\,e^2(\omega) + \varphi_7\sum_{i=1}^{N}\log\left(1 + \frac{|z_i|}{\zeta_7}\right), \tag{23}$$

where φ_7 and ζ_7 are positive constants. Then, the same as before, by applying the gradient method we get

$$\theta(\omega+1) = \theta(\omega) + \nabla(\omega)e(\omega)x(\omega) - \rho_7(\omega)\frac{\mathrm{sgn}[\theta(\omega)]}{1 + \zeta_7'|\theta(\omega)|}, \tag{24}$$

where $\rho_7(\omega) = (\nabla(\omega)\varphi_7)/(\zeta_7)$, and $\zeta_7' = 1/\zeta_7$. The RZA-VSS-LMS algorithm introduces a weighted zero-attracting effect that only affects the taps whose magnitudes are comparable to $1/\zeta_7'$, while taps with much greater magnitudes experience little shrinkage. This property leads to a reduction in bias compared to other algorithms.

III. Particle Swarm Optimization (PSO)

In the conventional PSO model, an individual is represented as a particle in a D-dimensional space, where its location and velocity are denoted by $C_i = (C_{i1}, C_{i2}, \dots, C_{iD})$ and $S_i = (S_{i1}, S_{i2}, \dots, S_{iD})$, respectively. The movement of each particle is governed by:

$$S_{id} = \epsilon \cdot S_{id} + c_1 \cdot \mathrm{rand}() \cdot (U_{id} - C_{id}) + c_2 \cdot \mathrm{rand}() \cdot \big(U_g - C_{id}\big), \tag{25}$$

$$C_{id} = C_{id} + S_{id}. \tag{26}$$

The inertia weight ϵ controls the trade-off between global exploration and local refinement, while the acceleration coefficients c_1 and c_2 determine how much influence the particle's personal best and global best have on velocity updates. In the given set of equations, c_1 and c_2 are positive constants, typically both set to 2 in this study [32], and rand() is a random function within the range [0, 1]. The vector $U_i = (U_{i1}, U_{i2}, \dots, U_{iD})$ represents the best prior position of particle i, known as $pbest$, which corresponds to the highest fitness value. Similarly, $gbest$ indicates the position of the best particle in the entire population, defined by the vector $U_g = \big(U_{g1}, U_{g2}, \dots, U_{gD}\big)$. Here, C_{id}, S_{id}, and U_{id} refer to the d^{th} dimension of the vectors C_i, S_i, and U_i, respectively. The inertia weight ϵ is a crucial parameter that accelerates the Particle Swarm Optimization (PSO) convergence speed. It is dynamically adjusted using the formula [33]:

$$\epsilon = \epsilon_{max} - \frac{iter \cdot (\epsilon_{max} - \epsilon_{min})}{iter_{PSO}}, \tag{27}$$

where ϵ_{max} and $\epsilon_{\min}$ are typically set to 0.9 and 0.4, respectively. The iteration number iter represents the current iteration, and $iter_{PSO}$ is the maximum number of iterations, and set to 200.

A. COMPARISON WITH OTHER OPTIMIZATION TECHNIQUES

The proposed PSO-based LMS algorithm and several commonly used optimization techniques in adaptive filtering, including

Genetic Algorithms (GA), Artificial Bee Colony (ABC), Differential Evolution (DE), Artificial Neural Networks (ANN), and Reinforcement Learning (RL), are compared below in terms of their key strengths and limitations.

- **Genetic Algorithms (GA):** Natural selection serves as the inspiration for evolutionary optimization methods known as genetic algorithms (GA). They are frequently employed in adaptive filtering for global optimization, particularly in situations involving a vast and intricate search space. Even while GA is reliable and capable of solving non-linear and non-convex optimization problems, it is typically more computationally costly and has a slower rate of convergence than PSO, especially in real-time applications [34], [35].
- **Artificial Bee Colony (ABC):** Artificial Bee Colony (ABC) is a swarm intelligence-based optimization algorithm inspired by the foraging behavior of honeybees. It is known for its simplicity and effectiveness in solving complex optimization problems. ABC is highly effective in exploring the search space and can handle multi-modal optimization problems better than PSO. However, ABC can be slower in convergence compared to PSO, especially in high-dimensional problems [36], [37].
- **Differential Evolution (DE):** Differential Evolution (DE) is a population-based optimization method that evolves solutions over generations using crossover, mutation, and selection procedures. DE is renowned for its quick convergence and great efficacy in resolving continuous optimization issues. Nevertheless, DE necessitates meticulous adjustment of its parameters (such as crossover rate and mutation factor), which might be difficult in adaptive filtering applications [38], [39].
- **Artificial Neural Networks (ANN):** Deep learning and other optimization methods based on Artificial Neural Networks (ANN) are being utilized more and more in adaptive filtering for challenging signal processing problems. ANN is useful for tasks like system identification and noise suppression since it can represent extremely non-linear systems. However, ANN is less appropriate for real-time applications than PSO since it needs a lot of data for training and is computationally costly [40], [41].
- **Reinforcement Learning (RL):** A machine learning method called reinforcement learning (RL) teaches an agent to make decisions by interacting with its surroundings. It has been used for dynamic optimization in adaptive filtering. Real-time adaptive filtering can benefit from RL's high degree of adaptability and ability to manage dynamic settings. In contrast to PSO, RL is more complicated to implement and demands a large amount of processing power [42], [43].
- **Convex Combination Adaptive Filters:** Convex combination adaptive filters improve performance in both stationary and non-stationary environments by combining two or more adaptive filters. These filters are appropriate for dynamic contexts because they may strike a balance between minimal steady-state error and quick convergence. However, they can be computationally costly and necessitate careful combination parameter tweaking [44], [45].
- **Bayesian Adaptive Filtering:** Bayesian adaptive filtering estimates a system's state using probabilistic models. It is especially helpful in applications where users already know something about the system. Bayesian techniques can provide probabilistic estimates of the filter parameters and are very good at managing uncertainty. But Bayesian techniques are computationally demanding and necessitate system information, which is not always available [46], [47].

As shown in Table 1, the suggested PSO-based LMS algorithm strikes a balance between computational efficiency, convergence speed, and ease of implementation, making it particularly well-suited for real-time adaptive filtering applications, even though each optimization technique has advantages of its own.

A. INTEGRATION OF PSO WITH LMS

The goal of integrating PSO with the LMS approach is to optimize the filter coefficients in a dynamic manner. The filter coefficients are modeled as particles in a multi-dimensional optimization space using the PSO framework in this method. In order to converge on an ideal set of weights that minimize the Mean Square Error (MSE) between the intended and actual output of the filter, these particles iterate through positions and velocities. Every particle is a potential solution for the filter coefficients, and the weight space may be explored effectively thanks to the iterative process of updating locations and velocities. The Mean Square Error (MSE), which is minimized during the optimization process, is used as the fitness function to evaluate the PSO's performance. The MSE can be written as follows:

$$MSE = 0.5\, e^2(\omega), \tag{28}$$

where $d(\omega)$ is the target signal and $y(\omega) = \theta^T(\omega)x(\omega)$ is the output of the LMS filter. The error signal is represented by $e(\omega)$, which is defined as $e(\omega) = d(\omega) - y(\omega)$.

IV. PROPOSED METHOD

By helping to choose the best weights for each iteration, the PSO optimizer is incorporated into the ZA-LLMS, RZA-LLMS, ZA-VSS-LMS, and RZA-VSS-LMS algorithms, which is a revolutionary method that accelerates the convergence process. PSO is specifically used to identify the optimal weights for each adaptive filter that minimize the cost function.

Table 1. Comparison of Various Optimization Techniques in Adaptive Filtering: Strengths, Weaknesses, and Applicability.

Optimization Technique	Strengths	Weaknesses	Applicability
PSO-based LMS	• Fast convergence. • Easy to implement. • Suitable for real-time systems.	• May get stuck in local minima. • Requires tuning of hyperparameters.	• Real-time signal processing. • Noise cancellation. • System identification.
Genetic Algorithms (GA)	• Robust for global optimization. • Handles non-linear problems.	• Computationally expensive. • Slower convergence.	• Large search spaces. • Non-convex optimization problems.
Artificial Bee Colony (ABC)	• Effective in exploring search space. • Handles multi-modal problems.	• Slower convergence in high-dimensional problems.	• Complex optimization problems. • Multi-modal optimization.
Differential Evolution (DE)	• Fast convergence. • Effective for continuous optimization.	• Requires careful parameter tuning	• Continuous optimization problems. • High-dimensional problems.
Artificial Neural Networks (ANN)	• Models non-linear systems. • Effective for complex tasks.	• Requires large datasets. • Computationally expensive.	• Noise cancellation. • System identification. • Non-linear signal processing.
Reinforcement Learning (RL)	• Highly adaptive. • Suitable for dynamic environments.	• Complex to implement. • Requires significant computational resources.	• Dynamic environments. • Real-time adaptive filtering.
Convex Combination Adaptive Filters	• Balances fast convergence and low MSE. • Suitable for dynamic environments.	• Requires careful tuning of combination parameters. • Computationally expensive.	• Dynamic environments. • Real-time signal processing.
Bayesian Adaptive Filtering	• Handles uncertainty well. • Provides probabilistic estimates.	• Computationally intensive. • Requires prior knowledge of the system.	• Systems with prior knowledge. • Probabilistic signal processing.

This optimization process is repeated for each iteration until the adaptive filters converge to their optimal weights. A comprehensive flowchart of the proposed method is depicted in Figure 2. The flowchart shows the steps involved in the PSO-based optimization of the adaptive filters, including the initialization of the PSO algorithm, the calculation of the fitness function, the updating of the particle velocities and positions, and the termination criterion. The flowchart denotes four identified points inside dashed circles, which are referred to as point 1, point 2, point 3, and point 4 in this context. It is evident that the proposed algorithm's complete iteration involves the processes from point 1 to point 3, which comprises two consecutive phases. The first phase, from point 1 to point 2, entails applying PSO before each iteration of the AF algorithm. The second phase, from point 2 to point 3, involves executing one iteration of the AF algorithm on the optimized data. Consequently, the value of θ at each point is as follows:

- Point 1: The value from the preceding iteration, or θ, starting value.
- Point 2: The first value of θ or the value obtained in the last iteration PSO-optimized.
- Point 3: The value of θ derived from one full optimization then adaption iteration.

- Point 4: The last value of θ (or, the filter) applied to data filtering.

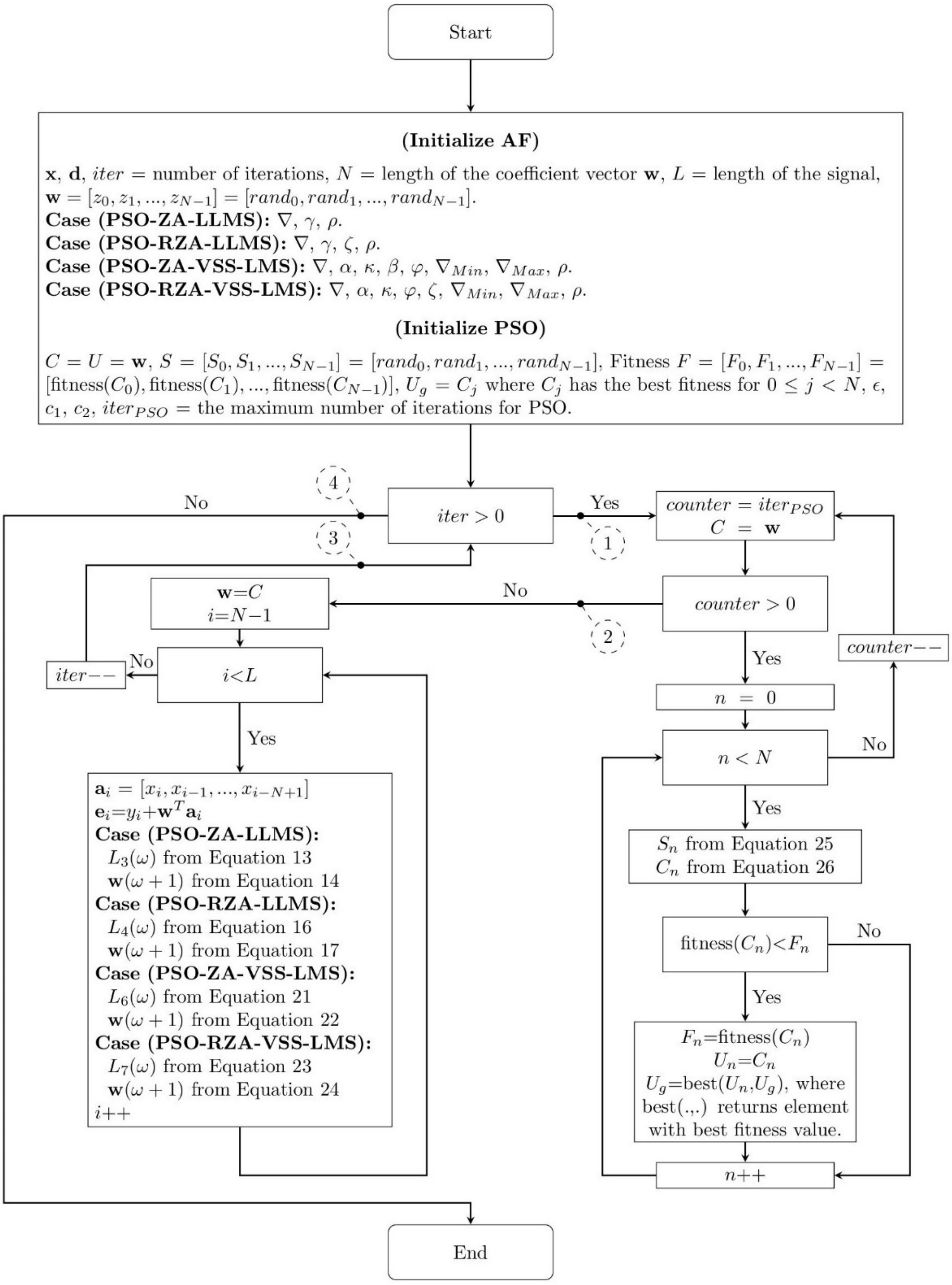


Figure 2. Flow chart for proposed algorithms.

V. Simulation Results and Discussion

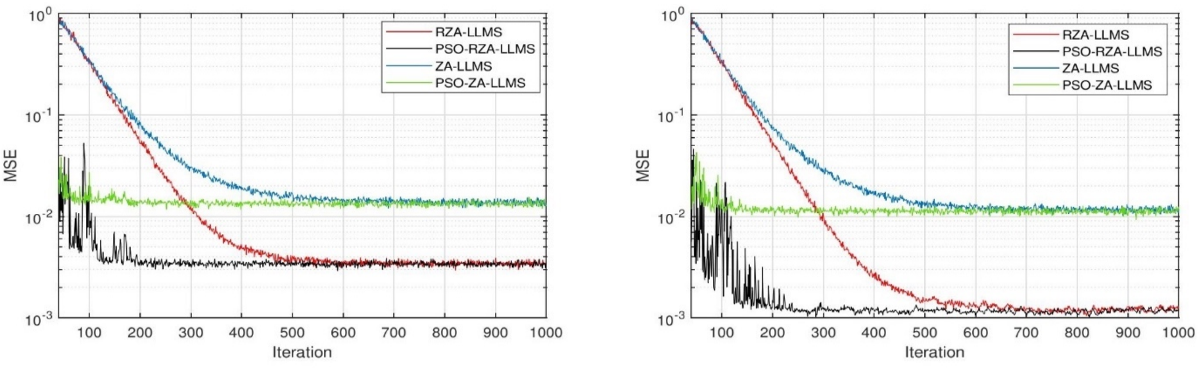
The suggested PSO-based algorithms are thoroughly compared in this part with those of the ZA-LLMS, RZA-LLMS, ZA-VSS-LMS, and RZA-VSS-LMS algorithms. Two different noise categories were used in order to verify the validity and robustness of the acquired data by means of comprehensive examinations. Characterised by a zero mean, a variance of 0.551, and a Signal-to-Noise Ratio (SNR) of 25, the first category is Additive White Gaussian Noise (AWGN). The second category consists in the use of Correlated Gaussian Sequence (CGS), commonly known as Colored Noise, which offers a more delicate and complex noise profile. In this part of the experiments, it is assumed that the input signal follows a white Gaussian distribution with a zero mean and a variance of one. In accordance with the symbols used in the equations in this paper, Table 2 presents a summary of the selected values from the literature, which are used as fixed values for models without PSO optimization and as initial values for models utilizing PSO. It also includes the parameter ranges used to constrain the optimization process.

Table 2. Suggested Ranges and Optimal Values for Parameters.

Parameter	Description	Suggested Range	Optimal Value
∇	Step size for LMS algorithms [18], [19]	[0.001, 0.1]	0.01
γ	Leakage factor for LLMS [14]	[0.001, 0.01]	0.002
γ_3	Zero-attracting parameter for ZA-LLMS [15]	[0.01, 0.1]	0.1
γ_4	Reweighted zero-attracting parameter for RZA-LLMS [17]	[0.01, 0.1]	0.01
ζ_4	Reweighted zero-attracting constant for RZA-LLMS [17]	[0.1, 1.0]	0.1
α_5	Decay factor for VSS-LMS [18]	[0.9, 0.99]	0.9
κ_5	Step size adjustment factor for VSS-LMS [18]	[0.01, 0.1]	0.05
∇_{min}	Minimum step size for VSS-LMS [18]	[0.001, 0.01]	0.01
∇_{max}	Maximum step size for VSS-LMS [18]	[0.01, 0.1]	0.04
φ_6	Zero-attracting parameter for ZA-VSS-LMS [21]	[0.01, 0.1]	0.1
φ_7	Reweighted zero-attracting parameter for RZA-VSS-LMS [21]	[0.001, 0.1]	0.001
ζ_7	Reweighted zero-attracting constant for RZA-VSS-LMS [21]	[0.1, 1.0]	0.1

The convergence behavior of the proposed hybrid approach, namely PSO-ZA-LLMS and PSO-RZA-LLMS, is visually depicted in Figure 3. The figure illustrates the performance of the algorithms in two distinct noise environments, namely AWGN as depicted in Figure 3a, and CGS noise as depicted in Figure 3b. Notably, the RZA-LLMS algorithm exhibits superior efficacy by effectively manipulating the weight coefficients, thereby mitigating bias to a greater extent when compared to the ZA-LLMS algorithm. Remarkably, both the PSO-ZA-LLMS and PSO-RZA-LLMS algorithms yield MSE values that are comparable to, or even lower than, those achieved by the ZA-LLMS and RZA-LLMS algorithms, respectively, in both noise environments. Furthermore, the PSO-ZA-LLMS and PSO-RZA-LLMS approaches exhibit accelerated convergence rates and significantly reduced MSE values in comparison to the ZA-LLMS and RZA-LLMS methods. Specifically, the PSO-ZA-LLMS and PSO-RZA-LLMS algorithms achieve steady-state performance at an early stage of the learning process, facilitated by the ability of PSO to enhance the weight coefficients and smoothly guide the MSE towards local minima with heightened efficiency. Conversely, both the ZA-LLMS and RZA-LLMS algorithms require significantly longer periods to reach the steady state. Moreover, the incorporation of PSO in the ZA-LLMS and RZA-LLMS algorithms enables a notable reduction in the computational time, resulting in approximately 400 fewer iterations compared to their non-PSO counterparts, namely ZA-LLMS and RZA-LLMS.

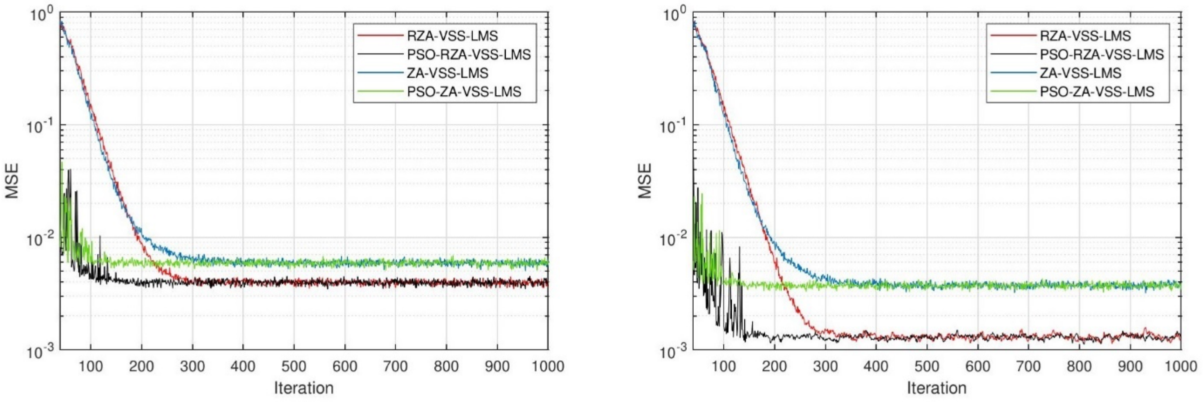
In order to establish the effectiveness of the proposed hybrid approach, it has been employed to assess the performance of the ZA-VSS-LMS and RZA-VSS-LMS algorithms. A similar trend is observed in Figure 4. Moreover, compared to their equivalents, the suggested PSO-ZA-VSS-LMS and PSO-RZA-VSS-LMS algorithms show remarkable traits like faster convergence rates and lower MSE values. This result is evidence of the strength of the hybrid method since it efficiently covers the capacity to understand several models, maximize their learning process, and interact smoothly with other algorithms.



(a) AWGN.

(b) CGS.

Figure 3: MSE to Iteration for ZA-LLMS, RZA-LLMS, PSO-ZA-LLMS, and PSO-RZA-LLMS using (a) Additive White Gaussian Noise (AWGN), and (b) Correlated Gaussian Sequence (CGS) noise.



(a) AWGN. **(b) CGS.**

Figure 4: MSE to Iteration for ZA-VSS-LMS, RZA-VSS-LMS, PSO-ZA-VSS-LMS, and RZA-VSS-LMS using (a) Additive White Gaussian Noise (AWGN), and (b) Correlated Gaussian Sequence (CGS) noise.

Table 3 shows the convergence iterations of the eight techniques over several SNR values. The results reveal that PSO causes a clear acceleration in the convergence of all algorithms and over all cases.

Table 3: Comparison of the convergence iterations and MSE of the 8 algorithms across various SNR values for AWGN.

SNR	**10**		**15**		**20**		**25**		**30**	
Measurement	**MSE**	**Iter. #**	**MSE**	**Iter. #**	**MSE**	**Iter. #**	**MSE**	**Iter. #**	**MSE**	**Iter. #**
ZA-LLMS [15]	0.131	597	0.046	607	0.019	643	0.013	663	0.012	711
PSO-ZA-LLMS	0.11	137	0.043	154	0.019	170	0.012	190	0.011	194
RZA-LLMS [17]	0.102	398	0.037	434	0.0093	459	0.0024	556	0.0013	602
PSO-RZA-LLMS	0.098	149	0.034	152	0.0089	192	0.0024	272	0.001	312
ZA-VSS-LMS [21]	0.134	201	0.057	222	0.01	283	0.0049	297	0.0021	314
PSO-ZA-VSS-LMS	0.125	97	0.055	140	0.01	189	0.003	195	0.0012	205
RZA-VSS-LMS [21]	0.102	261	0.046	284	0.011	317	0.005	324	0.0034	404
PSO-RZA-VSS-LMS	0.101	90	0.037	113	0.009	138	0.0042	179	0.0019	212

Figure 5 shows, at various SNR levels, the ratio of iterations needed to achieve MSE convergence in the PSO-optimized algorithms (PSO-ZA-LLMS, PSO-RZA-LLMS, PSO-ZA-VSS-LMS, RZA-VSS-LMS, respectively) compared to non-PSO versions (ZA-LLMS, RZA-LLMS, ZA-VSS-LMS, RZA-VSS-LMS, respectively). The results indicate that the use of PSO optimization accelerates convergence in all cases, with varying ratios. For instance, at SNR=10, PSO-ZA-LLMS requires only 23% of the iterations needed by ZA-LLMS to converge. Although Figure 5 shows that PSO-ZA-VSS-LMS and PSO-RZA-VSS-LMS require higher percentages compared to PSO-ZA-LLMS and PSO-RZA-LLMS, Table 3 demonstrates that they still require fewer iterations to converge.

VI. EVALUATION OF THE 2-D LMS-BASED ALGORITHMS IN 2D IMAGES

A. VALIDATION AND EXPERIMENTAL DESIGN

In order to evaluate the performance of the proposed technique on MRI scans, we applied several noise types to the original images presented in Figure 6. These noise variations are described as follows:

- **Gaussian Noise:** White noise with a mean of 0.3 and variance of 0.01.
- **Localvar Noise:** Zero-mean Gaussian noise with a local variance of 0.3.
- **Poisson Noise:** Representing photon-based noise commonly found in medical and astronomical imaging.
- **Salt & Pepper Noise:** Impulsive noise with a density of 0.3, simulating errors in data transmission or sensor malfunction.
- **Speckle Noise:** Often encountered in radar and ultrasonic imaging systems, a multiplicative noise model with a variance of 0.05.

We used the Peak Signal-to-Noise Ratio (PSNR) to evaluate the image restoration and noise lowering efficacy. Expressed in decibels, PSNR is a widely used statistic measuring the ratio between the noise present and the maximum achievable power of an image. A greater PSNR value denotes improved recovered image quality relative to original one.

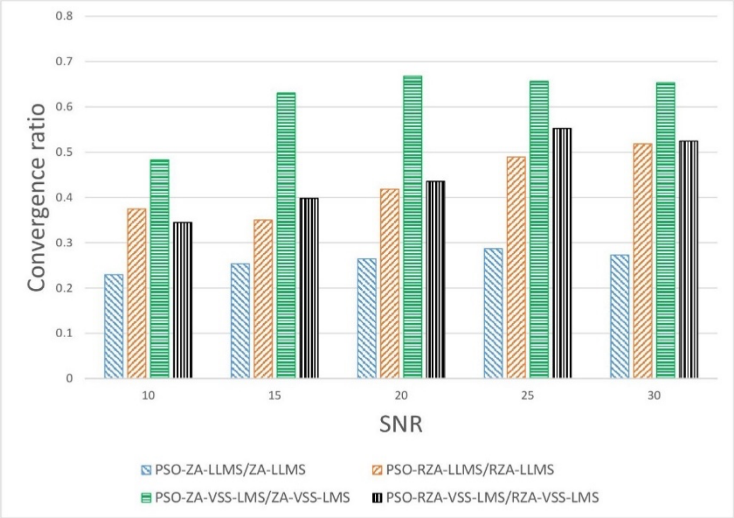
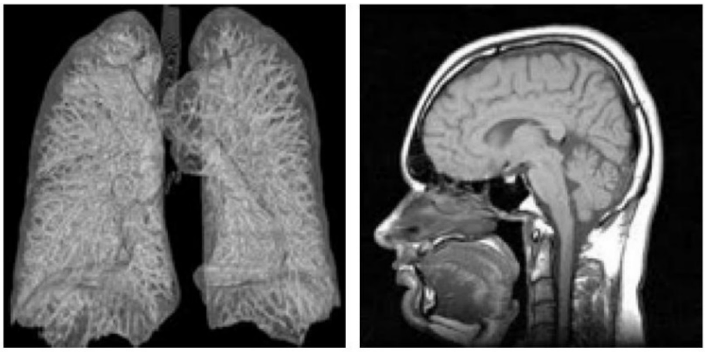


Figure 5: Comparison of convergence rates between algorithms with and without PSO optimization technique for AWGN.



(a) (b)

Figure 6: Original images used in the experiments.

B. RESULTS

Tables 4 and 5 show the Peak PSNR values for the images restored by both the suggested and conventional techniques. The outcomes allow one to make the following important observations:

- **Improvement with PSO-Enhanced Algorithms:** In terms of PSNR, the algorithms improved with Particle Swarm Optimization (PSO) routinely exceeded their conventional equivalents. For images damaged with Gaussian noise, for example, the RZA-LLMS-PSO technique attained a PSNR value of 18.0922 dB, as shown in Table 4, compared to 15.7893 dB acquired using the regular RZA-LLMS approach. Similarly, Table 5 shows that the PSO-augmented RZA-VSS-LMS algorithm produced a PSNR of 24.5831 dB under Gaussian noise conditions, which is substantially higher than the 20.5074 dB achieved by the non-PSO RZA-VSS-LMS approach. The PSO-RZA-VSS-LMS method outperformed others in preserving fine anatomical details for Speckle noise reduction, particularly in difficult areas like tissue borders and parenchymal textures in MRI data. The PSO-enhanced version, as seen in Table 5, got a PSNR of 24.56 dB, well above the baseline RZA-VSS-LMS, which got 20.95 dB, a 17.2% increase in signal quality. This improvement shows the capacity of the method to keep significant structural characteristics while lowering multiplicative artefacts. Visual outcomes (Table 7) show the clarity enhancement even more; PSO-RZA-VSS-LMS generates sharper edges and more homogenous textures in denoised photos. These quantitative and qualitative improvements highlight the efficiency of the model and especially fit it for diagnostic imaging uses since high-fidelity restoration of small structures is vital.

- **Algorithm Robustness Across Different Noise Variants:** The proposed algorithms demonstrated considerable robustness when applied to various noise types. For example, the ZA-VSS-LMS-PSO algorithm delivered notable PSNR values for both Poisson noise (16.9356 dB) and Speckle noise (16.4370 dB), as indicated in Table 4. This suggests that the algorithms are adept at managing both additive and multiplicative noise, positioning them as highly adaptable solutions for a diverse range of practical applications.
- **Visual Quality Enhancement:** Further confirmation of the effectiveness of the proposed algorithms is found in the visual quality improvements seen in the restored images, as presented in Tables 6 and 7. The images processed with PSO-enhanced algorithms exhibited sharper details and fewer distortion artifacts compared to those restored by the traditional methods. For example, as shown in Table 6, the image restored using the RZA-LLMS-PSO algorithm under Salt & Pepper noise conditions displays a significantly higher quality than the image restored using the conventional RZA-LLMS approach.

Table 4: PSNR values for the tested algorithms for the image in Figure 6a.

Noise Type	Gaussian	Localvar	Poisson	Salt & Pepper	Speckle
ZA-LLMS	15.9843	15.2826	15.1810	15.7862	15.5390
ZA-LLMS-PSO	16.2832	16.0262	15.7550	16.2376	15.9215
RZA-LLMS	15.7893	16.0250	15.9567	15.3292	16.0247
RZA-LLMS-PSO	18.0922	16.5063	16.6082	17.1189	16.6440
ZA-VSS-LMS	15.9636	16.2342	16.3827	15.7617	16.2120
ZA-VSS-LMS-PSO	16.4213	16.5370	16.9356	16.4390	16.4370
RZA-VSS-LMS	15.8754	16.1019	16.3433	15.6311	15.9219
RZA-VSS-LMS-PSO	16.4203	16.4362	16.7356	16.4395	16.4361

Table 5: PSNR values for the tested algorithms for the image in Figure 6b.

Noise Type	Gaussian	Localvar	Poisson	Salt & Pepper	Speckle
ZA-LLMS	18.4685	16.6679	16.7260	16.3003	16.4937
ZA-LLMS-PSO	18.5152	16.8115	16.7854	16.4749	16.5749
RZA-LLMS	18.4056	17.5042	17.3549	18.1443	17.5370
RZA-LLMS-PSO	20.2756	20.7498	20.7560	20.2489	20.6871
ZA-VSS-LMS	20.5119	20.9539	20.9358	20.7327	20.9812
ZA-VSS-LMS-PSO	24.5807	24.5587	24.5602	24.5490	21.2305
RZA-VSS-LMS	20.5074	21.0273	20.9358	20.1662	20.9545
RZA-VSS-LMS-PSO	24.5831	24.5609	24.5618	24.5513	24.5596

VII. ANALYSIS, LIMITATIONS, AND FUTURE WORKS

A. ANALYSIS AND DISCUSSION

The proposed PSO-based LMS adaptive filtering approach introduces a novel framework for optimizing filter coefficients, thereby improving convergence speed and reducing Mean Square Error (MSE). Compared to conventional LMS methods, the integration of PSO dynamically adjusts weights, preventing stagnation in local minima and enhancing adaptation efficiency. The experimental results validate the effectiveness of the proposed algorithms, particularly in synthetic noise environments (AWGN and CGS) and real-world MRI scans denoising.

Although the results show a notable increase in filtering accuracy and convergence rate, the following remarks draw attention to important aspects controlling performance:

- **Impact of PSO Optimization:** The adaptive exploration of the weight space of the PSO-enhanced LMS algorithms shows exceptional convergence characteristics. This advantage is particularly evident in cases where traditional LMS algorithms require extensive iterations to stabilize. As shown in Table 8, suboptimal PSO parameters can increase MSE by 29-75% and iterations by 26-63%, confirming the need for careful parameter selection.
- **Performance in Different Noise Conditions:** The proposed method demonstrates robust performance under Gaussian and colored Gaussian noise; however, further analysis is required to assess its effectiveness in handling non-Gaussian noise distributions and real-world signal variations.

B. LIMITATIONS OF THE STUDY

Despite the demonstrated improvements, the study has certain limitations that must be acknowledged:

- **Hyperparameter Sensitivity:** The choice of PSO parameters (e.g., inertia weight, number of particles, acceleration coefficients) significantly influences performance. A suboptimal configuration may lead to slower convergence or premature stagnation. As shown in Table 8, suboptimal PSO parameters can increase MSE by 29-75% and iterations by 26-63%, confirming the need for careful parameter selection.
- **Risk of Local Minima:** Although PSO mitigates local minima issues better than standard gradient-based approaches, it is still susceptible to convergence stagnation in complex search spaces.

Table 6: Retrieved images using the proposed and the compared algorithms for the image in Figure 6a.

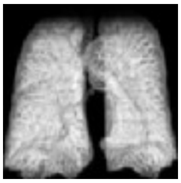
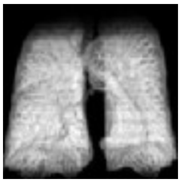
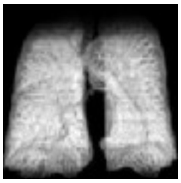
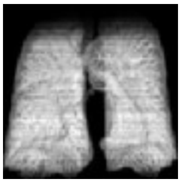
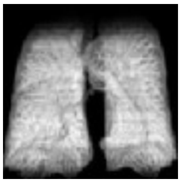
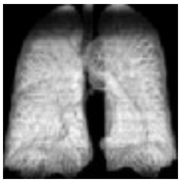
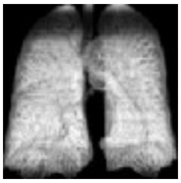
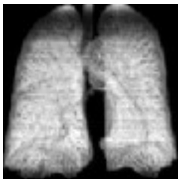
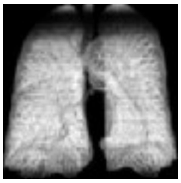
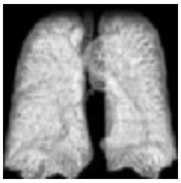
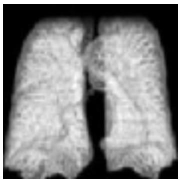
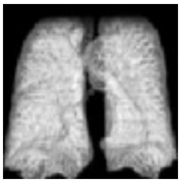
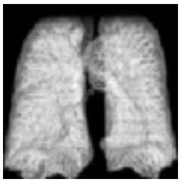
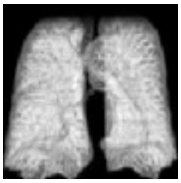
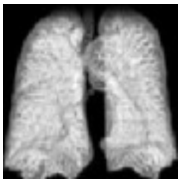
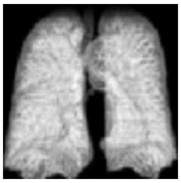
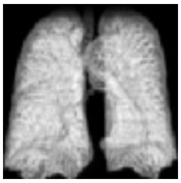
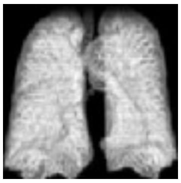
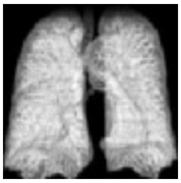
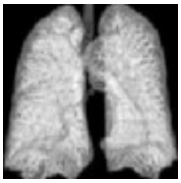
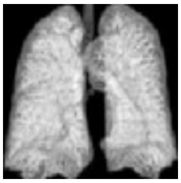
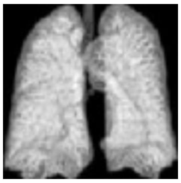
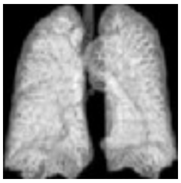
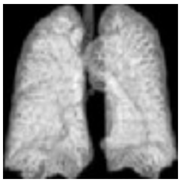
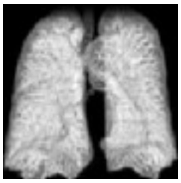
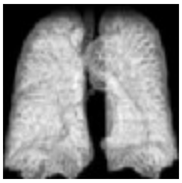
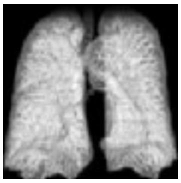
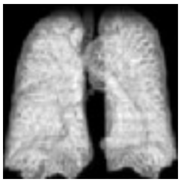
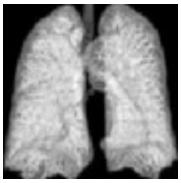
Technique	Gaussian	Localvar	Poisson	Salt&Pepper	Speckle
Noised Images					
ZA-LLMS					
ZA-LLMS-PSO					
RZA-LLMS					
RZA-LLMS-PSO					
ZA-VSS-LMS					
ZA-VSS-LMS-PSO					
RZA-VSS-LMS					
RZA-VSS-LMS-PSO					

Table 7: Retrieved images using the proposed and the compared algorithms for the image in Figure 6b

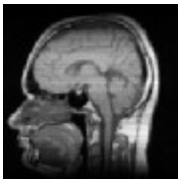
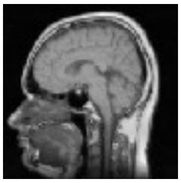
Technique	Gaussian	Localvar	Poisson	Salt&Pepper	Speckle
Noised Images					
ZA-LLMS					
ZA-LLMS-PSO					
RZA-LLMS					
RZA-LLMS-PSO					
ZA-VSS-LMS					
ZA-VSS-LMS-PSO					
RZA-VSS-LMS					
RZA-VSS-LMS-PSO					

Table 8: Ablation study evaluating the impact of individual PSO components on the performance of LMS-variant algorithms under AWGN conditions (SNR = 25 dB).

Algorithm	Configuration	MSE (Steady-State)	Iterations to Converge	% Faster vs. Non-PSO
PSO-ZA-LLMS	Full PSO	0.019	170	63%
	Fixed Inertia (ϵ = 0.5)	0.022	230	50%
	No Personal Best (c1 = 0)	0.025	290	37%
	No Global Best (c2 = 0)	0.027	310	33%
PSO-RZA-LLMS	Full PSO	0.0024	190	66%
	Fixed Inertia (ϵ = 0.5)	0.0031	240	57%
	No Personal Best (c1 = 0)	0.0038	300	46%
	No Global Best (c2 = 0)	0.0042	310	44%
PSO-ZA-VSS-LMS	Full PSO	0.003	195	34%
	Fixed Inertia (ϵ = 0.5)	0.0039	250	16%
	No Personal Best (c1 = 0)	0.0045	320	-6%
	No Global Best (c2 = 0)	0.0051	340	-12%
PSO-RZA-VSS-LMS	Full PSO	0.0042	179	45%
	Fixed Inertia (ϵ = 0.5)	0.0049	220	32%
	No Personal Best (c1 = 0)	0.0057	280	13%
	No Global Best (c2 = 0)	0.0063	300	7%

C. FUTURE RESEARCH DIRECTIONS

To further improve the proposed PSO-based LMS filtering approach, several potential research directions are identified:

- **Benchmarking Against Alternative Optimization Techniques:** While a theoretical comparison of PSO with Genetic Algorithms (GA), Differential Evolution (DE), Artificial Bee Colony (ABC), Artificial Neural Networks (ANN), and Reinforcement Learning (RL) has been provided, future work will include experimental benchmarking to validate performance trade-offs in different filtering tasks.
- **Advanced Hyperparameter Optimization:** Future research could include advanced hyperparameter tuning methods—such as Grid Search, Random Search, or Bayesian Optimization—to automatically identify ideal PSO parameters (e.g., inertia weight, acceleration coefficients), hence improving the robustness and generalization of the PSO-based adaptive filters. This would increase convergence consistency across datasets and lessen dependence on manual parameter selection.
- **Evaluation with Non-Gaussian Noise and Real-World Signals:** Broadening the study to encompass non-Gaussian noise distributions, larger real-world signal datasets, and dynamic telecommunication data would improve the robustness and generalizability of the proposed method.
- **Integration with Deep Learning Architectures:** The effectiveness of PSO-based LMS filtering in deep learning applications, such as Transformer-based denoising, LSTM-based adaptive filtering, and CNN-based signal restoration, should be investigated.
- **Ablation Studies to Assess PSO Contributions:** Conducting ablation studies by removing specific PSO components (e.g., adaptive weight selection, particle velocity updates) can provide deeper insights into the individual contributions of PSO to LMS performance.
- **Real-Time Implementation and Hardware Acceleration:** Implementing the PSO-LMS approach on FPGA, DSP, or GPU platforms can enable real-time deployment in latency-sensitive applications, such as speech enhancement and medical signal processing.

By tackling these topics, next research can improve PSO-based LMS filtering methods, increase their relevance in more general fields, and raise their performance in practical adaptive filtering applications.

VIII. CONCLUSION

This paper introduced a hybrid adaptive filtering framework that combines Particle Swarm Optimization (PSO) with various LMS-based algorithms, including ZA-LLMS, RZA-LLMS, ZA-VSS-LMS, and RZA-VSS-LMS. The integration of PSO significantly improved convergence speed and reduced Mean Square Error (MSE) compared to traditional LMS-based methods.

Experimental evaluations demonstrated the robustness of the proposed approach in both synthetic and real-world scenarios. In simulations with Additive White Gaussian Noise (AWGN) and Colored Gaussian Sequence (CGS), the PSO-enhanced methods outperformed conventional LMS variants, achieving up to a 67% reduction in iterations while maintaining lower MSE. Notably, the PSO-RZA-VSS-LMS algorithm achieved a PSNR of 24.58 dB under Gaussian noise conditions, surpassing its non-PSO counterpart at 20.51 dB. Similarly, under Salt & Pepper noise conditions, PSO-RZA-LLMS recorded 17.12 dB, compared to 15.32 dB with the standard RZA-LLMS, demonstrating superior denoising performance.

Further validation on 2D MRI images confirmed the adaptability of the proposed approach to image restoration tasks. Across multiple noise types, including Localvar noise, Poisson noise, Salt & Pepper noise, and Speckle noise, the PSO-augmented algorithms consistently outperformed their non-PSO counterparts in terms of PSNR and convergence speed. The experimental results revealed that PSO optimization not only enhanced filtering accuracy but also reduced computational complexity, with some PSO-enhanced methods requiring nearly 400 fewer iterations compared to their standard versions.

Despite these promising results, the study has certain limitations. While PSO enhances weight optimization, it may encounter local minima issues in high-dimensional filtering tasks, and its generalizability to broader signal processing applications, such as EEG analysis and telecommunication systems, remains an open research question. Additionally, the sensitivity of PSO to hyperparameter selection requires further investigation to ensure robust performance across different scenarios.

For future research, several directions are proposed. First, a comprehensive benchmarking of PSO-based LMS against alternative optimization techniques, such as Genetic Algorithms (GA) and Differential Evolution (DE), would provide deeper insights into its competitive advantages. Second, grid search and Bayesian optimization among other automated hyperparameter tuning techniques might improve model stability. Third, the flexibility of the suggested method would be enhanced by include real-world signal processing activities and non-Gaussian noise distributions in evaluation. Finally, investigating deep learning integrations with PSO-based LMS, such Transformer-based filtering or LSTM-based adaptive filtering, could open fresh opportunities for signal augmentation in challenging settings.

The outcomes of this work show the great possibilities of PSO-driven adaptive filtering as an effective substitute for real-time signal processing uses including dynamic noise cancellation, speech enhancement, and medical imaging. Future developments in hybrid learning methods and optimization strategies will help to improve and increase the relevance of the suggested strategy.

Authors' Contributions

In this study, author prepared and wrote the study.

Competing Interests

The authors declare that they have no competing interests.

References

[1] M. E. sayed M. Sakr and M. A. M. Hassan, Satellite tracking control system using optimal variable coefficients controllers based on evolutionary optimization techniques, El-Cezeri Journal of Science and Engineering, vol. 10, no. 2, pp. 326–348, 2023.

[2] M. K. Derdiman, Ayrık pso algoritması ile sehim kısıtı altında İki doğrultudaki kirişli döşemelerin güvenilirlik tabanlı optimizasyonu,'' El-Cezeri Journal of Science and Engineering, vol. 9, no. 1, pp. 49–64, 2022.

[3] Özdemir, S. "Oztürk, O. Şengül, and F. Kuncan, ''Position control of the suspended pendulum system with particle swarm optimization algorithm, El-Cezeri Journal of Science and Engineering, vol. 9, no. 2, pp. 669–679, 2022.

[4] Y. Bai, X. Gong, Q. Lu, Y. Song, W. Zhu, S. Xue, D. Wang, Z. Peng, and Z. Zhang, Application of adaptive filtering algorithm to the stability problem for double crystal monochromator. part i: Typical filtering algorithms, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, vol. 1048, p. 167924, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900222012165

[5] R. Karthick, A. Senthilselvi, P. Meenalochini, and S. Senthil Pandi, Design and analysis of linear phase finite impulse response filter using water strider optimization algorithm in fpga, Circuits, Systems, and Signal Processing, vol. 41, no. 9, pp. 5254–5282, Sep 2022. [Online]. Available: https://doi.org/10.1007/s00034-022-02034-2

[6] B. Durmuş, Infinite impulse response system identification using average differential evolution algorithm with local search, Neural Computing and Applications, vol. 34, no. 1, pp. 375–390, Jan 2022. [Online]. Available: https://doi.org/10.1007/s00521-021-06399-4

[7] G. Clark, S. Mitra, and S. Parker, Block implementation of adaptive digital filters, IEEE Transactions on Circuits and Systems, vol. 28, no. 6, pp. 584–592, 1981.

[8] Y. Yu, Z. Huang, H. He, Y. Zakharov, and R. C. de Lamare, Sparsity-aware robust normalized subband adaptive filtering algorithms with alternating optimization of parameters, IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 69, no. 9, pp. 3934–3938, 2022.

[9] G. Boidi, M. R. Da Silva, F. J. Profito, and I. F. Machado, Using machine learning radial basis function (rbf) method for predicting lubricated friction on textured and porous surfaces, Surface Topography: Metrology and Properties, vol. 8, no. 4, p. 044002, 2020.

[10] A. Singh, Adaptive noise cancellation, Dept. of Electronics & Communication, Netaji Subhas Institute of Technology, vol. 1, 2001.

[11] H. Kolivand, K. A. Akintoye, S. Asadianfam, and M. S. Rahim, Improved methods for finger vein identification using composite median-wiener filter and hierarchical centroid features extraction, Multimedia Tools and Applications, pp. 1–32, 2023.

[12] S. M. Wilson, M. K. Bohn, A. Madsen, T. Hundhausen, and K. Adeli, Lms-based continuous reference percentiles for 14 laboratory parameters in the caliper cohort of healthy children and adolescents,Clinical Chemistry and Laboratory Medicine (CCLM), 2023.

[13] B. Singh, M. Kandpal, and I. Hussain, Control of grid tied smart pv-dstatcom system using an adaptive technique, IEEE transactions on smart grid, vol. 9, no. 5, pp. 3986–3993, 2016.

[14] K. Mayyas and T. Aboulnasr, Leaky lms algorithm: Mse analysis for gaussian data, IEEE Transactions on Signal Processing, vol. 45, no. 4, pp. 927–934, 1997.

[15] M. S. Salman, Sparse leaky-lms algorithm for system identification and its convergence analysis, International Journal of Adaptive Control and Signal Processing, vol. 28, no. 10, pp. 1065–1072, 2014. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/acs.2428

[16] Y. Li and M. Hamamura, Zero-attracting variable-step-size least mean square algorithms for adaptive sparse channel estimation, International Journal of Adaptive Control and Signal Processing, vol. 29, no. 9, pp. 1189–1206, 2015. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/acs.2536

[17] M. S. Salman, A. A. Hameed, C. Turan, and B. Karlik, A new sparse convex combination of za-llms and rza-llms algorithms, in 2015 23nd Signal Processing and Communications Applications Conference (SIU), 2015, pp. 711–714.

[18] R. Kwong and E. Johnston, ‘‘A variable step size lms algorithm, IEEE Transactions on Signal Processing, vol. 40, no. 7, pp. 1633–1642, 1992.

[19] V. Mathews and Z. Xie, A stochastic gradient adaptive filter with gradient adaptive step size, IEEE Transactions on Signal Processing, vol. 41, no. 6, pp. 2075–2087, 1993.

[20] T. Aboulnasr and K. Mayyas, A robust variable step-size lms-type algorithm: analysis and simulations, IEEE Transactions on Signal Processing, vol. 45, no. 3, pp. 631–639, 1997.

[21] M. S. Salman, M. N. S. Jahromi, A. Hocanin, and O. Kukrer, A zero-attracting variable step-size lms algorithm for sparse system identification, in 2012 IX International Symposium on Telecommunications (BIHTEL), 2012, pp. 1–4.

[22] W. Jia, S. Kong, T. Cai, M. Li, Y. Jin, P. Wang, and Z. Dai, Steady-state performance analysis of the arctangent lms algorithm with gaussian input, IEEE Transactions on Circuits and Systems II: Express Briefs, 2023.

[23] H. Ferro, A combination of adaptive filters based on competitive learning principles, Available at SSRN 4345598, 2023.

[24] Y. He, J. Wei, Y. He, X. Rong, W. Guo, F. Wang, Y. Wang, and J. Liu, A process strategy planning of additive-subtractive hybrid manufacturing based multi-dimensional manufacturability evaluation of geometry feature, Journal of Manufacturing Systems, vol. 67, pp. 296–314, 2023.

[25] Á. A. Vázquez, J. G. Avalos, G. Sánchez, J. C. Sánchez, and H. Pérez, A comparative survey of convex combination of adaptive filters, IETE Journal of Research, vol. 69, no. 2, pp. 940–950, 2023.

[26] M. Martinez-Ramon, J. Arenas-García, A. Navia-Vázquez, and A. R. Figueiras-Vidal, An adaptive combination of adaptive filters for plant identification, in 2002 14th International Conference on Digital Signal Processing Proceedings. DSP 2002 (Cat. No. 02TH8628), vol. 2. IEEE, 2002, pp. 1195–1198.

[27] R. C. Eberhart and Y. Shi, Comparison between genetic algorithms and particle swarm optimization, in Evolutionary Programming VII, V. W. Porto, N. Saravanan, D. Waagen, and A. E. Eiben, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998, pp. 611–616.

[28] A. K. Mahapatra, N. Panda, and B. K. Pattanayak, Hybrid PSO (SGPSO) with the incorporation of discretization operator for training RBF neural network and optimal feature selection, Arabian Journal for Science and Engineering, vol. 48, no. 8, pp. 9991–10 019, Aug. 2023.

[29] F. Yun, H. Dong, C. Liang, T. Weimin, and T. Chao, Feature selection of XLPE cable condition diagnosis based on PSO-SVM, Arabian Journal for Science and Engineering, vol. 48, no. 5, pp. 5953–5963, May 2023.

[30] M. Saber, M. E. Ghoneim, and S. Kumar, Survey on design of digital fir filters using optimization models, Journal of Artificial Intelligence and Metaheuristics, vol. 2, no. 1, pp. 16–26, 2022.

[31] D. Krusienski and W. Jenkins, A particle swarm optimization-least mean squares algorithm for adaptive filtering, in Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, 2004., vol. 1, 2004, pp. 241–245 Vol.1.

[32] M. Chen, Y. Wang, P. Li, and H. Fu, Research on an improved pso algorithm with dual self-adaptation and dual variation, in 2022 IEEE International Conference on Mechatronics and Automation (ICMA), 2022, pp. 646–650.

[33] J. Xin, G. Chen, and Y. Hai, A particle swarm optimizer with multi-stage linearly-decreasing inertia weight, in 2009 International Joint Conference on Computational Sciences and Optimization, vol. 1, 2009, pp. 505–508.

[34] K. Deb, Multi-objective optimisation using evolutionary algorithms: an introduction, in Multi-objective evolutionary optimisation for product design and manufacturing. Springer, 2011, pp. 3–34.
[35] B. Alhijawi and A. Awajan, Genetic algorithms: Theory, genetic operators, solutions, and applications, Evolutionary Intelligence, vol. 17, no. 3, pp. 1245–1256, 2024.
[36] Q.-W. Chai, L. Kong, J.-S. Pan, and W.-M. Zheng, A novel discrete artificial bee colony algorithm combined with adaptive filtering to extract fetal electrocardiogram signals, Expert Systems with Applications, vol. 247, p. 123173, 2024.
[37] J. C. Bansal, H. Sharma, and S. S. Jadon, Artificial bee colony algorithm: a survey, International Journal of Advanced Intelligence Paradigms, vol. 5, no. 1-2, pp. 123–159, 2013.
[38] K. Price, R. M. Storn, and J. A. Lampinen, Differential evolution: a practical approach to global optimization, Springer Science & Business Media, 2006.
[39] E. Reyes-Davila, E. H. Haro, A. Casas-Ordaz, D. Oliva, and O. Avalos, Differential evolution: a survey on their operators and variants, Archives of Computational Methods in Engineering, pp. 1–30, 2024.
[40] H. Purwins, B. Li, T. Virtanen, J. Schlüter, S.-Y. Chang, and T. Sainath, Deep learning for audio signal processing, IEEE Journal of Selected Topics in Signal Processing, vol. 13, no. 2, pp. 206–219, 2019.
[41] A. Younesi, M. Ansari, M. Fazli, A. Ejlali, M. Shafique, and J. Henkel, A comprehensive survey of convolutions in deep learning: Applications, challenges, and future trends, IEEE Access, vol. 12, pp. 41 180–41 218, 2024.

El Cezeri
Journal of Science and Engineering
ISSN: 2148-3736

Research Article

Assessment of Kusum Biodiesel and Octanol Blends on Diesel Engine Performance and Emissions: A Comprehensive Study of Efficiency Trade-offs and Environmental Benefits

Kumaran P[1a], Sathiyaraj S[1b], Mahesh R[1c], Vignesh V[1d] and Vineeth A[1e]
[1]Aarupadai Veedu Institute of Technology, Vinayaka Mission's Research Foundation (Deemed to be University), Salem, India

kumaranp@avit.ac.in.

DOI : 10.31202/ecjse.1590321
Received: 28.11.2024 Accepted: 28.03.2025
How to cite this article:
Kumaran P, Sathiyaraj S, Mahesh R, Vignesh V and Vineeth A, " Assessment of Kusum Biodiesel and Octanol Blends on Diesel Engine Performance and Emissions: A Comprehensive Study of Efficiency Trade-offs and Environmental Benefits", El-Cezeri Journal of Science and Engineering, Vol: 12, Iss: 3, (2025), pp.(329-338).
ORCID: [a]0000-0003-2781-8999, [b]0000-0001-8401-8125, [c]0000-0002-3438-8715, [d]0009-0005-8644-4853, [e]0009-0000-3731-524X

Abstract This paper aims to investigate the effect of Kusum biodiesel and Octanol blends on Kirloaskar TV1 single cylinder diesel engine in terms of performance and emission. The tested blends, KB10O10D80, KB20O10D70, KB30O10D60, and KB40O10D50, revealed a significant decrease in emissions such as CO, HC, and smoke with an increase in biodiesel content at different engine loads. In particular, the KB10O10D80 blend, which had the lowest biodiesel content, recorded a CO reduction of 2.8% to 14.7%, HC reduction of 17.9% to 21.1% and smoke reduction of 3.45% to 18.18%, compared to diesel. However, the highest blend, KB40O10D50, recorded the highest emissions decrease with a reduction of CO by 27%, and a reduction of HC by 65.8% reduction in HC, and a 72.73% reduction in smoke at lower engine loads. But the study also found that brake thermal efficiency (BTE) reduced and brake specific energy consumption (BSEC) increased with the increase in biodiesel content. KB10O10D80 had the least reduction in BTE with the maximum reduction being 19.78% at 75% load and a moderate increase in BSEC, with an 11.66% increase at 100% load. However, the efficiency loss was the highest in KB40O10D50, but the emissions were reduced to the highest extent. Additionally, NOx emissions increased slightly, with KB10O10D80 showing a rise of 0.1% to 1.81%, and KB40O10D50 recording up to a 5.36% increase at full load. These findings underscore the trade-offs between environmental benefits and engine performance. The presence of Octanol, an oxygenated additive, improved combustion efficiency. However, the KB20O10D70 produced better performance and notably reduced emission.

Keywords: Diesel engine, Emissions, Kusum seed, n-Octanol, Performance.

I. INTRODUCTION

Depletion of renewable resources is an increasingly pressing issue, as the continuous extraction and consumption of natural resources, even those considered renewable, are reaching unsustainable levels. Biomass, wind, and solar energy are supposed to be replenished in a short time and are therefore categorized as renewable resources. But if used and managed inappropriately, they can be depleted, and this will cause long-term environmental and economic problems. One such alternative is biodiesel produced from non-edible waste materials such as seeds of trees like Kusum, scientifically known as Schleichera oleosa [1]. The Kusum seeds are usually thrown away as waste but the oil which is found in them can be used to produce biodiesel. Using such waste resources not only solves the problem of the exhaustion of traditional renewable resources, but also realizes the concept of circular economy, where waste materials are used to generate energy[2]. Through biodiesel from Kusum seed oil, the world can now turn to a renewable resource that is environmentally friendly and also at the same time deal with the issue of waste disposal. Schleichera oleosa is a prominent species in the Sapindaceae family and has multiple uses in medicine, agriculture, and industries. The Kusum tree is native to the tropical regions and is mostly located in India where it is an important source of income for the rural people. It is a deciduous or semi-evergreen tree, which is used ethnobotanically and as a host tree for lac culture [3]. The tree bears small, fleshy fruits known as berries that are either one or two- seeded. The Kusum tree is mainly located in India, but it is also present in other tropical countries, although the exact data on the distribution of the tree around the world are scarce. It is particularly common in Orissa and in the Darjeeling Himalayas, Andhra Pradhesh and Tamil Nadu where it is grown as a home garden tree. The climatic factors like rainfall and temperature affect the germination of seeds and growth of the tree which is important for its reproduction and spread. However, there is need to be addressed in terms of waste and harmful components. The waste produced from Kusum processing like most other agricultural and industrial processes if not well disposed can cause a lot of harm to the environment and the health of the people[4]. This response looks at the negative attributes of Kusum waste, possible environmental impacts and how these impacts can be prevented. Like other agricultural residues, Kusum waste may contain heavy metals including lead, cadmium, nickel, zinc and copper. These elements when present in high concentrations are a threat to the ecology and human health. For instance, in Akure, Nigeria, the study showed that the soils and sediments contained these metals and this poses a threat to human health and the environment [5]. Heavy metals in wastes can cause cytotoxic and genotoxic impacts on plants and animals. Research has revealed that metals such as Cu, Zn, and Pb can lead to chromosomal aberrations and other genetic changes in plants, which in turn affects the whole ecosystem. Disposal of Kusum waste in the wrong manner poses a threat to the soil and water resources. The accumulation of heavy metals and other pollutants in the soil can affect plant growth and enter the food chain, posing risks to human and animal health. Another strategy of handling hazardous waste is fractionation that involves sorting the waste into solid and liquid fractions. The solid part containing heavy metals can be safely disposed off while the liquid part can be utilized as fuel thus reducing on the conventional fuel usage and air borne emissions[6]. The investigations show that Kusum biodiesel can improve the engine efficiency and decrease some emissions but at the same time, it can raise NOx emissions. In general, the blends of Kusum biodiesel have been observed to have an enhanced BTE up to a certain blend ratio. For example, the biodiesel/diesel mixture of 20% Kusum biodiesel with 80% diesel (K-20 D-80) was found to be the best one as it increases BTE than that of diesel [7]. In the same way, the blending of hydrogen-compressed natural gas (HCNG) to KSOBD20 also enhanced the BTE to its maximum of 32. 09% at an injection pressure of 240 bar [8]. The study on the performance of n-Octanol as an additive in diesel engines has been investigated with the aim of improving the performance and reducing emissions, thus making it one of the best alternative fuel components. The study shows that n-Octanol has the potential of enhancing the efficiency of engines and lowering the emission of toxic gases but at the same time, it has some drawbacks that should be considered. This response will discuss the impact of n-Octanol on the diesel engine performance, emission and combustion using the findings from the above studies. Brake Thermal Efficiency (BTE): Literature surveys reveal that n-Octanol blends enhance brake thermal efficiency in a manner that is unchanging across experiments. This improvement is attributed to the fact that n-Octanol has more oxygen available for combustion as compared to heptane[9]. Nevertheless, due to the lower energy content of n-Octanol as compared to diesel, there is an increase in brake-specific fuel consumption [10]. Power Output: The presence of n-Octanol can result to a longer ignition delay because of its high latent heat of vaporization that decreases the initial combustion temperature and increases power [11].

The n-Octanol blends have been observed to decrease CO, HC and smoke opacity emissions by a large margin. This is mainly because n-Octanol is an oxygenated compound and thus, it undergoes complete combustion as compared to n-heptane [12]. Nitrogen Oxides (NOx) Emissions: In general, n-Octanol blends decrease NOx emissions compared with pure diesel, but increasing the proportion of n-Octanol in the blend increases the relative emissions of NOx. This is because of the formation of fuel rich zones that can raise local temperatures and formation of NOx [13]. Soot Formation and Oxidation: The presence of n-Octanol in biodiesel blends influences soot formation and oxidation phenomena. Even though the inception and coagulation rates are higher at higher n-Octanol to fuel ratios, the soot mass is lower because of higher oxidation [14]. Combustion Stability: n-Octanol has solubilising properties and helps in improving ignition qualities of fuel blends and increases stability of the blends. This is especially so in the blends with other alcohols or biodiesel where phase separation and low cetane numbers may be a problem [15]. In the light of the above findings, it can be concluded that n-Octanol has the potential of being used as a biodiesel additive; however, this comes with some trade-offs. The increase in BSFC due to the lower energy content of n-octanol may reduce some of the efficiency improvement. Furthermore, it is also found that the NOx emissions may rise with the increase in n-Octanol concentration and therefore, the blend ratios should be optimized to get the best balance of performance enhancement and reduced emissions [16]. In addition, the addition of n-Octanol to other additives like diethyl ether or nanoparticles is also known to improve combustion and emission characteristics, which mean that a multiple component system is desirable [17].

The Kusum biodiesel blends are seen to decrease the BSFC, which means better fuel economy. The KSOBD20 blend with HCNG recorded a BSFC of 0. 227 kg/kWh, which is less than that of diesel [18]. The B10 blend also had lower BSFC at full engine load conditions as shown in the study made by [19]. In general, the CO and HC emissions of Kusum biodiesel blends are lower than that of diesel. The KSOBD20 blend with HCNG reduced the CO and HC emissions to a large extent with the CO emission level reducing to 0. This was done by reducing the emission of CO to 013%, HC to 47 ppm [20]. One of the findings that is evident in most of the research is the rise in NOx emissions when using kusum biodiesel blends. This is due to the fact that biodiesel has higher oxygen content that leads to complete combustion and higher combustion temperature [21]. Kusum biodiesel blends are likely to decrease smoke and particulate emissions. The KSOBD20 with HCNG lowered the smoke emissions to 9% [22]. Likewise, the smoke density of the B10 blend was found to decrease as the proportion of biodiesel was raised [23]. The incorporation of hydrogen into the blends of kusum biodiesel has the effect of improving combustion features. The KSOBD20 blend with 15 lpm hydrogen got the CP of 69. 34 bar and an NHRR of 66.04 J/deg, which is an enhancement in the combustion efficiency [24]. It has been found that the best biodiesel blend is 20% Kusum biodiesel and 80% diesel because it has the best balance of performance and emissions [25]. Other sophisticated modeling approaches such as ANFIS-GA have been applied in estimating the performance of the engine and emissions, and in determining the right blending ratios [26].

II. NOVELTY STATMENT

While kusum biodiesel shows promise as an alternative fuel, challenges such as increased NOx emissions and the need for optimal blending ratios must be addressed. The use of higher alcohol additives like n-Octanol can enhance performance and reduce emissions, making kusum biodiesel a viable option for sustainable energy solutions. However, there is a lack of research n-Octanol blend with kusum biodiesel draws further attention towards this research where the engine performance were evaluated for KBOD blend

III. MATERIALS METHOD

Kusum seeds are first collected and cleaned to remove dust, debris, and other impurities. The cleaned seeds are then subjected to mechanical pressing, where they are crushed using a mechanical press to raw oil. Mechanical pressing is a physical process that does not involve chemical solvents, making it a cleaner extraction method. The resulting crude Kusum seed oil may contain impurities, free fatty acids (FFA), and other residues, making it necessary to refine or treat the oil before using it in the biodiesel production process. The crude oil is then washed to remove solid particles and sediments from the crude oil to produce cleaner oil for the next reactions.

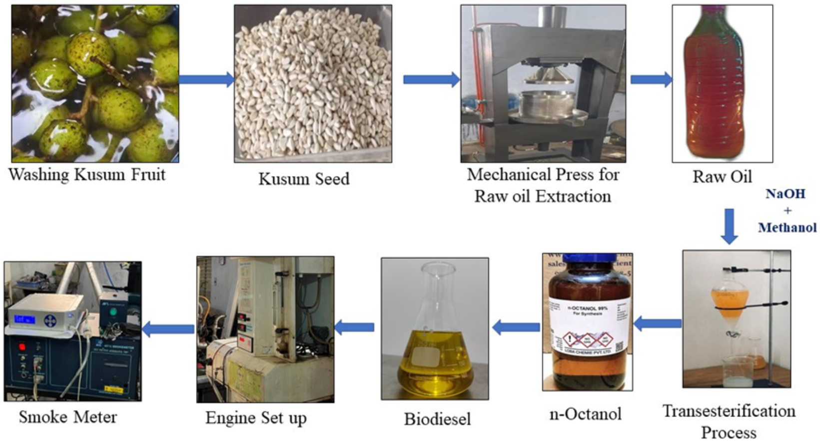


Figure 1: Schematic representation of process flow

The raw oil obtained from the mechanical pressing contained phospholipids, free fatty acids and water which will influence the efficiency of the transesterification process. The first process in pre-treatment is degumming where phospholipids and other impurities are separated and this is achieved through water treatment. The Kusum seed oil which has been pre-treated undergoes a chemical process known as transesterification to produce biodiesel. The oil is then blended with methanol and a catalyst sodium hydroxide (NaOH). The reaction occurs at the range of 60 to 65 °C, which is effective for the blending of the oil and methanol. The ratio of alcohol to the Kusum raw oil is normally 6:1 in order to achieve a stoichiometric reaction. The catalyst is soluble in alcohol and this is followed by the addition of this mixture to the oil. The reaction breaks down the triglycerides in the oil, forming methyl esters (biodiesel) and glycerol as a by-product. Once the transesterification reaction is complete, the reaction mixture is allowed to settle into two distinct layers. The upper layer consists of biodiesel (fatty acid methyl esters), while the lower layer contains glycerol, which is the by-product of the reaction. The separation happened by gravity based. The biodiesel is carefully decanted from the glycerol, which can be purified further. The glycerol layer may also contain traces of unreacted methanol, catalysts, and other impurities, which need to be removed before further use.

Table 1: Fuel composition and its nomenclature

Nomenclature	Kusum Biodiesel (KB)	n-Octanol (O)	Diesel (D)
KB10O10D80	10% by vol.	10% by vol.	80% by vol.
KB20O10D70	20% by vol.	10% by vol.	70% by vol.
KB30O10D60	30% by vol.	10% by vol.	60% by vol.
KB40O10D50	40% by vol.	10% by vol.	50% by vol.

The crude biodiesel obtained from the transesterification process still contains residual catalysts, methanol, soap, and glycerol that must be removed before the biodiesel is suitable for use. The purification process involved washing the biodiesel with warm water to dissolve and remove these impurities. The washing process can be done several times until the water comes out clear, ensuring that the biodiesel is free from contaminants. After washing, the biodiesel must be dried to eliminate any residual water. The overall process flow were depicted in the figure 1. After completing the transesterification, separation, and purification processes, the final product is ready for use. This Kusum biodiesel were blended with regular diesel and n-Octanol in different proportions as provide in the Table 1.

The theoretical yield of biodiesel was calculated based on the stoichiometry of the transesterification reaction, where one mole of triglyceride reacts with three moles of methanol to produce three moles of methyl esters (biodiesel) and one mole of glycerol. Assuming an average molar mass of Kusum oil triglycerides of approximately 885 g/mol, 2 liters of oil (1.84 kg, assuming a density of 0.92 g/mL) equated to about 2.08 moles of triglycerides. Using a 6:1 molar ratio, 6.24 moles of methanol were required, which corresponds to approximately 199.68 g of methanol (250 mL based on density). The theoretical mass of biodiesel was calculated as 1,823 g (or about 1.82 liters, assuming a density of 0.9 g/mL). The actual biodiesel yield was 1.7 liters, resulting in a yield percentage of 93.4% ((1.7 / 1.82) × 100). This high yield indicates efficient conversion, attributed to optimal reaction conditions, including accurate control of temperature, methanol ratio, and effective phase separation. The slight discrepancy between theoretical and actual yields is likely due to minor losses during washing and separation stages, demonstrating the effectiveness of Kusum oil as a viable feedstock for biodiesel production with high conversion efficiency. n-Octanol is obtained from the lobo chemie private limited, with 99% of purity. Purchased n-Octanol was added directly with the transesterified diesel to improve the fuel property at kept stirred with magnetic stirrer for 5minutes for homogeneous mixture. However, the n-Octanol content was limited to 100ml per litre for the various blends. The blends and its fuel composition are shown in the table 1.

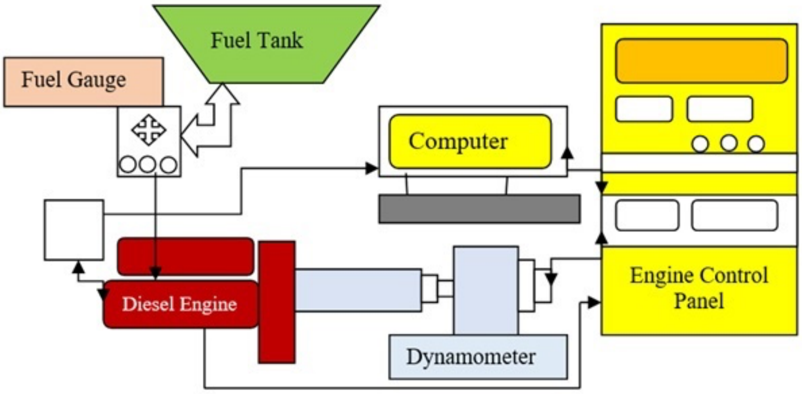


Figure 2: Experimental Setup

The experimental setup is shown in figure 2. In this study a single cylinder water-cooled Kirloskar TV1 engine was used with an engine capacity of 660cc and the fuel was injected at 21 bTDC and injection pressure at 210 bar with rated output of 5.2kW at 1500rpm. The inlet temperature was maintained at room temperature. Depending on the configuration, the load can range from zero to 5.5 kilowatts. The engine generates 5.5 kilowatts of power at a load of 0%, 25%, 50% and 100%. Diesel and biodiesel can be mixed with air in the fuel container, and the fuel mixture ratio can be adjusted using the fuel analyzer on the control screen. The AVL five-gas analyzer is used to measure gases emissions such as CO, CO_2, HC, O_2, and NO, while an AVL-developed smoke measuring instrument is used to assess the level of exhaust smoke.

IV. RESULT AND DISCUSSION

The engine was kept 15 minutes idle before each blend fuel operation. The engine kept to run till 15 minutes after attaining equilibrium, after final readings were taken and observed data were interpreted in this section.

4.1 Brake Thermal Efficiency

The brake thermal efficiency (BTE) of the Kusum biodiesel and Octanol blends (KB10O1D80, KB20O1D70, KB30O1D60, KB40O1D50) was compared with pure diesel across different engine loads at represented in the figure 3. At 0% load, BTE for diesel was 4.97%, while KB10O1D80 showed a slight decrease of 3.22%, and KB40O1D50 had a significant reduction of 17.46%. As the engine load increased to 25%, diesel had 19.1% BTE, with KB10O1D80 exhibiting a minimal drop of 3.09% and KB40O1D50 dropping by 18.32%. At 50% load, BTE for diesel was 25.7%, and KB10O10D80 experienced a moderate 3.11% decrease, while KB40O10D50 saw a more notable 17.87% reduction. Similarly, at 75% load, diesel reached 27.8%, with KB10O10D80 closely trailing at a 2.52% reduction, but KB40O10D50 lagging with a 19.78% decrease. At 100% load, diesel achieved 31.73%, whereas KB10O10D80 dropped by 2.24%, and KB40O10D50 significantly declined by 14.57%. Overall, KB10O10D80, containing 10% Kusum biodiesel and 10% Octanol, consistently performed closest to diesel across all loads, showing the smallest efficiency decrease, while KB40O10D50, with the highest Kusum biodiesel content (40%), had the greatest reduction in BTE. The lower BTE values with increased Kusum biodiesel content can be attributed to its lower calorific value compared to conventional diesel. Kusum biodiesel's reduced energy content per unit volume results in lower thermal efficiency, particularly at higher engine loads where more energy is demanded. The addition of Octanol, a higher oxygenated compound, helps improve the combustion process by providing more oxygen during combustion, enhancing fuel burn efficiency. However, the effect of Octanol is not enough to counterbalance the lower energy density of the higher Kusum biodiesel blends. This is evident in KB40O10D50, where a significant reduction in BTE is observed despite the presence of Octanol. On the other hand, KB10O10D80 maintains relatively higher thermal efficiency due to its higher diesel content, which helps sustain combustion performance. Overall, while Kusum biodiesel contributes to the sustainability of the blend, its increasing proportion leads to diminishing BTE, with Octanol serving to mitigate but not eliminate the efficiency losses.

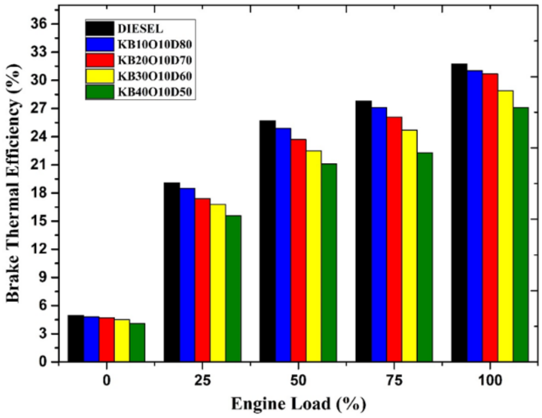


Figure 3: Brake thermal efficiency of KBOD blends at varied engine load

4.2 Brake specific Fuel Consumption

The Brake Specific Energy Consumption (BSEC) values for DIESEL, KB10O10D80, KB20O10D70, KB30O10D60, and KB40O10D50 show distinct variations across engine loads were depicted in the figure 4. At 0% load, DIESEL had a BSEC of 20342 kJ/kWh, while KB10O10D80 increased by 2.31%, and KB40O10D50 saw a 5.56% rise. At 25% load, DIESEL's BSEC was 162O10 kJ/kWh, with KB10O10D80 rising by 1.54%, and KB40O10D50 showing a 7.54% increase. For 50% load, DIESEL had 13200 kJ/kWh, and KB10O10D80 recorded a 6.44% increase, while KB40O10D50 rose by 8.97%. At 75% load, DIESEL reached 106O10 kJ/kWh, with KB10O10D80 increasing by 3.75%, and KB40O10D50 up by 11.75%. Lastly, at 100% load, DIESEL stood at 9841 kJ/kWh, while KB10O10D80 rose by 3.87%, and KB40O10D50 showed an 11.66% increase. Overall, the BSEC increases progressively with higher Kusum biodiesel content in all load conditions.

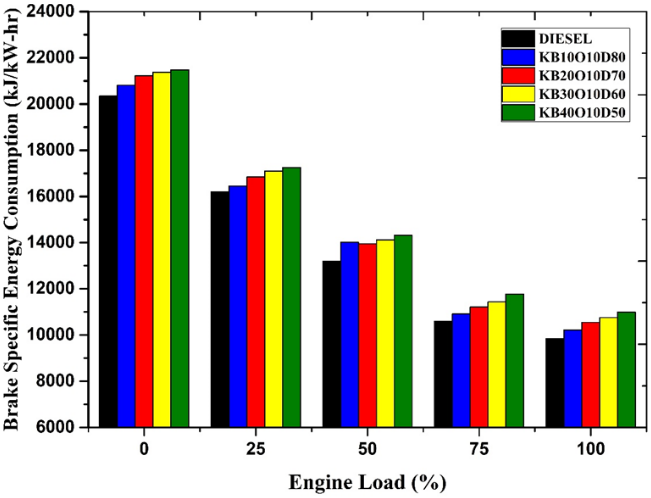


Figure 4: BSEC of KBOD blends at varied engine load

The rise in BSEC values with increasing Kusum biodiesel content is primarily due to the lower calorific value of biodiesel compared to conventional diesel. As more Kusum biodiesel is blended, more energy is required to produce the same amount of brake power, leading to a higher BSEC. Furthermore, the oxygenated nature of n-Octanol improves combustion efficiency, but it cannot fully offset the reduced energy density of Kusum biodiesel, particularly at higher blending ratios like KB40O10D50. This explains the progressively higher BSEC values observed with increasing biodiesel content. While lower blends like KB10O10D80 show relatively minor increases, higher blends such as KB40O10D50 exhibit more pronounced energy consumption increases, especially at higher loads, due to the significant difference in fuel properties from pure diesel.

4.3 Oxides of Nitrogen

The NOx emissions for DIESEL, KB10O10D80, KB20O10D70, KB30O10D60, and KB40O10D50 show changes across different engine loads were depicted in the figure 5. At 0% load, DIESEL recorded 999 ppm, while KB10O10D80 showed a slight increase of 0.1%, and KB40O10D50 exhibited a 2.1% rise. At 25% load, DIESEL emitted 1035 ppm, with KB10O10D80 showing a 0.58% increase, and KB40O10D50 rising by 3.57%. At 50% load, DIESEL emitted 1358 ppm, with KB10O10D80 increasing by 0.96%, and KB40O10D50 rising by 3.09%. At 75% load, DIESEL reached 1430 ppm, with KB10O10D80 showing a 1.33% increase, and KB40O10D50 rising by 4.96%.

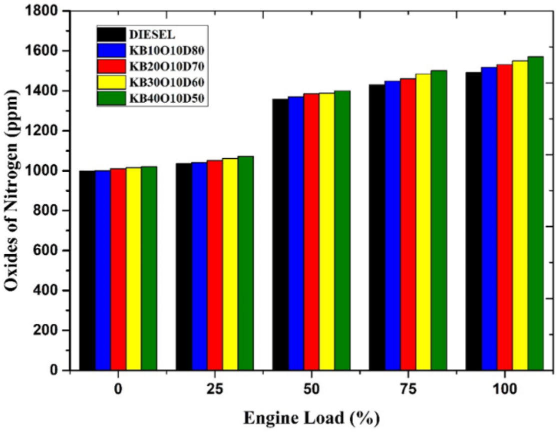


Figure 5: NOx emission of KBOD blends at varied engine load

At 100% load, DIESEL emitted 1491 ppm, with KB10O10D80 increasing by 1.81%, and KB40O10D50 showing a 5.36% rise. In general, NOx emissions show a progressive increase with higher Kusum biodiesel content across all engine load conditions. The increase in NOx emissions with higher blends of Kusum biodiesel can be attributed to the oxygenated nature of both biodiesel and n-Octanol, which enhance combustion efficiency and raise in-cylinder combustion temperatures. Elevated temperatures lead to higher thermal NOx formation, especially at higher engine loads where combustion intensity is greater. This is most evident in KB40O10D50, where the higher biodiesel proportion results in a significant rise in NOx emissions compared to diesel. The additional oxygen content promotes better fuel-air mixing, improving combustion but also raising peak temperatures, thereby increasing NOx formation. Although the presence of n-Octanol helps maintain combustion efficiency, it also contributes to the temperature rise, exacerbating NOx emissions in higher biodiesel blends. Therefore, while biodiesel blends improve the combustion process, the higher oxygen content and increased combustion temperatures explain the observed increase in NOx levels.

4.4 Smoke

The smoke emissions for DIESEL, KB10O10D80, KB20O10D70, KB30O10D60, and KB40O10D50 show significant variations across different engine loads were depicted in the figure 6. At 0% load, DIESEL recorded 11%, with KB10O10D80 showing a 18.18% decrease, and KB40O10D50 exhibiting a 72.73% reduction. At 25% load, DIESEL had 19%, with KB10O10D80 decreasing by 10.53%, and KB40O10D50 showing a 47.37% reduction. For 50% load, DIESEL emitted 29%, while KB10O10D80 saw a 3.45% decrease, and KB40O10D50 showed a 27.59% reduction. At 75% load, DIESEL had 41%, with KB10O10D80 recording a 7.32% decrease, and KB40O10D50 showing a 26.83% reduction. Lastly, at 100% load, DIESEL reached 59%, while KB10O10D80 decreased by 8.47%, and KB40O10D50 showed a 33.90% reduction. In summary, smoke emissions decrease progressively with higher Kusum biodiesel content across all engine loads, with the most significant reductions occurring in higher blends such as KB40O10D50.

The reduction in smoke emissions with increasing Kusum biodiesel content is primarily due to the oxygenated nature of both biodiesel and n-Octanol. These oxygenated compounds improve combustion by providing more oxygen during the fuel burn, leading to a more complete combustion process and significantly lowering particulate formation, which is the primary contributor to smoke. The oxygen content in biodiesel promotes a better fuel-air mixture, leading to reduced soot formation, especially at higher loads where diesel engines typically produce more smoke due to incomplete combustion. In higher blends like KB40O10D50, the reduced soot emissions are more pronounced, reflecting the increased oxygen availability from the biodiesel and Octanol mixture.

This results in lower particulate matter and reduced smoke formation compared to diesel, particularly at higher engine loads where combustion processes become more intense. Therefore, while the use of Kusum biodiesel in combination with n-Octanol helps improve combustion efficiency, it also significantly lowers smoke emissions across all engine load conditions.

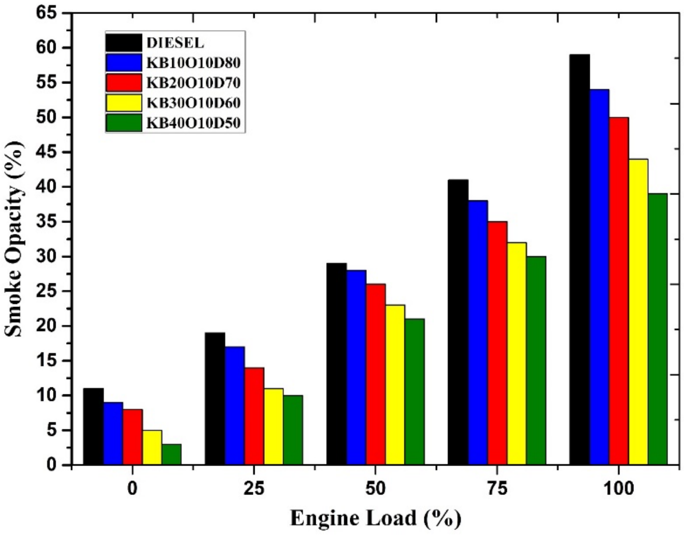


Figure 6: Smoke opacity of KBOD blends at varied engine load

4.5 Hydrocarbon emission

Comparing the hydrocarbon emissions of various blends (KB10O10D80, KB20O10D70, KB30O10D60, and KB40O10D50) to pure diesel reveals significant reductions at each engine load level were depicted in the figure 7. At 0% engine load, KB10O10D80 shows a 21.1% decrease in emissions compared to diesel, KB20O10D70 shows a 36.8% decrease, KB30O10D60 shows a 44.7% decrease, and KB40O10D50 shows a 65.8% decrease. At 25% engine load, emissions decrease by 20.6% for KB10O10D80, 38.2% for KB20O10D70, 44.1% for KB30O10D60, and 52.9% for KB40O10D50. At 50% engine load, the reductions are 14.8% for KB10O10D80, 22.2% for KB20O10D70, 44.4% for KB30O10D60, and 48.1% for KB40O10D50. At 75% engine load, the emissions drop by 20% for KB10O10D80, 37.1% for KB20O10D70, 45.7% for KB30O10D60, and 57.1% for KB40O10D50. Finally, at 100% engine load, KB10O10D80 has 17.9% lower emissions, KB20O10D70 has 35.9% lower emissions, KB30O10D60 has 41% lower emissions, and KB40O10D50 has 59% lower emissions compared to diesel. These numerical comparisons illustrate the significant reduction in hydrocarbon emissions as the proportion of Kusum biodiesel increases in the blend, demonstrating the effectiveness of these blends in reducing emissions relative to conventional diesel fuel.

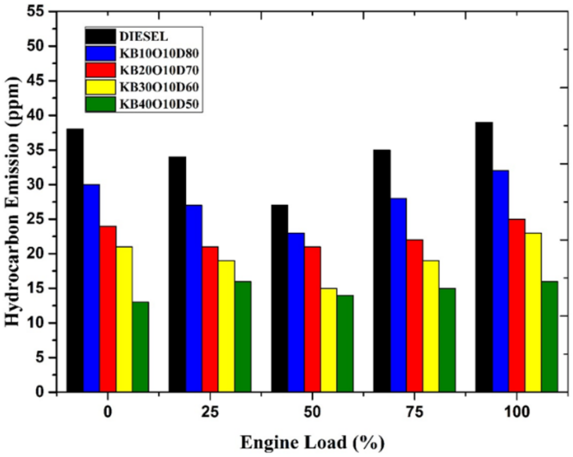


Figure 7: Hydrocarbon emission of KBOD blends at varied engine load

As Kusum biodiesel is introduced into the blend, emissions notably decrease due to biodiesel's higher oxygen content, which enhances combustion efficiency and reduces unburned hydrocarbons. The addition of N-octane also contributes to this improvement by increasing the octane number, which promotes more efficient combustion. The reduction in emissions increases from 10% Kusum biodiesel blend (KB10O10D80) to 40% Kusum biodiesel blend (KB40O10D50) and maximum reduction is observed in case of hydrocarbon emissions which is about 59% lower than that of pure diesel at full load condition. This consistent reduction of hydrocarbon emissions indicates that biodiesel content in the blends enhances the combustion process, which can be attributed to the oxygenated nature of biodiesel and the contribution of N-octane to the combustion quality. These factors not only help in reduction of hydrocarbon emissions but might also have a direct impact on Brake Thermal Efficiency (BTE) because of better combustion of fuel. Therefore, the use of more Kusum biodiesel in the fuel mix is beneficial to both the environment and performance, making it possible to use it in the reduction of emissions and the enhancement of engine efficiency.

4.6 Carbon monoxides

A comparison of the carbon monoxide (CO) emissions of the blends (KB10O10D80, KB20O10D70, KB30O10D60, and KB40O10D50) with that of pure diesel reveals a decrease in the emission at different engine loads were depicted in the figure 8. At 0% engine load, KB10O10D80 records a 7. 4% lower CO emissions than diesel, The KB20O10D70 has a 12. KB20O10D80 is down by 3%, KB30O10D60 is down by 16% and finally, KB40O10D50 is down by 21%. At 25% engine load, CO emissions decrease by 2.8% for KB10O10D80, 9.9% for KB20O10D70, 16.9% for KB30O10D60, and 23.9% for KB40O10D50. At 50% engine load, the reductions are 4.8% for KB10O10D80, 12.7% for KB20O10D70, 19% for KB30O10D60, and 27% for KB40O10D50. At 75% engine load, the CO emissions drop by 14.7% for KB10O10D80, 17.6% for KB20O10D70, 19.1% for KB30O10D60, and 23.5% for KB40O10D50. At 100% engine load, KB10O10D80 has 13.8% lower emissions compared to diesel, while KB20O10D70 shows an 8.5% decrease, KB30O10D60 shows an 18.1% decrease, and KB40O10D50 shows a 13.8% decrease. These numerical comparisons illustrate the general trend of reduced CO emissions with increasing Kusum biodiesel content, highlighting the potential of these blends to lower CO emissions relative to conventional diesel fuel.

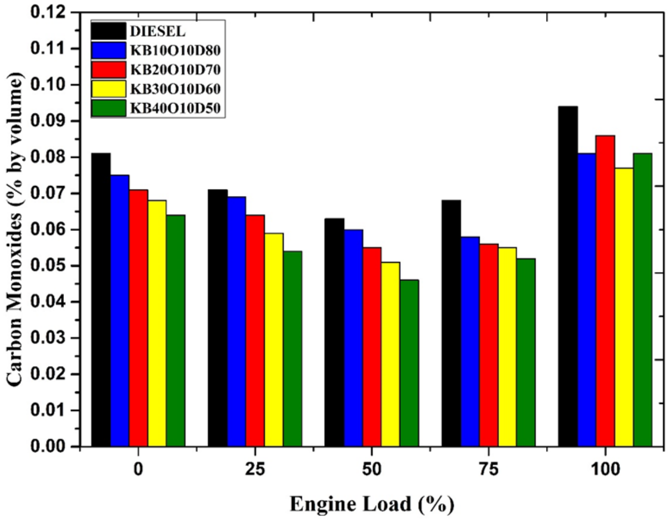


Figure 8: Carbon monoxide emission for KBOD blends across varied engine load

The decrease in carbon monoxide emissions with increase in Kusum biodiesel content can be explained by the following factors of fuel properties and combustion process. Kusum biodiesel has higher oxygen content than the normal diesel, and this results to efficient combustion and hence less production of CO which is associated with incomplete combustion. Because of the incorporation of oxygen in the biodiesel molecules, there is better oxidation of carbon during the combustion process thus reducing on the emission of CO. In the same regard, the addition of N-octane in the blends enhances the fuel's octane number, which enhances combustion efficiency through stable and efficient combustion especially under varying load conditions. Consequently, blends containing a higher percentage of Kusum biodiesel and N-octane have better combustion characteristics and reduced CO emission. This is in concordance with the observed decrease in emissions as demonstrated by the biodiesel containing oxygenated biodiesel and the improved combustion characteristics from N-octane in relation to pure diesel fuel.

V. CONCLUSION

The study demonstrates that the incorporation of Kusum biodiesel and Octanol blends into diesel fuel offers significant environmental benefits, notably in reducing key emissions such as carbon monoxide, hydrocarbons, and smoke, while simultaneously affecting brake thermal efficiency and brake specific energy consumption. The blends KB10O10D80, KB20O10D70, KB30O10D60, and KB40O10D50 exhibit a progressive reduction in emissions with increasing biodiesel content across various engine loads. However, the increase in biodiesel content also corresponds to a decrease in BTE and an increase in BSEC due to the lower calorific value of Kusum biodiesel compared to conventional diesel. The blend KB10O10D80, with the lowest biodiesel content, consistently performed closest to diesel in terms of thermal efficiency, showing the smallest efficiency decrease, while the highest blend, KB40O10D50, demonstrated the greatest reduction in BTE but the most significant improvements in emission reductions KB10O10D80 shows a Closest to diesel across all loads, showing the smallest decrease in BTE with a significant reduction in BTE, with a maximum decrease of 19.78% at 75% load.BTE decreases as the Kusum biodiesel content increases due to its lower calorific value compared to diesel. KB10O10D80 Shows a moderate increase in BSEC across all loads, indicating higher energy consumption compared to diesel. It exhibits the highest increase in BSEC, with up to an 11.66% rise at 100% load. BSEC increases with higher Kusum biodiesel content, primarily due to the lower energy density of biodiesel.

KB10O10D80 Shows a 17.9% to 21.1% decrease in HC emissions compared to diesel, depending on the engine load with highest reduction in HC emissions, with up to 65.8% lower emissions at 0% load. HC emissions decrease progressively as the Kusum biodiesel content increases, reflecting improved combustion. KB10O10D80 Shows a reduction in smoke emissions by 3.45% to 18.18% across various engine loads compared to diesel and exhibits the largest reduction in smoke, with up to 72.73% lower emissions at 0% load. Smoke emissions decrease significantly with higher Kusum biodiesel content due to better combustion and reduced soot formation. KB10O10D80: Shows a reduction in CO emissions by 2.8% to 14.7% across various engine loads compared to diesel. It demonstrates the most significant decrease in CO emissions, with up to 27% lower emissions at 50% load.CO emissions decrease consistently as the Kusum biodiesel content increases, indicating more complete combustion. KB10O10D80 Shows a slight increase in NOx emissions by 0.1% to 1.81% across different loads compared to diesel. It exhibits the highest increase in NOx emissions, with up to 5.36% higher emissions at 100% load. NOx emissions increase with higher Kusum biodiesel content, likely due to the higher in-cylinder temperatures and oxygen content leading to greater thermal NOx formation. These findings highlight the tradeoffs between environmental benefits and engine performance. The presence of n-Octanol in the blends contributes positively by enhancing combustion efficiency due to its oxygenated nature, which helps mitigate some of the efficiency losses associated with biodiesel's lower energy density. However, this mitigation is not sufficient to fully offset the reduced energy content of the higher biodiesel blends, particularly in KB40O10D50, which showed the greatest efficiency losses but also the highest reduction in CO, HC, and smoke emissions. The increase in NOx emissions with higher biodiesel content suggests that further optimization, possibly through engine tuning or the use of NOx reduction technologies, may be necessary to balance the benefits and drawbacks of these biodiesel blends. Overall, Kusum biodiesel and Octanol blends present a viable alternative to conventional diesel, offering a pathway to reducing emissions while maintaining acceptable engine performance, especially at lower biodiesel blend ratios.

Authors' Contributions

In this study, Kumaran P created the idea, conduction of experiment and Writing the original draft. Whereas Vignesh V and Vineeth A performed a literature review. Mahesh R and Sathiyaraj S done the supervision and validation of data

Competing Interests

The authors declare that they have no conflict of interest.

References

[1] S M. Veeraraghavan, et al., Influence of Hydrogen induction on performance and emission characteristics of an agricultural diesel engine fuelled with cultured Scenedesmus obliquus from industrial waste. Process Safety and Environmental Protection 187, 1576-1585, 2024

[2] K B. Parimi, et al., Investigation on a diesel engine fueled with hydrogen-compressed natural gas and Kusum seed biodiesel: Performance, combustion, and emission approach., Heat Transfer ,52(3), 2671-2693,2024

[3] Sakthimurugan, V., and S. Madhu. Novel Scenedesmus obliquus algae biofuel emission and performance characterise in Si-DLC coated diesel engine. Materials Today: Proceedings 77, 490-496, 2023

[4] S M. Veeraraghavan, et al., Waste-recovered quaternary blends: Enhancing engine performance through hydrogen induction by varied injection timing and pressure for sustainable practices., International Journal of Hydrogen Energy 87, 227-237, 2024

[5] K. T. Krishna, and V. Sakthimurugan, Various compression ratio strategies on introducing TiO2 in diesel-Palm oil blends in compression engine against diesel fuel for enhanced engine brake horsepower." AIP Conference Proceedings. Vol. 2822. No. 1. AIP Publishing, 2023

[6] C. Kishore, S. Yashvir, and N. Prateek, Comparative performance analysis on the DI diesel engine running on Karanja Kusum and Mahua methyl esters. Materials Today: Proceedings 46, 10496-10502,2021

[7] P. K. Praneel, and K. S. Vinod, Experimental analysis of diesel engine run on biodiesel derived from kusum oil. Materials Today: Proceedings 47, 2669-2672, 2021

[8] A. G. Poshetti, and M. S. Tandale. Experimental Study on CI Engine Performance for Optimum Blending Ratio of Blended Kusum Biodiesel. Advances in Energy Research, Vol. 2: Selected Papers from ICAER 2017. Springer Singapore, 2020

[9] A K. Yadav, A. Kumar, et al., Experimental investigations of performance and emissions characteristics of Kusum (Schleichera oleosa) biodiesel in a multi-cylinder transportation diesel engine. Waste and biomass valorization 8, 1331-1341, 2017

[10] K. Jayaraman, et al., Assessing the impact of sargassum algae biodiesel blends on energy conversion in a modified single-cylinder diesel engine with a silica-incorporated diamond-like coated piston. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects 46(1), 6882-6897, 2024

[11] J. Ahn, et al. Impact of Using n-Octanol/Diesel Blends on the Performance and Emissions of a Direct-Injection Diesel Engine. Energies 17(11), 2691, 2024

[12] K. Pachaiyappan, et al. Enhancing Combustion Efficiency: Utilizing Graphene Oxide Nanofluids As Fuel Additives With Tomato Oil Methyl Ester in CI Engines. Turkish Journal of Engineering, vol. 8, no. 4, pp. 720-8, 2024.

[13] J. Li, et al., Analysis of soot formation and oxidation processes in the diesel engine fueled by n-Octanol/biodiesel blends based on a detailed soot population balance model. Fuel, 358, 130376, 2024.

[14] M. Singh, et al. Effect of n-pentanol with novel water hyacinth biodiesel-diesel ternary blends on diesel engine performance and emission characteristics. Vietnam Journal of Chemistry ,2024.

[15] J. Li, et al. Blending n-Octanol with biodiesel for more efficient and cleaner combustion in diesel engines: a modeling study. Journal of Cleaner Production 403, 136877, 2023.

[16] G. Thippeshnaik, et al., Experimental investigation of compression ignition engine combustion, performance, and emission characteristics of ternary blends with higher alcohols (1-Heptanol and n-Octanol)., Energies 16(18), 6582, 2023

[17] AI. El -Seesy, et al., Enhancement of the combustion and stability aspects of diesel-methanolhydrous methanol blends utilizing n-octanol, diethyl ether, and nanoparticle additives. Journal of Cleaner Production 371, 133673, 2022

[18] M. Aneeque, et al., The combined effect of alcohols and Calophyllum inophyllum biodiesel using response surface methodology optimization., Sustainability 13.13 (2021): 7345.

[19] L. Atepor, The Effect of Bitter Kola (Garcinia) Biodiesel and n-Octanol Additives on Performance, Emission and Combustion Characteristics of Diesel Engines. Sustainable Education and Development 9. Springer International Publishing, 2021.

[20] S. Kodate, Vitthal, Y. Ajay Kumar, and G. N. Kumar. Combustion, performance and emission analysis of preheated KOME biodiesel as an alternate fuel for a diesel engine. Journal of Thermal Analysis and Calorimetry 141(6), 2335-2345, 2020

[21] B S. Chauhan, Naveen Kumar, and Haeng Muk Cho. A study on the performance and emission of a diesel engine fueled with Jatropha biodiesel oil and its blends. Energy 37(1), 616-622, 2012

[22] A. Singh, et al., An experimental investigation of emission performance of heterogenous catalyst jatropha biodiesel using RSM. Case Studies in Thermal Engineering 25, 100876, 2021

[23] IM R.Fattah, et al. Biodiesel production, characterization, engine performance, and emission characteristics of Malaysian Alexandrian laurel oil., RSC advances 4(34), 17787-17796, 2014

[24] J. Padmanaban, et al. Sustainability study and SWOT analysis of mixed biofuel blends in engine at various injection pressure analysed by experimentally and statistically., Scientific Reports 14(1), 31574, 2024

[25] D. Petkovi´c, et al. Neuro fuzzy estimation of the most influential parameters for Kusum biodiesel performance., Energy 229, 120621, 2021

[26] O. Khan, et al. Modelling of compression ignition engine by soft computing techniques (ANFIS-NSGA-II and RSM) to enhance the performance characteristics for leachate blends with nano-additives., Scientific Reports, 13(1), 15429, 2023.

El Cezeri
Journal of Science and Engineering
ISSN: 2148-3736

Research Article

Multicamera-Based Indoor Localization and Path Optimization for Mobile Robots Using ArUco Markers

Abdulhamit Sevgi[1a], Hasan Erdinç Koçer[2b]
[1] Mechatronic Program, OSTİM Technical University Ankara, Türkiye
[2] Electrical & Electronics Engineering, Selçuk University, Konya, Türkiye

abdulhamit.sevgi@ostimteknik.edu.tr

DOI : 10.31202/ecjse.1614568
Received: 06.01.2025 Accepted: 07.04.2025
How to cite this article:
Abdulhamit Sevgi, Hasan Erdinç Koçer, "Multicamera-Based Indoor Localization and Path Optimization forMobile Robots Using ArUco Markers ", El-Cezeri Journal of Science and Engineering, Vol: 12, Iss: 3, (2025), pp.(339-355).
ORCID: [a]0000-0003-3567-848X; [b] 0000-0002-0799-2140.

Abstract : In indoor environments where GPS signals are not available, determining the positions of mobile robots is challenging. Common methods include SLAM, marker-based localization, inertial measurement units (IMU),and hybrid positioning systems (HPS). SLAM enables simultaneous localization and mapping, while marker-based localization uses specific markers for position detection. IMUs track motion through velocity, acceleration, and angular velocity, and HPS combines sensors for improved accuracy. This study develops a route planning and motion optimization method using ArUco markers from the University of Córdoba. Detected via image processing, these markers guide robots by calculating the shortest and safest paths to the targets. Multiple cameras enhance motion range and vision, while an automatic pan adjustment addresses overlaps and alignment issues, ensuring seamless image integration. The proposed method demonstrates the potential of multicamera systems for reliable indoor navigation, offering promising applications in the industrial and service domains.

Keywords : Mobile Robots, Motion Optimization, Multi-Camera Systems, ArUco Markers, Indoor Localization.

1 Introduction

The localization and autonomous navigation of mobile robots have become pivotal areas of research, fostering the development of various systems and algorithms for position determination. Notable examples include Global Positioning System (GPS), Local Positioning Systems (LPS), and Hybrid Positioning Systems (HPS). GPS, which relies on satellite data to provide positional information, is widelyied in outdoor settisettings as military operations, meteorology, vehicle monitoring, cartography, and agriculture. This system delivers precise data on loonelocity, and timing. However, GPSHowever,door use due to constraints such as signal atsuch such such asn, lack of line-of-sight, and diminished accuracy. In contrast, LPS can determine the location of objects in a Cartesian coordinate frame, which is particularly suitable for indoor environments [1], [2], [3]. It is an important component of electronic performance and monitoring systems that facilitate accurate and efficient data collection to monitor external loads. Techniques commonly associated with LPS include Wi-Fi, Bluetooth Low Energy (BLE), Radio Frequency Identification (RFID), and Ultra Wideband (UWB). Although LPS offers advantages in terms of precision, power efficiency and indoor coverage, it is more complex and costly than GPS due to the need for dedicated infrastructure and hardware deployment. Hybrid Positioning Systems (HPS) integrate GPS and LPS technologies to provide seamless tracking in both indoor and outdoor environments. Using multiple data sources, HPS overcomes the limitations of GPS in confined or urban areas, improves positional accuracy, and expands the applicability of the system in various environments [4]. Localization systems are generally classified as outdoor and indoor positioning. For outdoor environments, GPS is preferred as a low-cost solution with high accuracy. In contrast, for indoor environments where GPS is ineffective, Indoor Positioning Systems (IPS) are essential. IPS uses a variety of technologies, including magnetic sensors, inertial sensors, acoustic sensors (audible, ultrasonic, and acoustic sounds), optical sensors, radio frequency-based sensors, and image processing with marker-based systems [3].

Image processing-based systems are based on the processing and analyzing of video data, employing fixed and mobile camera systems for localization [5]. Fixed camera systems use stationary cameras to track moving objects, determining their positions based on prominent features in the camera's field of view. Mobile camera systems, equipped with localization hardware, involve two stages: an offline phase where images of the environment are captured and processed to extract unique features, and an online phase where captured images are matched with stored features to determine the camera position.

Modern mobile robots are often equipped with onboard camera systems [6], providing an effective tool for determining their position within the environment. Visual localization is a subset of trigonometric methods, where vision systems consisting of one or two cameras detect markers using computer vision techniques [7], [8], [9]. Robots can derive information about the physical characteristics of the environment using these tools combined with various algorithms. For single-camera systems, the robot's position is calculated on the basis of marker positions in the environment captured during movement. Determining the robot's pre-movement coordinates requires knowledge of at least two markers, although static robot positions can be identified relative to a single marker, particularly when specialized 2D markers are used. Knowing the calibration parameters of the camera lens allows one to calculate the displacement and rotation of the marker in the image plane.

Several types of 2D marker systems have been developed for augmented reality applications, including ARStudio, ARToolkit, and ARTag [10]. These systems utilize plane patterns recognized by algorithms and ensure that marker patterns cannot be created by rotating another used pattern in the environment. These requirements are met by the ArUco marker system, which includes a library of functions for detection and localization. The monochromatic design of ArUco markers allows for efficient and rapid detection within a frame, with minimal computational effort. Visual localization using reference markers has been employed in UAV navigation and landing algorithms [11], [12] and mobile robot localization [13]. In addition, studies have explored landmark-based navigation, focusing on feature extraction and recognition of static environmental characteristics [14].

In this study, a localization and motion optimization algorithm is developed based on data obtained from ArUco markers. The algorithm calculates efficient routes to predefined targets, enabling the effective and reliable movement of mobile robots in indoor environments. The results of the proposed approach have been observed and analyzed.

2 Experimental Methods

2.1 ArUco Tags

ArUco markers are widely used in the field of computer vision for target detection, pose estimation, and position tracking. In this study, 5×5 markers are used, with markers numbered 00–05 designated for vehicles and markers numbered 11–17 assigned as targets. Identifying the pixel coordinates of all markers allows one to accurately determine the real-world positions of the vehicles within the local coordinate system. Figure 1 provides a visual representation of the ArUco codes used in this study.

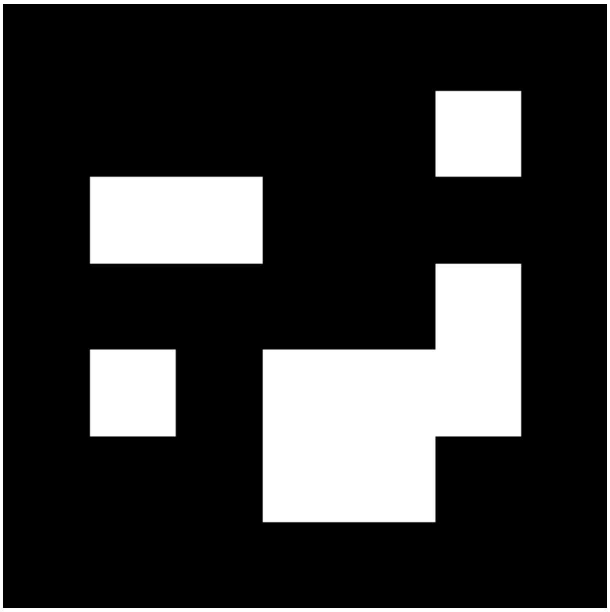


Figure 1: ArUco Markers.

The perspective transformation is used to determine the positions of vehicles [15]. By performing a perspective transformation using the pixel coordinates of the markers and their corresponding world coordinates, the two-dimensional information of the markers can be directly obtained. Since the pixel coordinates are proportional to the world coordinates, simple operations on the pixel coordinates enable the calculation of both the world coordinates and the orientation of the vehicles. In this study, ArUco markers are utilized to identify the positions of the four edges of robots and targets.

2.2 System Components

The robot developed for this study has three wheels: two active drive wheels located at the rear and one swiveling caster wheel in the front. It is designed with dimensions of $140 \times 140 \times 120,$ mm (Width × Length × Height) to accommodate all components and is manufactured using a 3D printer. The structural parameters are detailed in the mathematical models. Figure 2 illustrates one of the mobile robots produced.

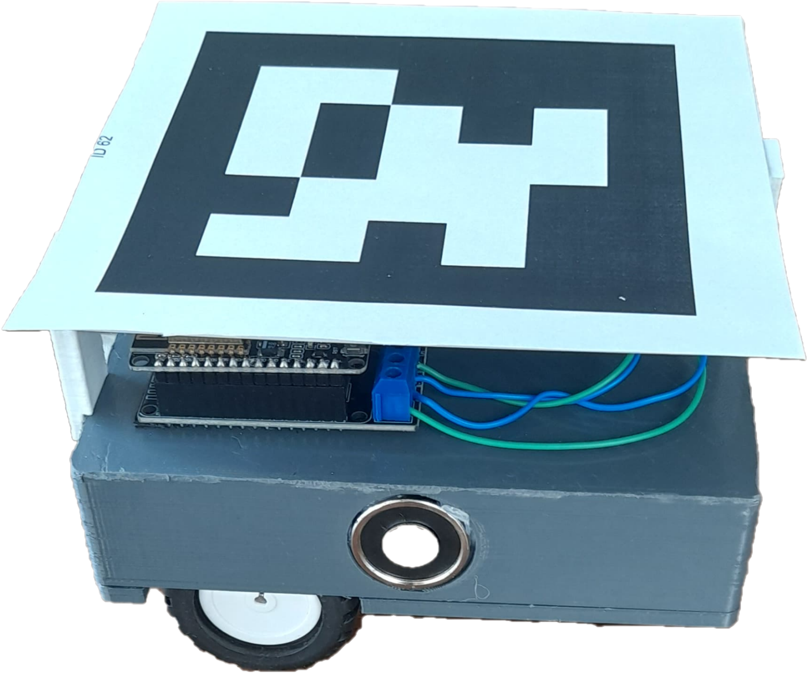


Figure 2: Produced Mobile Robot.

Communication within the system is facilitated by the ESP8266 controller board. These devices utilize the Wi-Fi communication protocol, which enables operation in low-power wireless networks. All robots and the main controller are connected to a single Wi-Fi network and are assigned static IP addresses. The other electronic components of the system are illustrated in Figure 3.

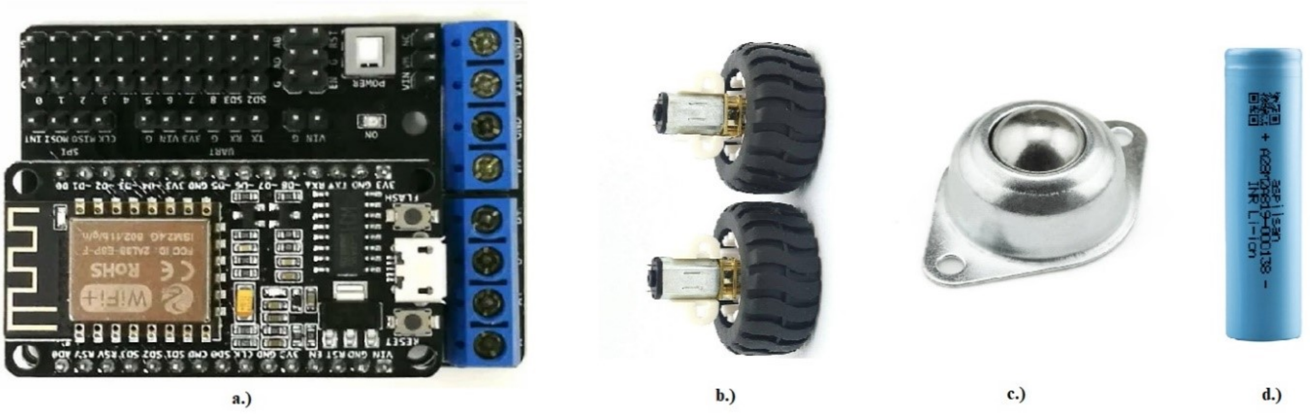


Figure 3: System Components: (a) ESP8266 and L293D motor driver board, (b) N20 motor and wheel, (c) Swivel caster wheel, (d) LiPo battery.

2.3 Kinematic Model of the Robot

Robots are controlled and classified according to their movement mechanisms. The motion of a robot is influenced by several factors, including maximum speed, total mass, wheel structure, differential mechanisms, and other variables. Although the precise interaction between load-bearing skeletal elements can be complex, joint behavior is typically estimated using linear and rotational kinematic joints. In this study, we plan the design and implementation of a two-wheeled mobile robot with differential drive, following the kinematic and dynamic simulation models available in the literature. Velocity decomposition is employed to address the robot's kinematic model. The velocity distributions of the right wheel and the robot body are illustrated in Figure 4.

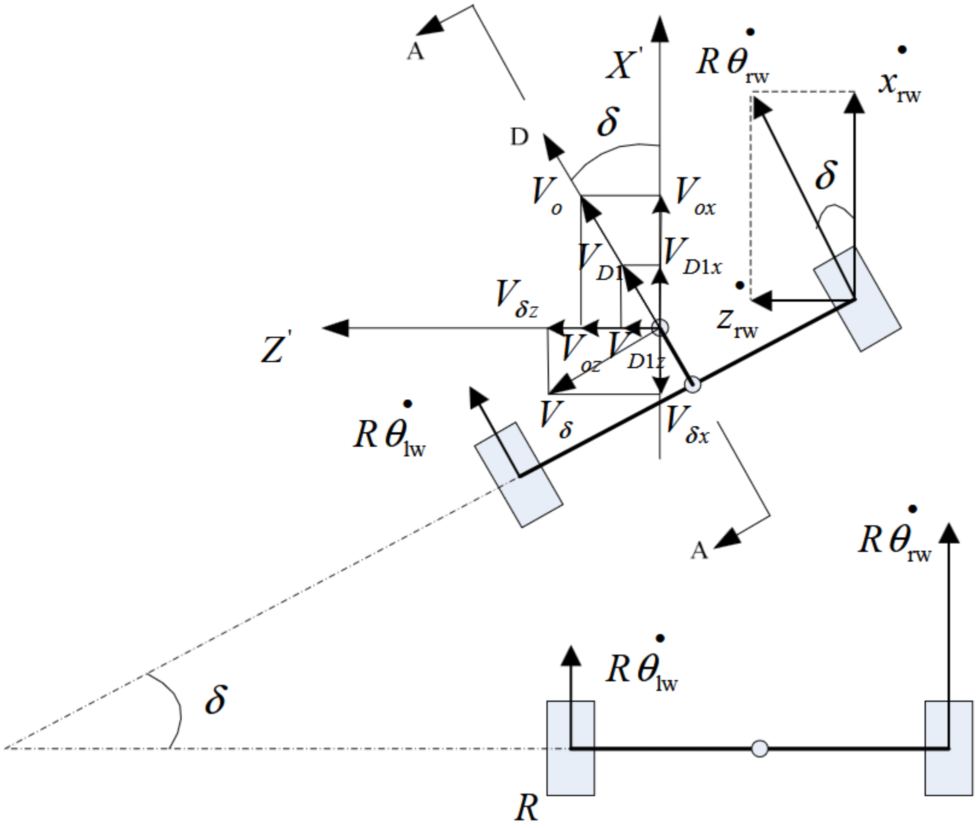


Figure 4: Velocity distributions of the robot [16].

In this robot, the vertical axis that passes through the midpoint of the two wheels represents the positive Y-axis, the direction of the robot's movement indicates the positive X-axis, and the axis of the two wheels corresponds to the positive Z-axis. The cross-sectional view (AA) of the robot is illustrated in Figure 5.

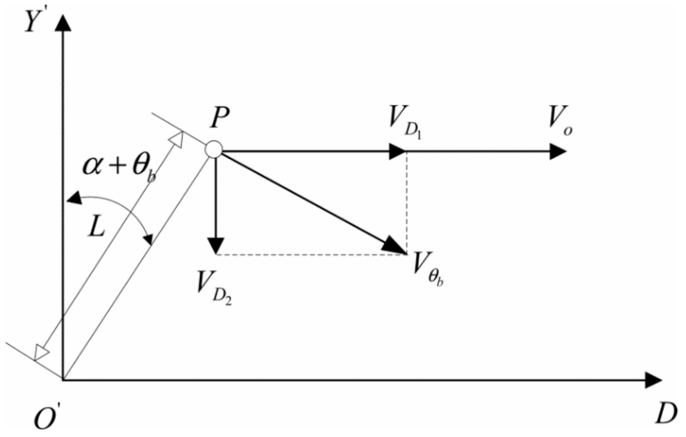


Figure 5: AA Cross-Sectional View of the Robot [16].

From Figures 4 and 5, the following parameters are defined:

- $\dot{R}_{\mathrm{rw}}$: Linear velocity of the right wheel (m/s).
- $\dot{x}_{\mathrm{rw}}$: Velocity of the right wheel along the X-axis (m/s).
- $\dot{z}_{\mathrm{rw}}$: Velocity of the right wheel along the Z-axis (m/s).
- δ: Deviation angle between the robot's current direction and its initial direction (rad).
- $\dot{\theta}_{\mathrm{rw}}$: Angular velocity of the right wheel (rad/s).
- R: Radius of the wheel (m).
- $\dot{\theta}_{\mathrm{lw}}$: Angular velocity of the left wheel (rad/s).
- V_{θ_b}: Rotational velocity of the center of mass around the OX axis (rad/s).
- L: Distance between the center of mass and the axle of the two wheels (m).
- V_{D2}: Velocity of V_{θ_b} along the OY axis (m/s).
- V_{D1}: Velocity of V_{θ_b} along the OD axis (m/s).
- V_o: Velocity of the robot's chassis center (m/s).
- V_δ: Rotational velocity of the center of mass around the OY axis (m/s).

The velocity analysis of the right wheel, as illustrated in Figure 4, is presented in Equation 2:

$$\begin{aligned} \dot{x}_{\mathrm{rw}} &= R\dot{\theta}_{\mathrm{rw}} \cos\delta, \\ \dot{y}_{\mathrm{rw}} &= 0, \\ \dot{z}_{\mathrm{rw}} &= R\dot{\theta}_{\mathrm{rw}} \sin\delta. \end{aligned} \tag{1}$$

Using the same method, the velocity of the left wheel is calculated as follows:

$$\begin{aligned} \dot{x}_{\mathrm{lw}} &= R\dot{\theta}_{\mathrm{lw}} \cos\delta, \\ \dot{y}_{\mathrm{lw}} &= 0, \\ \dot{z}_{\mathrm{lw}} &= R\dot{\theta}_{\mathrm{lw}} \sin\delta. \end{aligned} \tag{2}$$

Equations 1 and 2 define the velocity components of the right and left wheels. The velocity equation for the right wheel decomposes the linear velocity at the wheel's contact point with the ground into two components. The velocity component along the X-axis is calculated by multiplying the angular velocity by the radius of the wheel and projecting it onto the horizontal plane based on the robot orientation angle. Similarly, the velocity component along the Z-axis is derived by multiplying the angular velocity by the radius of the wheel and determining its projection onto the vertical plane according to the orientation of the robot. The velocity along the Y-axis is assumed to be zero, as the robot operates strictly within the horizontal plane. The equations for the left wheel follow the same principles as those for the right wheel, but they account for the independent angular velocities of each wheel.

V_{D1}, V_o, and V_δ are decomposed along the X-axis as V_{D1x}, V_{ox}, and $V_{\delta x}$, along the Z-axis as V_{D1z}, V_{oz}, and $V_{\delta z}$, and along the Y-axis as V_{D1y}, V_{oy}, and $V_{\delta y}$. The velocity equation along the X-axis is:

$$\begin{aligned} V_{D1x} &= V_{D1} \cos\delta, \\ V_{D1} &= V_{\theta b} \cos(\alpha + \theta_b), \\ V_{\theta b} &= L(\alpha + \theta_b). \end{aligned} \tag{3}$$

$$\begin{aligned} V_{ox} &= V_o \cos\delta, \\ V_o &= \frac{R\theta^*_{i\omega} + R\theta^*_{w\omega}}{2}. \end{aligned} \tag{4}$$

$$\begin{aligned} V_{\delta x} &= -V_\delta \sin\delta, \\ V_\delta &= \dot{\delta} L \sin(\alpha + \theta_b), \\ \dot{\delta} &= \frac{\dot{\theta}_{l\omega} - R\dot{\theta}_{r\omega}}{D}. \end{aligned} \tag{5}$$

To determine the total velocity of the robot's body along the X-axis, various velocity components are combined. The first component represents the projection of linear motion onto the X-axis, which varies according to the orientation angle. The second component corresponds to the velocity of the center of the chassis, projected on the basis of the orientation angle. Lastly, the third component arises from the rotational motion of the robot's body center and contributes to the velocity along the X-axis. The total velocity along the X-axis is expressed as the sum of these three components, modeling the combined effects of linear motion and rotational dynamics. Following these equations, the decomposed velocities along the X, Y, and Z axes are integrated to establish the complete kinematic model of the robot, ensuring consistency between the wheel velocities and the overall motion of the robot. This unified approach simplifies the analysis and facilitates accurate motion control.

$$\begin{aligned} V_x &= V_{D1x} + V_{ox} + V_{\delta x} \\ &= L(\alpha + \theta_b)\cos(\alpha + \theta_b)\cos\delta + \frac{R(\dot{\theta}_{l\omega} + \dot{\theta}_{r\omega})}{2}\cos\delta \\ &\quad - \frac{R(\dot{\theta}_{l\omega} - \dot{\theta}_{r\omega})}{D} L\sin(\alpha + \theta_b)\sin\delta. \end{aligned} \tag{6}$$

These relationships are established using the appropriate formula.

The velocity along the Y-axis is expressed as:

$$V_y = V_{D2} = -V_{\theta b}\sin(\alpha + \theta_b) \quad = -L(\alpha + \dot{\theta}_b)\sin(\alpha + \theta_b). \tag{7}$$

The velocity component of the robot's body along the Y-axis is derived by considering the effects of both inclination and body angle. The first term represents the linear velocity component along the Y-axis, which is influenced by the inclination angle and the tilt of the body. The second term represents the velocity component on the Y-axis of the center of the chassis as the robot moves. This equation helps to understand how the vertical movement of the center of the body is affected when the robot operates on an inclined surface.

The velocity along the Z-axis is expressed as:

$$\begin{aligned} V_{D1z} &= V_{D1}\sin\delta, \\ V_{D1} &= V_{\theta b}\cos(\alpha + \theta_b), \\ V_{\theta b} &= L(\alpha + \dot{\theta}_b). \end{aligned} \tag{8}$$

$$\begin{aligned} V_{oz} &= V_o \sin\delta, \\ V_o &= \frac{R\dot{\theta}_{l\omega} + R\dot{\theta}_{r\omega}}{2}. \end{aligned} \tag{9}$$

$$\begin{aligned} V_{\delta Z} &= V_\delta \cos\delta, \\ V_\delta &= \dot{\delta} \cdot L\sin(\alpha + \theta_b), \\ \dot{\delta} &= \frac{R\dot{\theta}_{l\omega} - R\dot{\theta}_{r\omega}}{D}. \end{aligned} \tag{10}$$

The overall velocity along the Z-axis is:

$$\begin{aligned} V_Z &= V_{D1z} + V_{oz} + V_{\delta Z} \\ &= L(\alpha + \dot{\theta}_b)\cos(\alpha + \theta_b)\sin\delta + \frac{R(\dot{\theta}_{l\omega} + \dot{\theta}_{r\omega})}{2}\sin\delta \\ &\quad + \frac{R(\dot{\theta}_{l\omega} - \dot{\theta}_{r\omega})}{D} L\sin(\alpha + \theta_b)\cos\delta. \end{aligned} \tag{11}$$

The total velocity of the robot along the Z-axis is determined by integrating the linear velocity, the rotational velocity around its center, and additional components influenced by the orientation angle. First, the projection of linear motion onto the Z-axis is calculated. Second, the velocity component resulting from the robot's rotational motion about its center is considered along this axis. Third, an additional velocity component is included to account for changes in the orientation angle. These equations illustrate how the robot moves along the Z-axis while changing direction, elucidating the shifts and orientation adjustments that occur during rotation. Based on the provided kinematic equations, the rotational angles of the robot's right and left wheels are presented in Equation 12:

$$\begin{aligned} \text{Left Angle} &= \phi = \theta_{lw} - \theta_b, \\ \text{Right Angle} &= \varphi = \theta_{rw} - \theta_b. \end{aligned} \tag{12}$$

These equations illustrate the relationship between the angular motions of the robot's left and right wheels and the tilt angle of its body. The difference in angular displacement between the two wheels determines the relative inclination of the body. As the angular movement of the body is influenced by the independent motion of the wheels, the body stability fluctuates based on the interaction of these variables. After determining the angle values, the center point of the robot's chassis and the deviation angle can be calculated using Equations 13 and 14, respectively:

$$p = \frac{p_{lw} + p_{rw}}{2} = \frac{R\theta_{lw} + R\theta_{rw}}{2} \quad = \frac{R}{2}(\phi + \varphi) + R(\alpha + \theta_b). \tag{13}$$

This equation calculates the travel distance of the robot. The total displacement of the robot is influenced by the combined angular motions of both the right and left wheels, as well as variations in the body's tilt angle. By multiplying the total angular movement of the wheels by the radius of the wheel, the total forward movement of the center of the chassis is determined. Furthermore, on inclined surfaces or under unbalanced conditions, an additional displacement component resulting from the body's angular tilt is included in the equation.

$$\delta = \frac{p_{lw} - p_{rw}}{D} = \frac{R\theta_{lw} - R\theta_{rw}}{D} \quad = \frac{R}{D}(\phi - \varphi). \tag{14}$$

This equation determines the robot's change in direction. The variation in the robot's orientation angle is calculated on the basis of the difference in angular motion between the right and left wheels. One of the key parameters influencing the turning angle is the distance between the two wheels, known as the wheelbase width. If both wheels rotate at the same speed, the robot moves in a straight line; however, if one wheel rotates faster than the other, the robot changes direction along a curved path. This equation facilitates the analysis of the robot's turning behavior on the basis of its wheel speeds. These equations illustrate that, by using the robot's kinematic model, the desired motion can be accomplished by controlling the velocities of the two wheels [16].

After deriving the kinematic models of the robots, one of the most critical steps in motion planning is the development of a navigation system. In this study, the navigation system comprises two video cameras and ArUco markers, as illustrated in Figure 6.

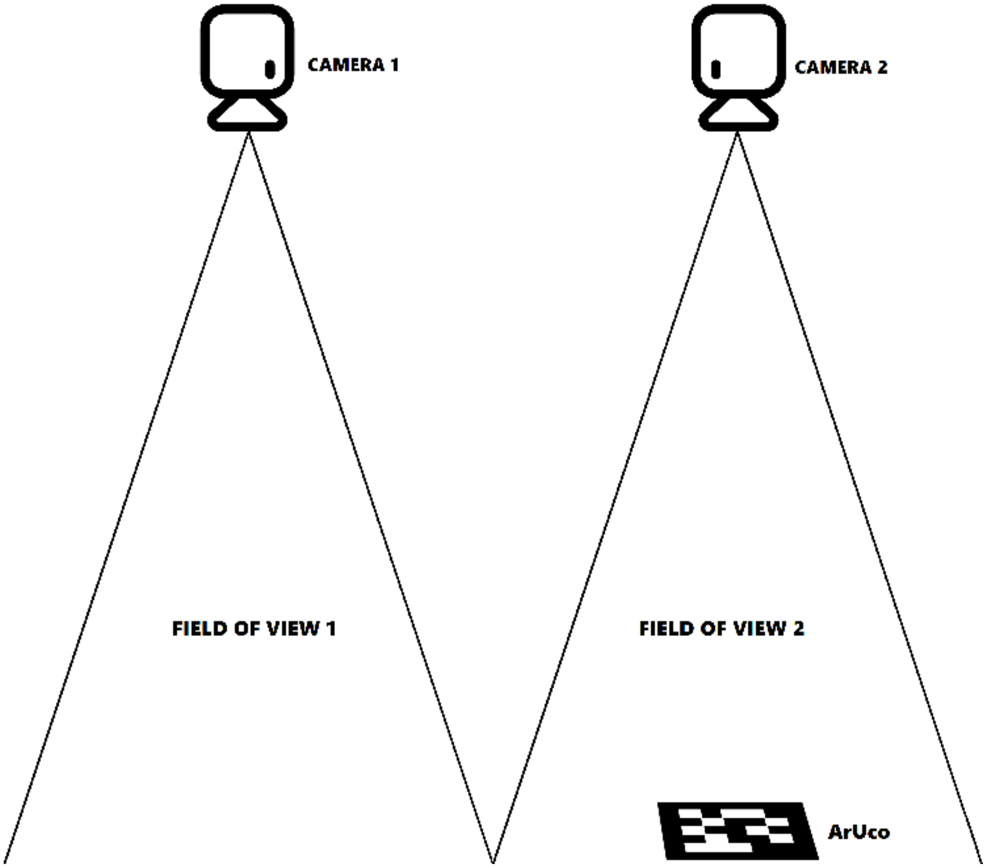
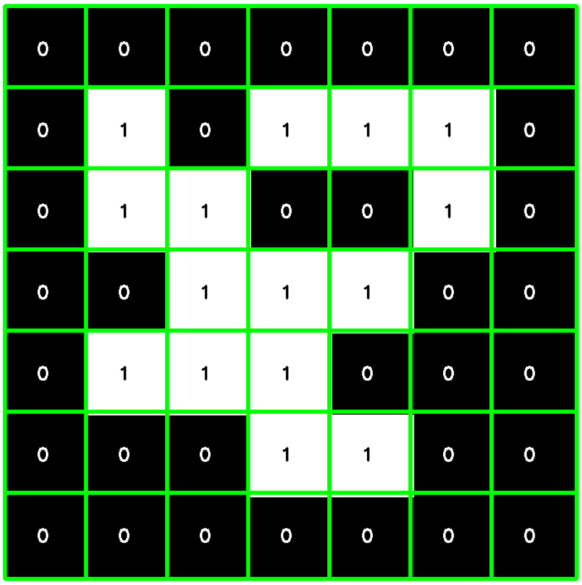


Figure 6: Cameras and System Field of View.

To track the robots, two Logitech C505e cameras are used to merge their fields of view. During the image merging process, the overlap between the fields of view of the cameras is identified, and any deviations in camera panning are corrected. In this context, feature points are extracted using the Scale-Invariant Feature Transform (SIFT) algorithm. The percentage of overlap is calculated and this information is used to ensure the precise alignment of the cameras [17]. Given that the cameras are mounted at a height of 3 meters, the operational area they can capture measures 3 x 6 m² (Figure 7). This area is organized to accommodate multiple robots, each equipped with a unique matrix reference marker.

Figure 7: System Workspace.

The size of the ArUco markers is 5 × 5. Python programming language, in conjunction with OpenCV libraries, is used to process the ArUco markers. By placing an ArUco marker on top of the robot, aligned with its center of mass, the robot's position and orientation can be accurately determined. ArUco markers are square-shaped and feature a unique binary matrix along with a border for each marker, as shown in Figure 1. This internal matrix consists of ones and zeros, which define the unique identifier of the marker—black squares represent 0, while white squares represent 1, as illustrated in Figure 8. The marker border is black and is equivalent to one bit in width. Of the 100 possible markers that the library can generate, 20 are utilized in this study.



0	0	0	0	0	0	0
0	1	0	1	1	1	0
0	1	1	0	0	1	0
0	0	1	1	1	0	0
0	1	1	1	0	0	0
0	0	0	1	1	0	0
0	0	0	0	0	0	0

Figure 8: Numbered and Segmented Visualization of Black and White Squares.

Before detecting markers, calibration is performed to extract camera parameters. This process utilizes a chessboard pattern for camera calibration. A large chessboard with dimensions of 8 × 6 squares and an inner square size of 35 mm is used. To ensure accurate calibration, the chessboard is positioned at various locations within the camera frame, as illustrated in Figure 9. Upon completion of this process, the camera calibration matrix and distortion coefficients are calculated. The calibration matrix includes the focal lengths of the camera lens and the position of the projection center. These parameters facilitate the translation of the 2D coordinates of an object in the captured image into its corresponding 3D coordinates in the real world. Furthermore, the distortion coefficients correct for the geometric distortions introduced by the camera optics in the captured images.

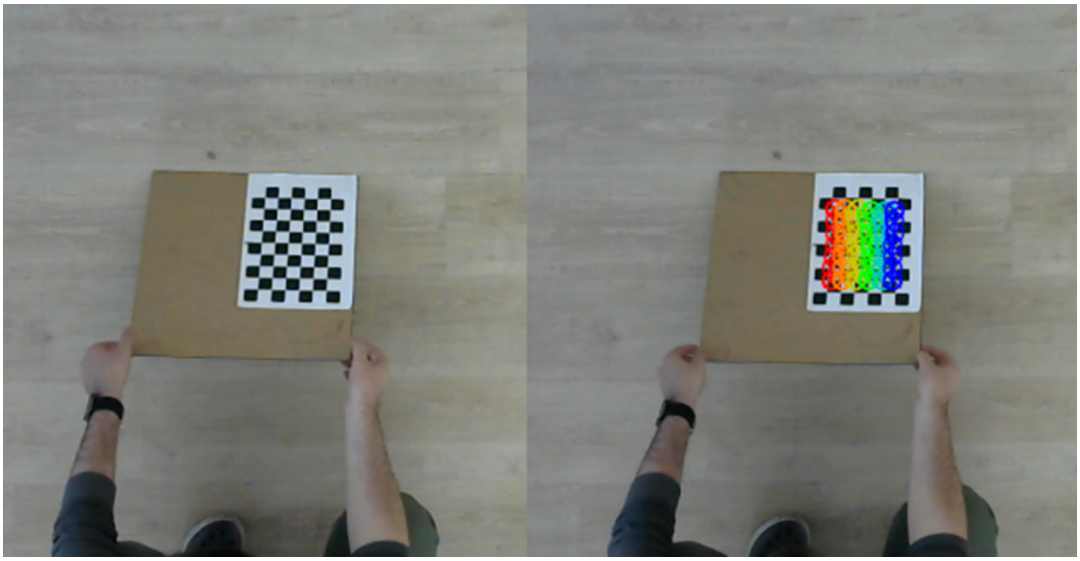


Figure 9: Calibration Process Using a Chessboard.

For a camera mounted at a height of 3 meters, a resolution of 1280 × 720 is adequate to track markers within the workspace. At this resolution, the numerical identifier of each marker is clearly visible, with the X-axis represented in red and the Y-axis in green, as illustrated in Figure 10. Identification of an ArUco tag involves several transformation steps applied to the original image. Initially, the image was converted to grayscale. Next, a binarization process is performed in which each pixel is converted to black or white based on its intensity. Finally, a contour analysis is performed to identify and isolate objects with a defined pattern. The identified marker is rendered in white, while the background remains black.

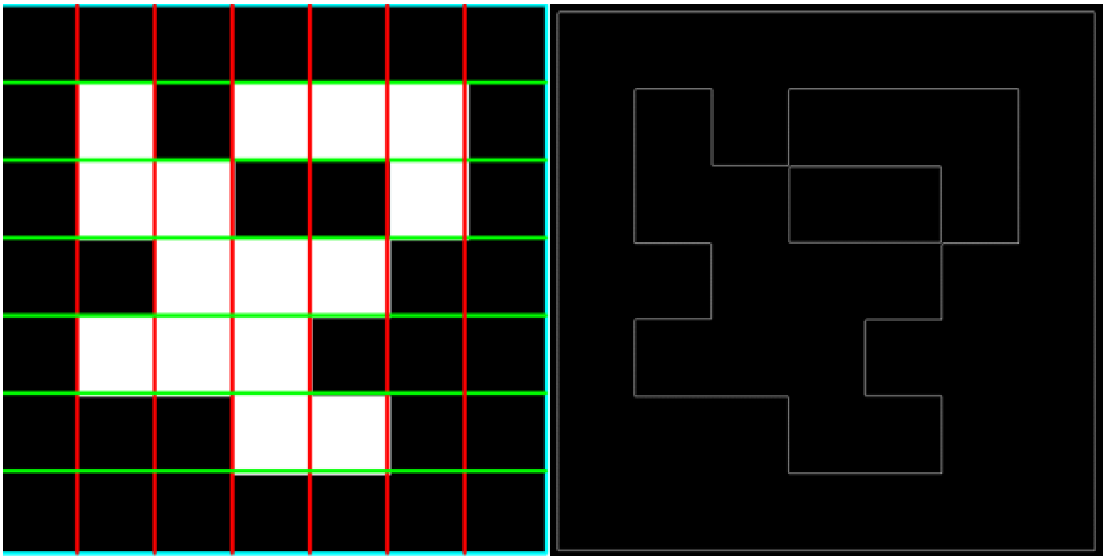
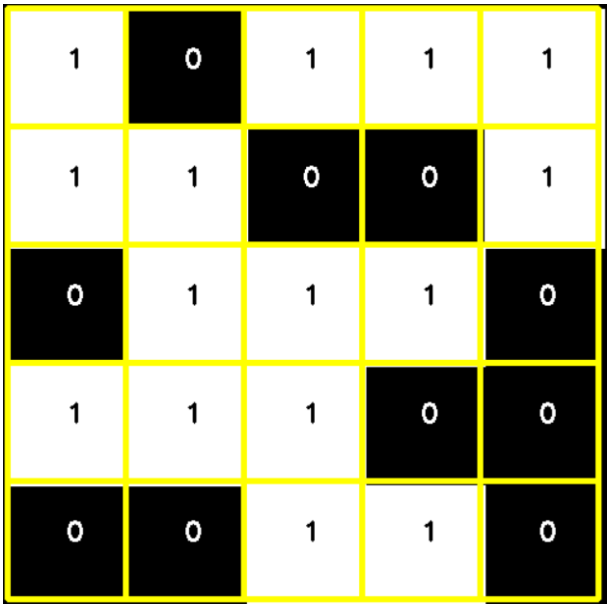


Figure 10: Segmented ArUco with Vertical-Horizontal Gridlines and Identified Borders.

Subsequently, polygon approximations are derived from the identified contours, and a predefined 5 × 5 grid is generated for the binary encoding analysis, as illustrated in Figure 11. In this grid, the black and white areas are assigned values of 0 and 1, respectively, enabling the determination of the marker's ID in matrix format.



1	0	1	1	1
1	1	0	0	1
0	1	1	1	0
1	1	1	0	0
0	0	1	1	0

Figure 11: Visualization with 5×5 Grid Applied.

Using this matrix, the system can determine the center position and orientation of the marker relative to the working plane, as illustrated in Figure 12.

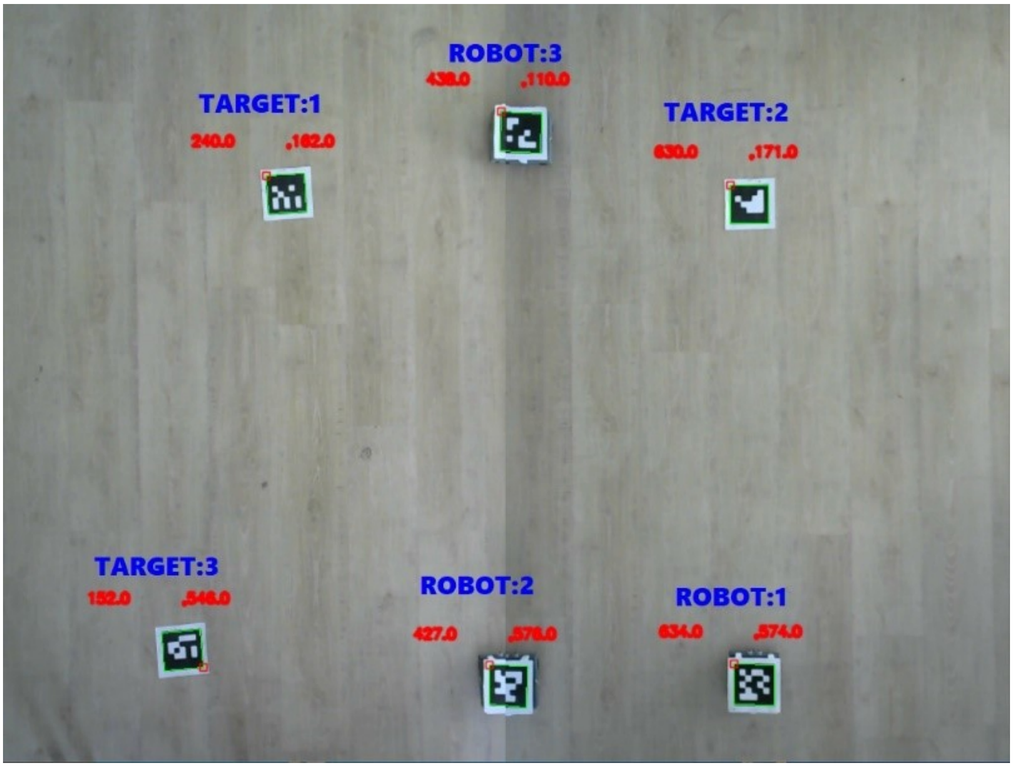


Figure 12: Positioning of Robots and Targets Using ArUco Codes by Cameras.

The pixel-based positional data of the center points of the identified robots and targets have been obtained. Using this information, the distance between the robots and their targets can be calculated employing the Euclidean formula:

$$\text{Distance} = \sqrt{(x_{\text{target}} - x_{\text{robot}})^2 + (y_{\text{target}} - y_{\text{robot}})^2}. \tag{15}$$

Here, x_{robot} and y_{robot} represent the coordinates of the current position of the robot, while x_{target} and y_{target} denote the coordinates of the target position. This calculation determines the linear distance from the robot's current position to the target position. The orientation angle of the robot relative to the target is calculated as follows:

The vector $\mathbf{r}_1$, representing the orientation towards the target, is defined as:

$$\mathbf{r}_1 = (x_{\text{target}} - x_{\text{current}}, y_{\text{target}} - y_{\text{current}}). \tag{16}$$

The vector $\mathbf{r}_2$, representing the robot's current orientation, is defined as:

$$\mathbf{r}_2 = (x_{\text{front}} - x_{\text{current}}, y_{\text{front}} - y_{\text{current}}). \tag{17}$$

The closure angle θ between the two vectors $\mathbf{r}_1$ (orientation towards the target) and $\mathbf{r}_2$ (current orientation of the robot) is calculated using the dot product formula:

$$\theta = \arccos\left(\frac{\mathbf{r}_1 \cdot \mathbf{r}_2}{\|\mathbf{r}_1\|\|\mathbf{r}_2\|}\right). \tag{18}$$

Here, x_{front} and y_{front} represent the position coordinates of the front part of the robot. The calculated angular error indicates the necessary angular correction to enhance the robot's orientation accuracy. If the angle is expressed in degrees:

$$\text{Angular Error} = \theta \times \frac{180}{\pi}. \tag{19}$$

These calculations have been used to determine the positional and orientation errors required for the robot to accurately and efficiently reach the target position. Once the robots' positioning within the system is finalized, communication between the main controller and the robots is established through Wi-Fi signals. The general operational block diagram of the system is illustrated in Figure 13.

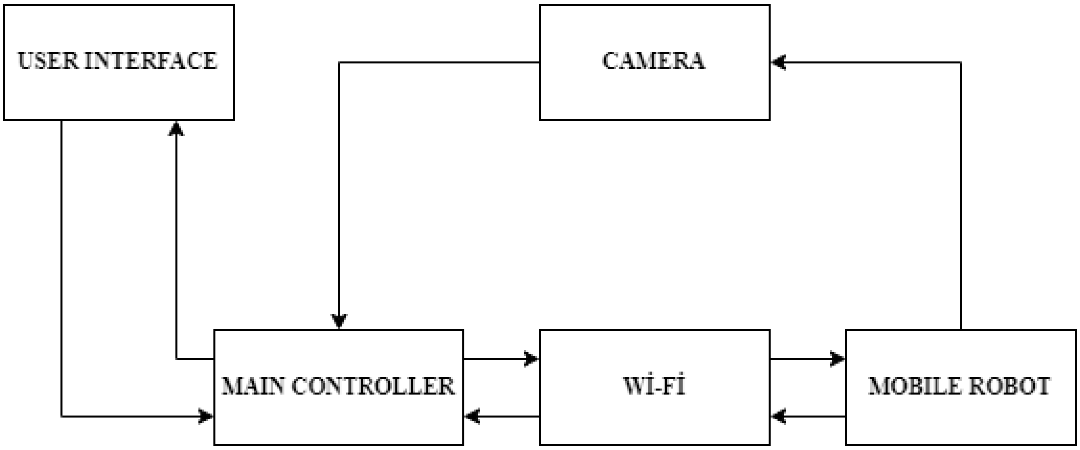


Figure 13: System Block Diagram.

As illustrated in the block diagram, the positions of the robots and targets are determined by the main controller using data from the cameras. The user interface allows the operator to select which robot should approach which target. Subsequently, movement commands are transmitted to mobile robots via the Wi-Fi communication protocol. An internal speed control mechanism has been designed and implemented for each mobile robot, utilizing two PID controllers, one for each wheel. This setup enables speed adjustments based on the distance and angular differences between the robots and their targets, ensuring that the robots accurately reach the desired positions. The flow chart of the system's operational algorithm is presented in Figure 14

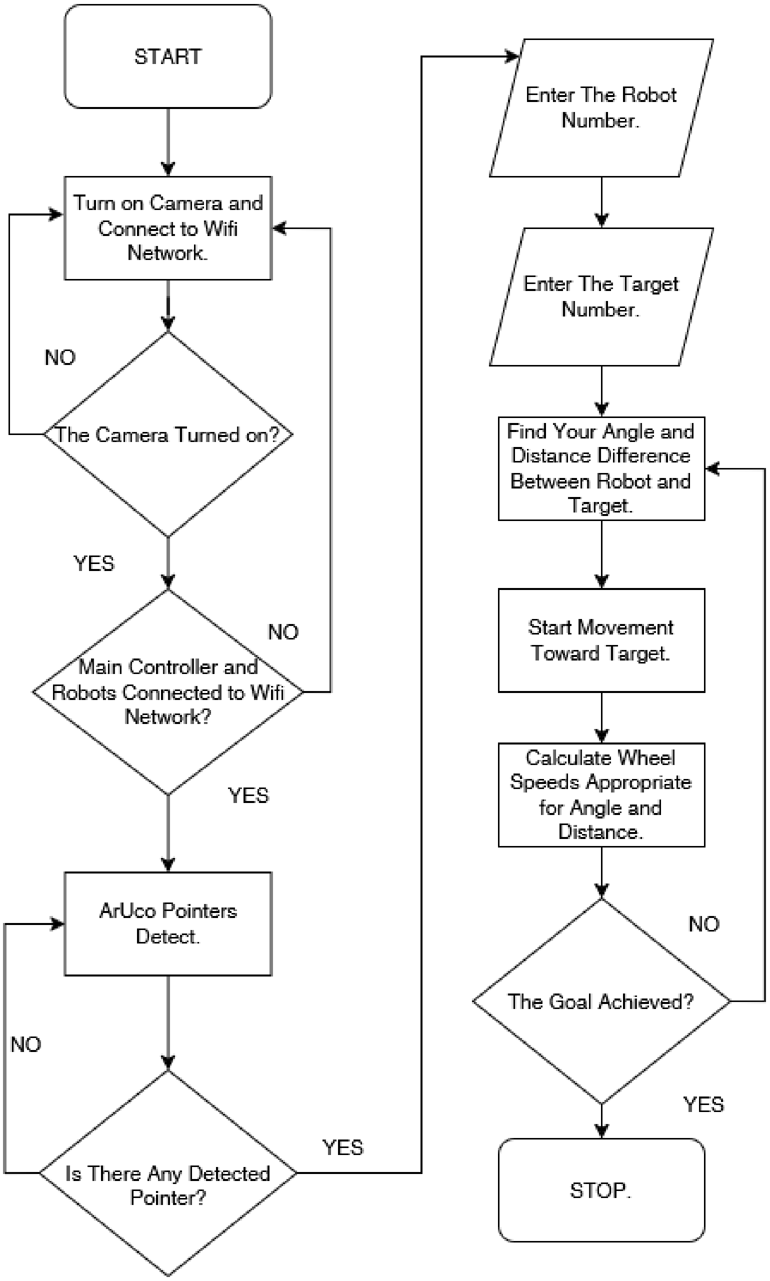


Figure 14: Algorithm Flowchart.

3 Results and Discussion

Experimental studies are conducted to evaluate the accuracy of robots in reaching predefined targets. In this context, the movement of the selected robot towards a designated target, as specified through the user interface, is examined in detail. The graphical user interface (GUI), developed using Python, significantly enhances the execution, monitoring, and control of the experiments. The system identifies ArUco markers in the environment, detects their positional information, and records the data on a centralized main controller. The data transmitted to the main controller integrate seamlessly with the robots' control software, ensuring continuous access to the position and orientation information of the markers.

During the experiment, the distances and angular differences between the robots and their targets are analyzed in detail. In Figure 12, a real-time snapshot of three different robots and their designated targets. Through these images, robots' positions, orientations, and distances from their targets are continuously monitored and analyzed. Based on these data, speed and directional adjustments are made according to the robots' kinematic models, ensuring that each robot reaches its target efficiently. Once the robots successfully reach their targets, the system marks the tasks as completed, concluding the process. Such systems demonstrate significant potential, particularly in fields such as industrial robotics and autonomous vehicles. In this study, two experimental scenarios are designed. The first scenario involves one robot and three targets within the field of view of the camera. This experiment is carried out 30 times for each robot, resulting in a total of 90 trials. A snapshot of the robot reaching the first target is presented in Figure 15.

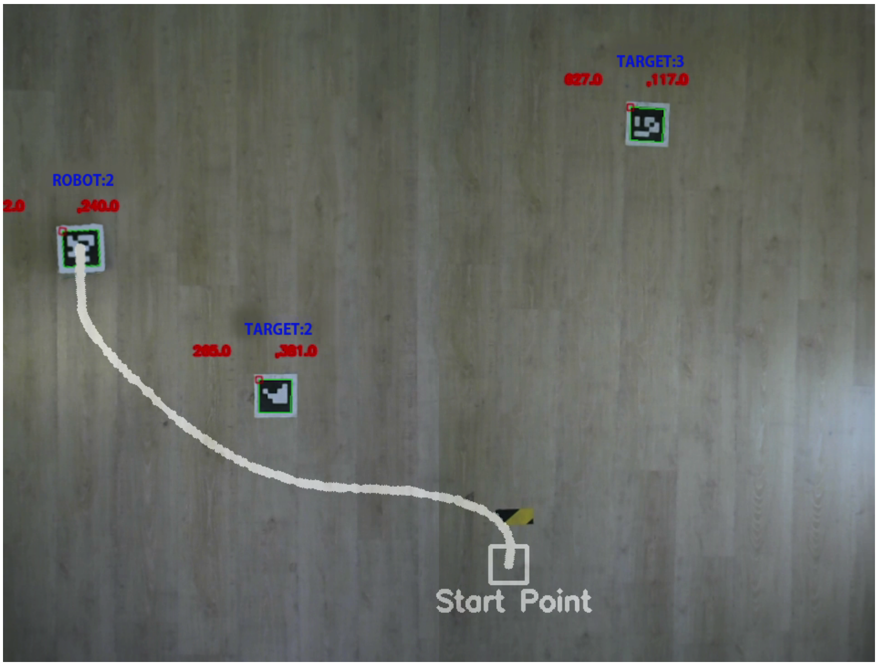


Figure 15: Experimental Scenario Where the Robot Reaches the First Target.

In the second experimental scenario, three robots are programmed to move toward three distinct targets. During this phase, a total of 30 trials are conducted. The robots are initially directed toward linear targets and then toward lateral targets in a mixed sequence, and the results are recorded. In the experimental setup that involves multiple robots and targets in the field, the robots are dispatched sequentially to their designated targets. The experimental results are illustrated in Figure 16.

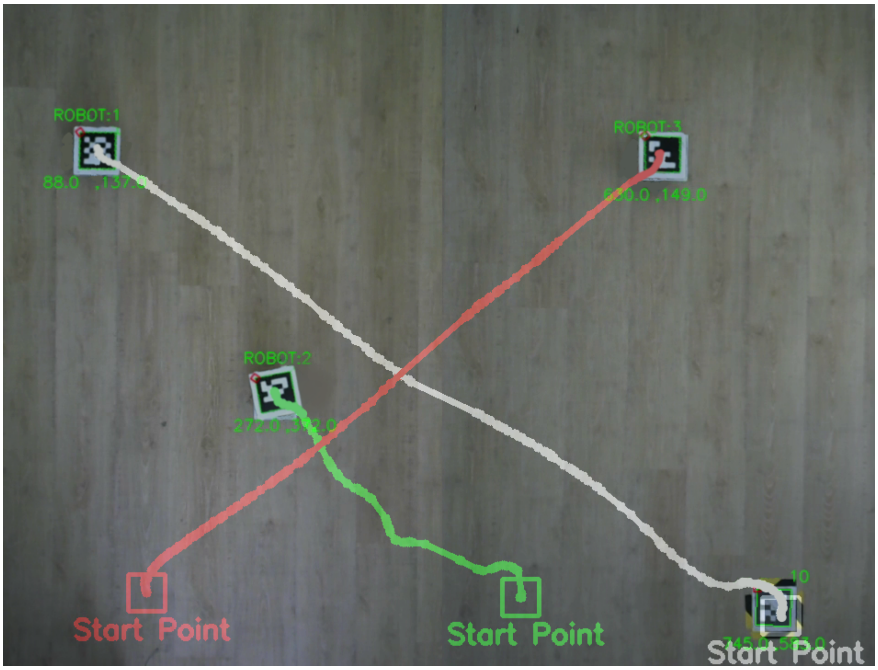


Figure 16: Experimental Scenario Where Robots Move Sequentially Toward and Reach Their Targets.

The developed positioning system effectively ensures that robots reach their targets in all experimental trials. Compared to other positioning systems documented in the literature, the proposed system is both cost-effective and capable of delivering more precise results. The findings of the experimental study are presented in Table 1.

Table 1: Error Rates of Experimental Studies (mm).

Experiment	Single Robot Experiments			Collective Robots		
	Robot 1	Robot 2	Robot 3	Robot 1	Robot 2	Robot 3
1	5.00	2.00	11.00	14.00	13.00	19.00
2	8.00	5.00	13.00	9.00	12.00	10.00
3	10.00	7.00	5.00	8.00	13.00	15.00
4	4.00	5.00	2.00	5.00	13.00	18.00
5	5.00	4.00	6.00	16.00	7.00	19.00
6	6.00	5.00	7.00	18.00	14.00	18.00
7	3.00	12.00	8.00	8.00	12.00	13.00
8	16.00	3.00	2.00	6.00	17.00	18.00
9	11.00	8.00	5.00	1.00	10.00	16.00
10	4.00	5.00	7.00	16.00	8.00	15.00
11	13.00	11.00	11.00	6.00	17.00	18.00
12	13.00	14.00	2.00	16.00	14.00	17.00
13	5.00	5.00	4.00	8.00	21.00	17.00
14	2.00	12.00	11.00	11.00	14.00	11.00
15	6.00	11.00	11.00	14.00	13.00	19.00
16	13.00	6.00	12.00	5.00	13.00	18.00
17	8.00	6.00	8.00	16.00	8.00	15.00
18	2.00	4.00	6.00	16.00	14.00	17.00
19	11.00	2.00	5.00	8.00	12.00	13.00
20	7.00	7.00	5.00	18.00	14.00	18.00
21	14.00	5.00	11.00	14.00	11.00	14.00
22	2.00	14.00	16.00	16.00	13.00	18.00
23	4.00	10.00	3.00	11.00	11.00	14.00
24	11.00	6.00	6.00	18.00	14.00	18.00
25	11.00	2.00	5.00	16.00	13.00	5.00
26	12.00	11.00	4.00	6.00	8.00	5.00
27	8.00	13.00	10.00	9.00	13.00	12.00
28	6.00	13.00	5.00	1.00	10.00	16.00
29	5.00	4.00	8.00	16.00	7.00	19.00
30	5.00	5.00	4.00	9.00	12.00	24.00
Single Average Error	7.70	7.20	7.10	11.20	12.40	15.60
Average Error		7.30			13.10	

The robot 1 positioning error graph after reaching the given target points in individual experimental trials is presented in Figure 17.

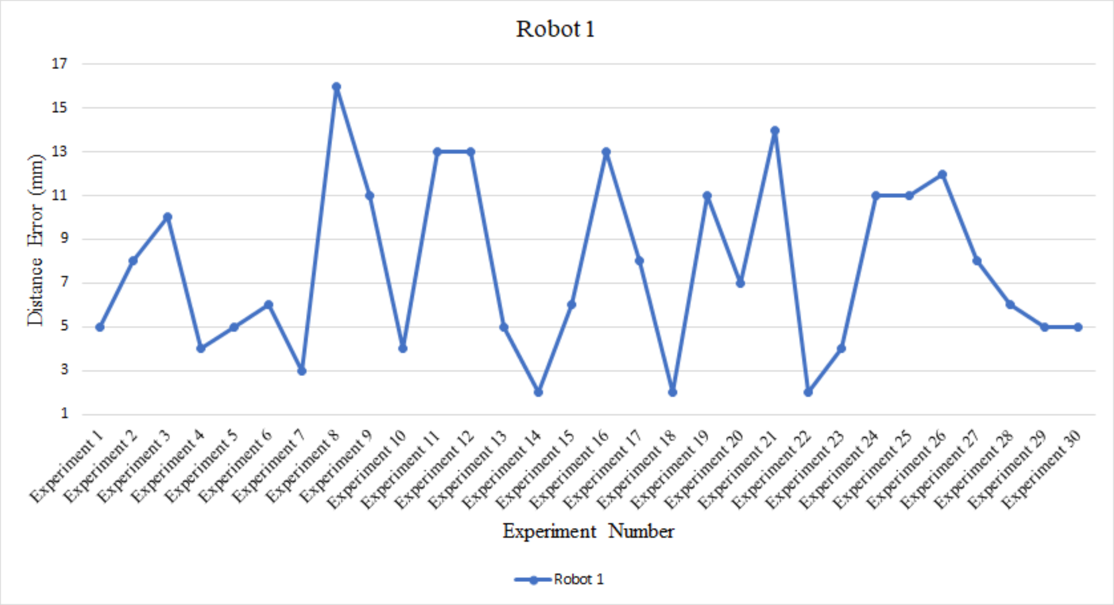


Figure 17: Positioning Error Rates of Robot 1.

The robot 2 positioning error graph after reaching the given target points in individual experimental trials is presented in

Figure 18.

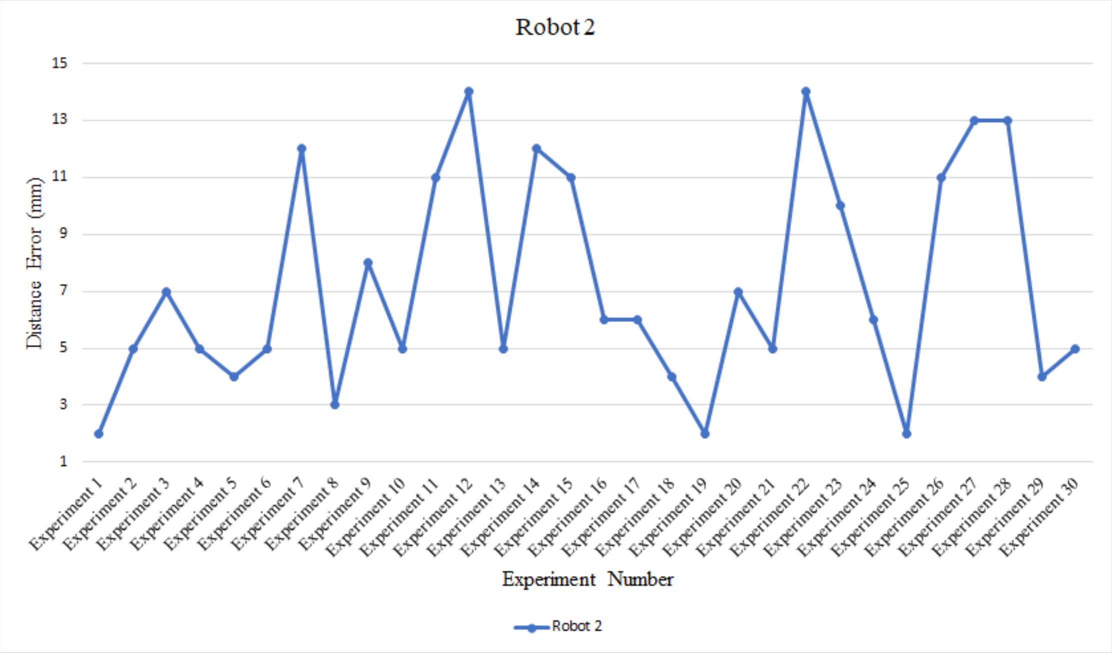


Figure 18: Positioning Error Rates of Robot 2.

The robot 3 positioning error graph after reaching the given target points in individual experimental trials is presented in Figure 19.

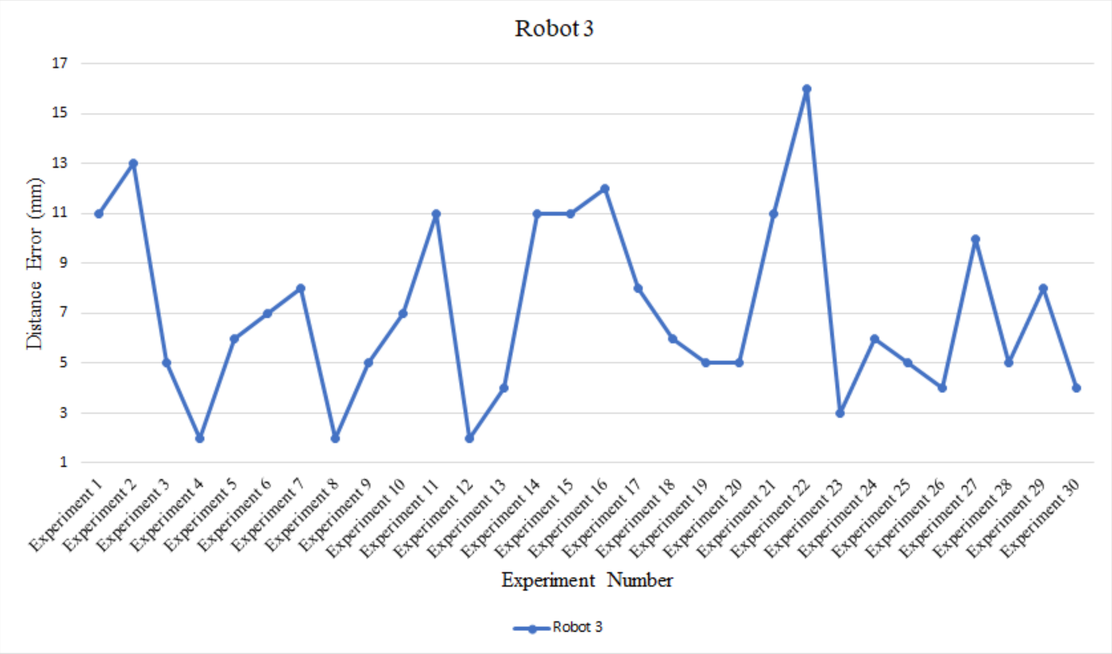


Figure 19: Positioning Error Rates of Robot 3.

The positioning error graph of all robots after reaching the given target points in multi-robot experimental trials is presented in Figure 20.

The experimental results indicate that the system positioning error averaged 7 mm in individual trials and 13 mm in

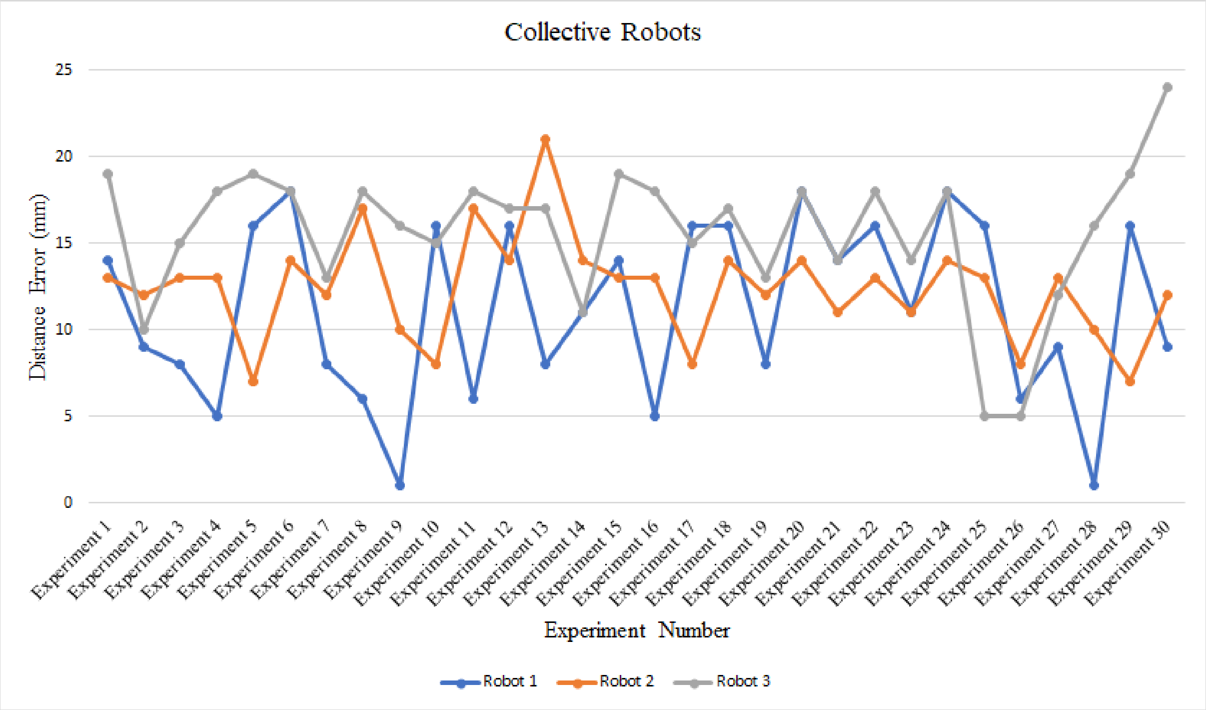


Figure 20: Positioning Error Rates of All Robots in the Multi-Robot Experimental Scenario.

multirobot trials. Considering that the average error rates in similar studies in the literature are generally above 100 mm, these findings demonstrate that our system exhibits significantly superior performance compared to its counterparts. In this context, we plan to continuously improve the components of the system, the testing environment, and the potential applications developed in this study. In particular, a multi-agent heterogeneous testing environment will be established to enable coordinated movement among robots. For this purpose, advanced robotic behaviors such as obstacle avoidance, consensus-based algorithms, and robotic formations will be implemented within the testing environment. Comprehensive tests will be conducted to improve the performance of the system and evaluate its effectiveness in various scenarios. Based on the data obtained, further enhancements will be made to the system. A comparison of similar studies in the literature with the system developed in this study is presented in Table 2.

Table 2: Comparison of Similar Studies in the Literature.

Method	Average Distance Error (mm)
Wi-Fi [18]	2310
Bluetooth [19]	15
IMU [20]	400
UWB [21]	40
Computer Vision [22]	30
Our Method	7

In conclusion, the exceptional performance and future development potential of the system developed in this study signify substantial advancements compared to existing research in the literature. Future investigations will concentrate on further optimizing the system and integrating it into new application domains. These efforts aim to empower robotic systems to function in a more intelligent, efficient, and harmonious manner.

Acknowledgments

This study includes the results of the Ph.D. thesis of the first author.

Authors' Contributions

Both authors contributed equally, read and approved the final manuscript.

Competing Interests

The authors declare that they have no competing interests.

References

[1] H. Sahib Hasan, M. Hussein, S. Mad Saad, and M. Azuwan Mat Dzahir, “An Overview of Local Positioning System: Technologies, Techniques and Applications,” *International Journal of Engineering & Technology*, vol. 7, no. 3.25, p. 1, Aug. 2018. DOI: 10.14419/ijet.v7i3.25.17459.

[2] A. Francis, A. Akinnusotu, and A. Adeyemi, “Global Positioning System and It’s Wide Applications,” *Continental J. Information Technology*, vol. 9, no. 1, pp. 22–32, Oct. 2015.

[3] N. Syazwani C. J., A. Wahab, N. Sunar, H. Syed, K. Y. Wong, and Y. Aun, “Indoor Positioning System: A Review,” *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 6, Jan. 2022. DOI: 10.14569/ijacsa.2022.0130659.

[4] S. Bastiaens, M. Alijani, W. Joseph, and D. Plets, “Visible Light Positioning as a Next-Generation Indoor Positioning Technology: A Tutorial,” *IEEE Communications Surveys & Tutorials*, vol. 26, no. 4, pp. 2867–2913, Fourth Quarter 2024. DOI: 10.1109/COMST.2024.3372153.

[5] X. Zhu, W. Qu, T. Qiu, L. Zhao, M. Atiquzzaman, and D. O. Wu, “Indoor Intelligent Fingerprint-Based Localization: Principles, Approaches and Challenges,” *IEEE Communications Surveys & Tutorials*, vol. 22, no. 4, pp. 2634–2657, Fourth Quarter 2020. DOI: 10.1109/COMST.2020.3014304.

[6] X. Feng, S. Liu, Q. Yuan, J. Xiao, and D. Zhao, “Research on wheel-legged robot based on LQR and ADRC,” *Scientific Reports*, vol. 13, no. 1, p. 15122, Sep. 2023. DOI: 10.1038/s41598-023-41462-1.

[7] O. Marques, *Practical Image and Video Processing Using MATLAB*®, Hoboken, NJ, USA: John Wiley & Sons, Inc., 2011. DOI: 10.1002/9781118093467.

[8] A. Babinec, L. Jurišica, P. Hubinsky, and F. Duchon, “Visual Localization of Mobile Robot Using Artificial Markers,” *Procedia Engineering*, vol. 96, pp. 1–9, 2014. DOI: 10.1016/j.proeng.2014.12.091.

[9] G. Bradski and A. Kaehler, *Learning OpenCV*, O’Reilly Media, Inc., 2008.

[10] M. Fiala, “ARTag, a fiducial marker system using digital techniques,” *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005, pp. 590–596. DOI: 10.1109/CVPR.2005.74.

[11] E. Tola, “Real-Time UAV Pose Estimation and Tracking Using FPGA Accelerated April Tag,” RIT Digital Institutional Repository, 2021. URL: *https : //repository.rit.edu/theses/*10853.

[12] M. Nahangi, A. Heins, B. McCabe, and A. P. Schoellig, “Automated Localization of UAVs in GPS-Denied Indoor Construction Environments Using Fiducial Markers,” *Proceedings of ISARC*, Jul. 2018. DOI: 10.22260/isarc2018/0012.

[13] I. Ullah, D. Adhikari, H. Khan, M. Shahid Anwar, S. Ahmad, and X. Bai, “Mobile robot localization: Current challenges and future prospective,” *Computer Science Review*, vol. 53, Aug. 2024. DOI: 10.1016/j.cosrev.2024.100651.

[14] A. Sampathkrishna, “ArUco Marker-based localization and Node graph approach to mapping,” *arXiv.org*, Aug. 2022. URL:*https : //arxiv.org/abs/*2208.09355.

[15] J. Kim, Y. Jeong, H. Lee, and H. Yun, “Marker-Based Structural Displacement Measurement Models with Camera Movement Error Correction Using Image Matching and Anomaly Detection,” *Sensors*, vol. 20, no. 19, p. 5676, Oct. 2020. DOI: 10.3390/s20195676.

[16] H. Bin, L. W. Zhen, and L. H. Feng, “The Kinematics Model of a Two-Wheeled Self-Balancing Autonomous Mobile Robot and Its Simulation,” *Proceedings of the International Conference on Computer Engineering and Applications*, 2010, pp. 64–68. DOI: 10.1109/ICCEA.2010.169.

[17] W. Burger and M. J. Burge, “Scale-Invariant Feature Transform (SIFT)”, *Texts in Computer Science*, pp. 709–763, Jan. 2022. DOI: 10.1007/978-3-031-05744-1_25.

[18] W. Cui, Q. Liu, L. Zhang, H. Wang, X. Lu, and J. Li, “A robust mobile robot indoor positioning system based on Wi-Fi,” *International Journal of Advanced Robotic Systems*, vol. 17, no. 1, p. 172988141989666, Jan. 2020. DOI: 10.1177/1729881419896660.

[19] D. R. Philips, E. Salami, H. Ramiah, and J. Kanesan, “Location Accuracy Optimization in Bluetooth Low Energy (BLE) 5.1-Based Indoor Positioning System (IPS)—A Machine Learning Approach,” *IEEE Access*, vol. 11, pp. 140186–140201, 2023. DOI: 10.1109/ACCESS.2023.3338358.

[20] F. Jiang, D. Caruso, A. Dhekne, Q. Qu, J. J. Engel, and J. Dong, “Robust Indoor Localization with Ranging-IMU Fusion,” *arXiv*, Cornell University, Jan. 2023. DOI: 10.48550/arxiv.2309.08803.

[21] M. Faeik, S. Junginger, T. Roddelkopf, and K. Thurow, “UWB-Based Real-Time Indoor Positioning Systems: A Comprehensive Review,” *Applied Sciences*, vol. 14, no. 23, pp. 11005–11005, Nov. 2024. DOI: 10.3390/app142311005.

[22] S. Roos-Hoefgeest, I. A. Garcia, and R. C. Gonzalez, “Mobile robot localization in industrial environments using a ring of cameras and ArUco markers,” *Proceedings of IECON 2021-47th Annual Conference of the IEEE Industrial Electronics Society*, Toronto, ON, Canada, 2021, pp. 1–6. DOI: 10.1109/IECON48115.2021.9589442.

El Cezeri
Journal of Science and Engineering
ISSN: 2148-3736

Research Article

Real-Time Detection of Vehicle Queue States in Urban Traffic Using Deep Learning

Ahsen Battal[1a], Yunus Emre Avci[1b], Adem Tuncer[2c]
[1] Dept. of Computer Engineering, Institute of Graduate Studies, Yalova University, Yalova, Türkiye
[2] Dept. of Computer Engineering, Engineering Faculty, Yalova University, Yalova, Türkiye

battalahsen@gmail.com

DOI : 10.31202/ecjse.1755333
Received: 01.08.2025Accepted: 16.09.2025
How to cite this article:
Ahsen Battal, Yunus Emre Avci, Adem Tuncer, "Real-Time Detection of Vehicle Queue States in Urban Traffic Using Deep Learning", El-Cezeri Journal of Science and Engineering, Vol: 12, Iss: 3, (2025), pp.(356-364).
ORCID: [a]0000-0002-4824-5889; [b]0000-0003-3921-7162; [c]0000-0001-7305-1886.

Abstract : Traffic congestion and vehicle queue formation at signalized intersections represent critical challenges in modern urban transportation systems, requiring accurate real-time detection methods for effective traffic management. This study presents a deep learning-based approach for real-time vehicle queue state classification that integrates You Only Look Once (YOLO) object detection with Simple Online Real-time Tracking (SORT) algorithms using standard traffic camera footage. The proposed system performs multi-class vehicle classification, real-time vehicle tracking with unique ID assignment, and speed estimation through camera calibration techniques, achieving 16.42 FPS average processing speed across diverse video scenarios. A comprehensive queue state detection methodology is developed that categorizes traffic conditions into three categories: Heavy traffic, stable flow, and free flow based on the analysis of average speeds of the detected vehicles, excluding motorcycles and bicycles due to their distinct traffic behavior patterns. Experimental validation across several test datasets encompassing both high and low resolutions demonstrates robust vehicle detection performance across all vehicle classes. Speed estimation accuracy ranges from 89% to 99%, validated against vehicle counting and tracking in designated traffic lanes, providing essential data for queue analysis. The system achieves vehicle counting accuracy ranging from 78.57% to 100% across different scenarios. The system offers a cost-effective alternative to traditional sensor-based methods by utilizing existing traffic-surveillance infrastructure, making it suitable for widespread deployment in intelligent transportation systems. Results indicate the proposed approach successfully detects queue states in real-time conditions across diverse traffic scenarios, from heavy congestion to free flow conditions. This research advances computer vision-based traffic monitoring by demonstrating the practical effectiveness of integrated object detection and tracking algorithms, contributing to improved traffic flow optimization and congestion management.

Keywords : Deep learning, YOLO, SORT, traffic queue detection, speed detection, vehicle counting

1 Introduction

Traffic congestion has become a critical global challenge in urban areas, causing significant economic, environmental, and social impacts. The accurate measurement of vehicle queues at signalized intersections represents a fundamental challenge in modern Intelligent Transportation Systems (ITS), with direct implications for traffic flow optimization and congestion management. As urban traffic volumes continue to escalate globally, the demand for precise, cost-effective, and real-time traffic state detection systems has intensified significantly. As more people move to urban areas, the need for smart traffic solutions becomes even more urgent. Traditional queue detection methods including inductive loop sensors, manual observation, and basic computer vision algorithms suffer from significant limitations; high infrastructure, limited spatial coverage, poor real-time performance, and inadequate accuracy under varying environmental conditions. ITS are proving to be one of the best ways to fix these issues by tackling the root causes of traffic and finding solutions. Many studies show how effective ITS can be in detecting traffic jams. One of crucial elements in traffic management is the identification of queuing, which directly informs applications like the estimation of the level of service and the optimization of traffic signal control.

Recent advances in deep learning have shown significant promise for queue length estimation in traffic. A notable study [1] developed a hybrid Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) to accurately estimate queue lengths at intersections, achieving significant error reduction compared to traditional methods. The CNN components excel at spatial feature extraction from traffic images, while LSTM networks capture temporal dependencies in traffic flow patterns. Additionally, LSTM was used to estimate short-term arrival models and long-term traffic demand trends for lane utilization rate forecasting. Compared to other studies, the developed integrated deep learning model demonstrated high performance in estimating queue lengths in individual lanes. However, it was inadequate for setting the duration, start time, and sequence of signal groups to within a few seconds or minutes for both isolated signalized intersections and area traffic. Traffic control and signal planning are known methods to mitigate traffic congestion and reduce delays. A major challenge in optimizing traffic signal scheduling is accurately predicting traffic conditions before the start of the next forecast cycle. In a study [2] utilizing real-time traffic data and forecasting cycles, LSTM was applied to forecast queue lengths for the upcoming cycle. Additionally, a sequential model-based optimization technique was implemented to prevent overfitting and to select optimal hyperparameters. Experiments using the traffic control system dataset aimed to estimate vehicle queue lengths only for straight movement. The study did not incorporate an adaptability optimization technique, focusing instead on fixed cycle times. Umair et al. [3] proposed a CNN-based approach for estimating vehicle queue length in urban traffic scenarios using low-resolution traffic videos. The queue length was estimated based on the total number of vehicles waiting at a signal, with stopped vehicles detected using Deep SORT-based object tracking. Due to the powerful and accurate CNN-based detection and monitoring, the estimated queue length using cameras was effective. The study conducted a comprehensive analysis of vehicle detection models, including YOLOv3, YOLOv4, YOLOv5, SSD, ResNet101, and InceptionV3, ultimately selecting YOLOv4 as the primary model due its superior accuracy and robustness.

Advanced sensor technologies have emerged as promising solutions for precise queue detection. A significant study [4] introduced a method using roadside LiDAR data, achieving an average accuracy of 98%. This method processes LiDAR data for real-time vehicle tracking and addresses issues like occlusion and package loss. However, several limitations were identified: the detection range of a single LiDAR sensor constraints queue length measurement, requiring additional sensors for longer queues, assumptions about vehicle speed and length could impact detection accuracy, and LiDAR performance may degrade under adverse weather conditions.

Probe vehicles, equipped with various technological tools to gather data from the road environment, are commonly used for calculating and estimating traffic capacity, density, and queue length on roads. These vehicles play a significant role in areas such as queue length and traffic volume estimation. In a study [5] focused on calculating traffic capacity and queue length, different Bayes-based approaches were developed using probe vehicles in each lane. These approaches, which estimate penetration and penetration rate (the ratio of probe vehicles to other vehicles on the highway), were used to calculate queue lengths for straight-going and right-turning lanes. The authors noted that the developed model achieved high prediction accuracy and could be utilized in traffic signal control. A more sophisticated nonparametric approach was developed that moves beyond traditional assumptions of random arrivals and parameter estimation. The main objective is to create straightforward, analytical, nonparametric models that estimate queue lengths at traffic signals on a cycle-by-cycle basis using partial queue observations from probe vehicles. A crucial feature is that this approach doesn't rely on assumptions about random arrivals or the need to estimate fundamental parameters such as market penetration rates or arrival rates. Its simplicity and comparable accuracy to more complex parametric models make it a valuable contribution to the field, particularly for real-time applications where primary parameters are difficult to obtain dynamically. However, its current limitations regarding oversaturated conditions and overflow queues highlight areas for future research [6].

Drone-based computer vision has shown particular promise for comprehensive traffic analysis. Zhou et al. [7] introduced a method for determining the tail profiles using high-resolution data from various sources. The study focused on three key components: signal status estimation, queue profile identification, and lane detection. The developed algorithms were validated using a real-world dataset collected by drones, with results indicating that the methodology effectively extracted tail profile information from the raw drone data. The pNEUMA dataset was employed in this validation process. Traditional computer vision approaches have also been explored, though with mixed results. Vector Auto Regression (VAR) model [8] was developed to estimate queue lengths in individual lanes for mixed traffic typical of developing countries. The model uses Passenger Car Units to account for different vehicle types and includes lane-changing behavior analysis. Using drone-collected data and time series analysis, the model achieved high accuracy (R-squared = 0.97, MAPE = 21.55%) by identifying arrival flow, discharge flow, and lane changes as key factors with a three-time lag dependency according to the authors. However, the model had limitations, including poor performance in zero-queue situations, occasional negative estimates, reduced accuracy for queues over 50 meters due to lane-changing effects, and signal timing proved unsuitable as an input variable. Integration of multiple data sources has emerged as a sophisticated approach to queue length estimation.

The recent increase in the deployment of License Plate Recognition (LPR) detectors has enabled the utilization of their data for more advanced applications, such as calculating queue lengths at intersections, inferring vehicular trajectories, and estimating overall traffic conditions and emissions [9]. Liu et al. [10] developed a Random Forest (RF) based real-time queue length estimation method utilizing Global Position System (GPS) and LPR data. GPS data provided vehicle stop positions, while the RF model was trained to predict these stopping positions using traffic flow features extracted from LPR data. The estimated stopping positions were also used to calculate the cyclic maximum length for each approach lane. Similarly, Zhan et al. [11] proposed a lane-based queue length estimation model utilizing LPR data. The model incorporates a Gauss-based interpolation method for each lane to reconstruct missing information for unrecognized or mismatched vehicles. A car tracking-based simulation is then applied to estimate real-time queue lengths for each lane, using vehicle arrival and departure information.

Advanced mathematical modeling approaches have been developed to leverage vehicle trajectory data. One study [12] focused on real-time queue length estimation at signalized intersections using vehicle trajectory data, employing shock wave theory to model the dynamic queue forming and dissipating processes without relying on signal timing, arrival patterns, or penetration rates. By identifying inflection points in vehicle trajectories, the method estimates stopping and discharging shock waves, providing robust queue length estimates under various conditions. However, it cannot estimate queue lengths without probe vehicles, and faces challenges when upstream disruptions affect traffic flow.

Al Okaishi et al. [13] proposed a queuing system consisting of two main steps: detecting queues using the square difference method to identify motion in target areas, followed by vehicle identification using the Single Shot Detector (SSD) algorithm when no movement is detected. The main limitation of this model is its instability; the queue length resets to zero with any detected motion at the front of the queue, necessitating a sufficient transition time to ensure continuous vehicle movement.

Conventional vehicle queue length detection methods typically depend on static sensors, manual counting, or basic algorithms that cannot provide the accuracy and real-time performance needed for contemporary traffic management systems. While existing approaches have shown promise, several limitations persist: sensor-based methods require expensive infrastructure and have limited coverage, existing computer vision approaches often lack real-time performance or comprehensive evaluation across diverse conditions, and most studies focus on either detection or tracking separately, without optimized integration.

This study addresses key gaps in the literature by introducing an accurate and cost-effective method for estimating traffic queue lengths using conventional traffic cameras. The proposed approach leverages a deep learning framework that combines YOLO for real-time vehicle detection with SORT for robust object tracking. By processing video data from existing camera infrastructure, the system enables real-time vehicle counting, speed estimation, and dynamic queue length measurement. This integrated solution provides essential insights for efficient traffic monitoring and intelligent traffic management. The results demonstrate the system's effectiveness and practical viability for real-world traffic management applications. The study begins with vehicle classification—categorizing vehicles into cars, motorcycles, trucks, buses, and bicycles—using labeled visual datasets, followed by queue state detection. The paper is organized as follows: Section 2 presents the methodologies for speed estimation and queue length detection. Section 3 details and discusses the experimental results. Finally, Section 4 concludes the paper by summarizing the key findings, highlighting their implications for traffic management, and proposing directions for future research.

2 Methodology

This study presents a comprehensive real-time queue detection system that integrates deep learning-based vehicle detection, multi-object tracking, and speed estimation algorithms. The proposed methodology consists of six main stages: video stream processing, frame extraction, vehicle detection and classification, lane-based tracking, speed estimation, and queue state analysis. The system utilizes YOLOv5 for object detection, SORT algorithm for vehicle tracking, and implements a rolling window approach for queue state determination based on estimated vehicle speeds.

The methodology begins with video stream processing and frame extraction, followed by lane boundary definition using predefined reference points. Subsequently, vehicles are detected and classified using the YOLOv5 model, with each vehicle assigned a unique identifier through the SORT tracking algorithm. The system calculates individual vehicles' speeds using camera calibration and Euclidean distance measurements, finally determining queue states through average speed analysis within a rolling window framework. Average speeds of 20 km/h or lower indicate heavy traffic; speeds between 20 km/h and 55 km/h (inclusive) represent stable flow; and speeds above 55 km/h indicate free flow conditions. The Figure 1 illustrates the flowchart of the proposed queue state detection system.

2.1 Vehicle Speed Detection

YOLOv5m model [14] for robust vehicle detection and classification in the study. The model was trained on a comprehensive dataset combining multiple sources to achieve optimal performance in various traffic conditions. In our previous study [15], an initial dataset was created using several images from two different datasets [16], [17], targeting vehicle class detection and lane-based counting. The dataset was subsequently augmented with additional images from another source [18]. The model successfully identified five distinct vehicle classes: car, motorcycle, truck, bus, and bicycle. Performance evaluation across three different training configurations is presented in Table 1 with optimal results achieved at 150 epochs, demonstrating the model's effectiveness in multi-class vehicle recognition under diverse traffic scenarios.

Following vehicle classification, the system implements lane-based vehicle tracking using the SORT algorithm [19]. SORT employs a Kalman filter for state estimation, using a constant velocity motion model, and associates detections to tracks via

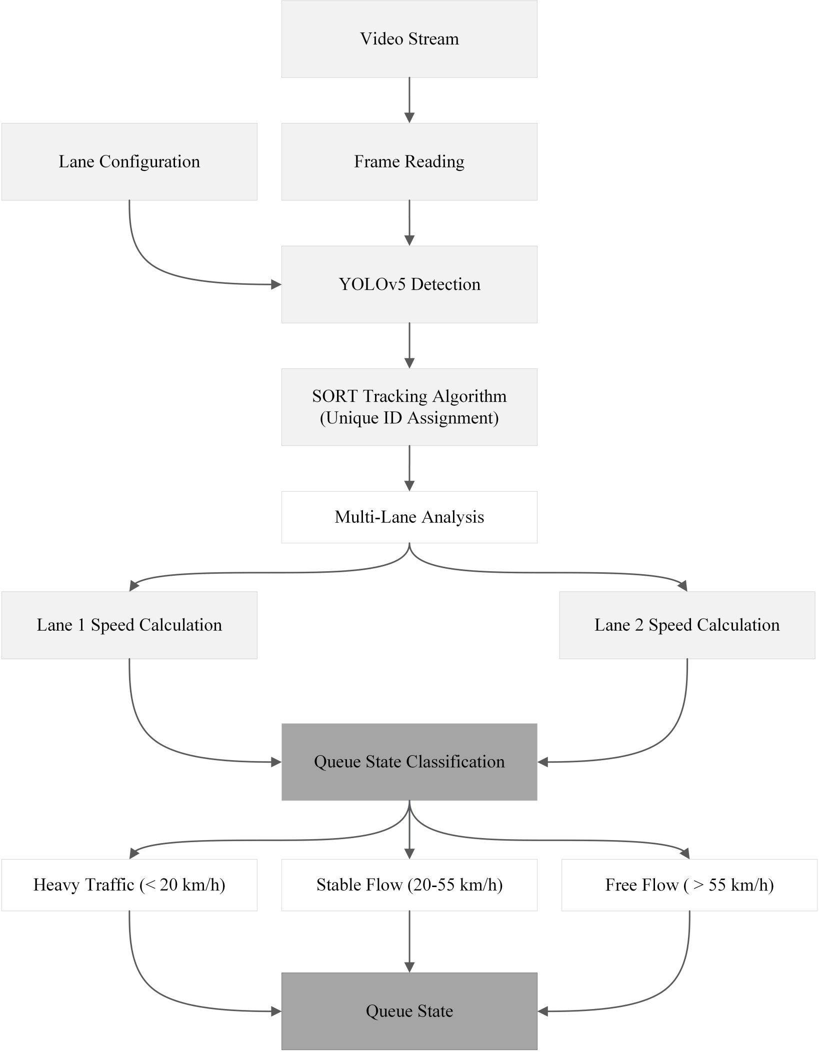


Figure 1: Flowchart of the proposed queue state detection system

Table 1: Performance results from the training dataset

Epoch	Precision	Recall	mAP
80	0.86	0.91	0.93
100	0.94	0.85	0.92
150	0.86	0.87	0.93

the Hungarian algorithm with an intersection over union (IoU) threshold. Each detected vehicle receives a unique identifier, enabling continuous tracking across frames.

Speed estimation plays a crucial role in queue state detection. As illustrated in Figure 2, several key parameters are extracted before speed calculation. These include the initial and final center point coordinates of the vehicle within the region of interest, the pixel-based length of the vehicle at its initial and final positions (with only the initial length used in the computation), and the timestamps of the vehicle's first and last detection in the video. Using these parameters, camera calibration—based on the method by [20]—is applied to convert pixel measurements into real-world distance, enabling the calculation of the actual distance traveled by the vehicle in meters. This calibrated distance, combined with the time difference, is then used to estimate the vehicle's speed.

$$k = \frac{L_{actual}}{L_{pixel}} \tag{1}$$

where k represents the camera calibration factor, L_{actual} is the real-world vehicle length (meters), and L_{pixels} is the pixel-based vehicle lengths. The system utilizes class-specific average vehicle lengths to enhance calibration accuracy across different

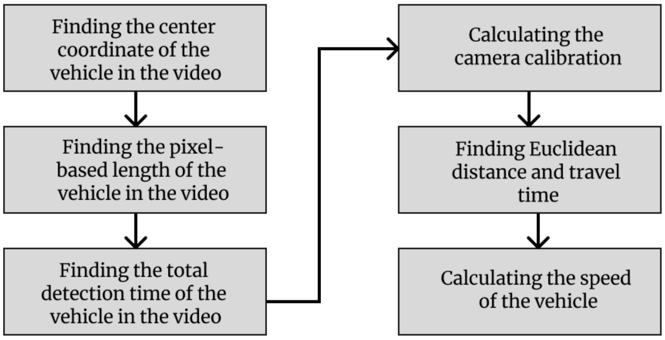


Figure 2: Vehicle speed detection diagram

vehicle types. In the study, the assumed average lengths were 4.5 meters for cars, 8.1 meters for trucks, 9.5 meters for buses [21].

$$d_{pixel} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{2}$$

Once the camera calibration is computed, the initial and final center point coordinates are used to calculate the distance the vehicle travels throughout the video. The Euclidean formula in Equation (2) was used to find this length. Since the result of this calculation is in pixels, it is converted to meters by multiplying it by the camera calibration as given in Equation (3).

$$d_{meter} = k \times d_{pixel} \tag{3}$$

In the speed estimation phase, the time taken by the vehicle is as important as the distance taken. During this period, the total movement time is calculated as the difference between the timestamps of the vehicle's initial and final detections as given in Equation (4).

$$t = t_2 - t_1 \tag{4}$$

After computing all necessary parameters, the estimated speed is determined by dividing the distance traveled by the elapsed time. Since the initial speed is calculated in meters per second (m/s), it is converted to kilometers per hour (km/h) by multiplying by 3.6 as given in Equation (5).

$$v = (\frac{d_{meter}}{t} \times 3.6) \tag{5}$$

Table 2: Results on speed detection obtained from a dataset consisting of videos recorded by the authors

Video	Actual Speed (km/h)	Estimated Speed (km/h)	Accuracy
Vehicle-1	30	30.83	%97.23
Vehicle-2	50	46.51	%93.02
Vehicle-3	70	70.60	%99.14
Vehicle-4	100	89.11	%89.11

The speed estimations were validated using recorded videos in which the actual speeds of the vehicles were known. These estimated speeds, measured in kilometers per hour, are compared with the actual speeds in Table 2. For instance, a vehicle traveling at an actual speed of 30 km/h was estimated to be moving at 30.83 km/h. Additional tests demonstrated accuracy rates of 97.23%, 93.02%, 99.14%, and 89.11%, respectively.

2.2 Queue State Detection

The approximate speeds of each vehicle are calculated and recorded for queue state detection. The traffic state $Q(t)$ is determined based on the average speed of vehicles in the detection zone and is defined in Equation (6). If fewer than three vehicles are detected (or no vehicles are present), insufficient data is available for reliable queue state assessment, and the system defaults to classifying the traffic state as free flow.

$$Q(t) = \begin{cases} \text{Heavy traffic,} & \text{if } 0 \text{ km/h} \leq \bar{v}(t) \leq 20 \text{ km/h} \\ \text{Stable flow,} & \text{if } 20 \text{ km/h} < \bar{v}(t) \leq 55 \text{ km/h} \\ \text{Free flow,} & \text{if } 55 \text{ km/h} < \bar{v}(t) \end{cases} \tag{6}$$

where $\bar{v}(t)$ denotes the average speed of detected vehicles at time t.

After vehicle detection, counting, and speed estimation, the system proceeds to classify the traffic queue state. This process employs a rolling window approach, where the average speed of the three most recently detected four-wheeled vehicles is calculated as given in Equation (7) to determine the queue state. The window then shifts by excluding the earliest detected vehicle and incorporating the next one, calculating the average speed of the updated set (i.e., the second, third, and fourth vehicles) to reassess the queue condition as given in Equation (8).

$$\bar{v}(t) = \frac{1}{3}\sum_{i=1}^{3} v_i \tag{7}$$

$$\bar{v}(t+1) = \frac{1}{3}\sum_{i=2}^{4} v_i \tag{8}$$

Two-wheeled vehicles like motorcycles and bicycles are not included in traffic queue analysis for both theoretical and practical considerations. Unlike larger vehicles, these two-wheelers follow different speed and traffic flow patterns and use lanes differently. They can weave between stopped cars through lane-splitting and filtering, which means they don't follow the standard first-come, first-served queue rules that apply to cars and trucks. By leaving out motorcycles and bicycles from queue analysis, researchers can better measure the actual congestion levels that impact the main flow of traffic. The system identifies traffic queues by calculating the average speed of vehicles traveling in specific lanes and uses this data to assess queuing conditions. For all these reasons, two-wheeled vehicles were not included in the queue state analysis in the study.

3 Experimental Results

The comprehensive evaluation of the proposed traffic queue detection system was conducted across 11 video sequences, encompassing diverse traffic scenarios ranging from free-flowing conditions to severe congestion. The evaluation framework employs multiple performance metrics to assess system effectiveness across three primary dimensions: computational efficiency, detection accuracy, and queue classification performance.

Table 3: Dataset Specifications and Performance Metrics

Video	Resolution	Duration(s)	FPS	YOLO Detection(ms)	SORT Tracking(ms)
Video-1	1080x1920	127.1	14.36	44.47	0.63
Video-2	1080x1920	28.4	14.39	44.58	0.68
Video-3	480x848	43.6	21.81	41.34	0.40
Video-4	480x848	78.4	21.42	42.14	0.43
Video-5	480x848	76.9	21.70	41.62	0.40
Video-6	1080x1920	51.6	14.59	43.5	0.40
Video-7	1080x1920	74.8	14.56	43.24	0.38
Video-8	1080x1920	43.2	14.45	44.33	0.88
Video-9	1080x1920	84	13.99	45.88	0.53
Video-10	1080x1920	51	14.65	44.37	0.63
Video-11	1080x1920	45	14.78	45.92	0.52

As presented in Table 3, the experimental dataset consists of 11 videos with varying specifications to evaluate the robustness of the system under different recording conditions. The dataset includes both high-resolution videos (1080×1920 resolution) and low-resolution videos (480×848 resolution), with durations ranging from 28.4 to 127.1 seconds. Frame rates varied between 13.99 and 21.81 FPS, providing comprehensive coverage of typical traffic surveillance scenarios. The computational performance analysis reveals optimal resource allocation, with YOLOv5 detection requiring an average of 44.2 ms per frame, while SORT tracking operations consumed only 0.52 ms on average. This distribution demonstrates that 98.8% of computational time is dedicated to object detection, 1.2% to tracking operations, indicating highly efficient algorithm integration. The system successfully processed videos at varying frame rates while maintaining consistent detection performance.

Equations (9) and (10) are used to calculate the accuracy of the estimated number of vehicles and speed calculation in vehicle counting [22].

$$Error(\%) = \left(\frac{|Detected\ Number - Actual\ Number|}{Actual\ Number}\right) \times 100 \tag{9}$$

$$Accuracy(\%) = 100 - Error(\%) \tag{10}$$

The SORT algorithm successfully tracked a total of 441 vehicles across all video streams, with individual video performance ranging from 15 to 84 vehicles. The system demonstrated consistent tracking capabilities across varying traffic densities, with lane-specific vehicle counts providing granular insights into traffic distribution patterns. In particular, the left-right lane distribution analysis revealed traffic flow imbalances in several scenarios, such as video 1 and video 9, indicating asymmetric traffic patterns commonly observed in real-world scenarios. The results demonstrate high accuracy in vehicle counting, generally achieving more than 90%. According to the Table 4, 45 vehicles were detected, although only 42 vehicles actually passed through the area in video 8. The system may overestimate the number of vehicles present in reality, as shown in video 8. Possible reasons for this include occlusion, where a vehicle in the frame is temporarily blocked by another vehicle; when the occluded vehicle reappears, the system fails to recognize it as the same vehicle and instead detects it as a new one. Additionally, factors such as lighting conditions, viewing angles, and partial visibility may prevent the correct re-identification of previously detected vehicles. Such issues (ID reassignments and challenges in multi-object trajectory tracking) are commonly reported in the literature [23].

Table 4: Results regarding vehicle counting

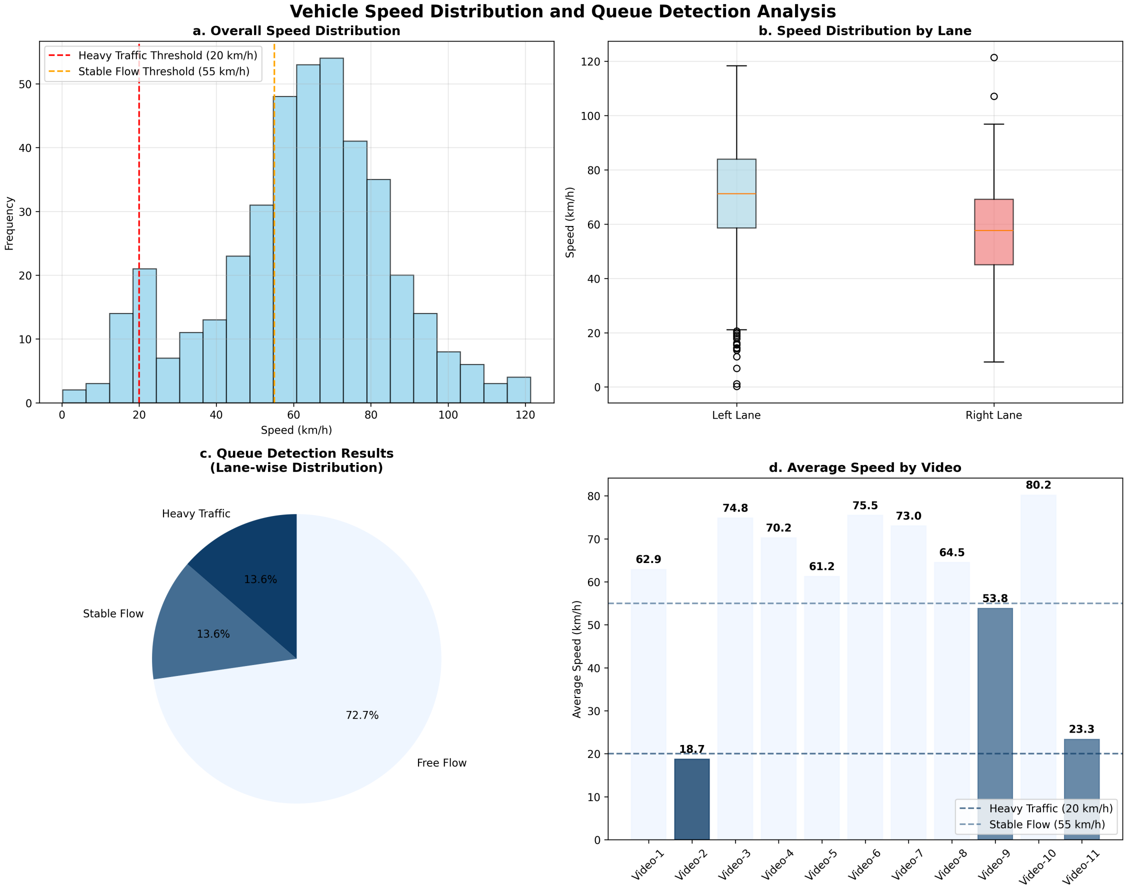
Video	Number of Detected Vehicles		Number of Actual Vehicles		Accuracy (%)	
	Right	Left	Right	Left	Right	Left
Video-1	38	46	37	47	97.29	97.87
Video-2	6	9	6	9	100	100
Video-3	14	8	14	9	100	88.88
Video-4	16	11	19	14	84.21	78.57
Video-5	16	21	17	21	94.11	100
Video-6	20	24	19	24	94.70	100
Video-7	23	29	22	30	95.65	96.67
Video-8	20	25	17	25	82.35	100
Video-9	21	30	22	30	95.45	100
Video-10	15	20	15	20	100	100
Video-11	18	10	19	11	94.73	90.90

Traffic flow analysis results are illustrated in Figure 3 from 11 test videos processed using the YOLOv5+SORT algorithm. The histogram of overall vehicle speed distribution across all videos, with queue classification thresholds indicated that heavy traffic threshold at 20 km/h and a stable flow threshold at 55 km/h, showing a normal distribution with a mean speed of 62.5 km/h and a standard deviation of 22.51 km/h. The distribution demonstrates the system's capability to capture the full spectrum of urban traffic conditions. The comparative box plot analysis of speed distributions between the left lane and right lane illustrates typical traffic flow patterns where the left lane maintains higher average speeds. The queue state classification results are presented as a pie chart, demonstrating system accuracy with balanced detection across three categories: heavy traffic (13.6%, 3 of 22 lanes), stable flow (18.19%, 4 of 22 lanes), and free flow (68.2%, 15 of 22 lanes).

Table 5: Vehicle Speed Analysis and Queue Detection Results by Video

Video	Right Lane Vehicles	Left Lane Vehicles	Right Avg Speed (km/h)	Left Avg Speed (km/h)	Right Queue State	Left Queue State
Video-1	38	46	55.76	68.59	Free flow	Free flow
Video-2	6	9	18.42	18.9	Heavy traffic	Heavy traffic
Video-3	14	8	62.24	93.65	Free flow	Free flow
Video-4	16	11	61.15	84.73	Free flow	Free flow
Video-5	16	21	58.37	62.94	Free flow	Free flow
Video-6	20	24	65.67	84.33	Free flow	Free flow
Video-7	23	29	64.41	79.12	Free flow	Free flow
Video-8	20	25	55.15	72.27	Free flow	Free flow
Video-9	21	30	52.41	54.66	Stable flow	Stable flow
Video-10	15	20	72.87	85.63	Free flow	Free flow
Video-11	18	10	29.62	12.04	Stable flow	Heavy traffic

The comprehensive speed analysis and queue state detection results are summarized in Table 5, showcasing the system's capability to perform real-time queue state classification. The speed values given in the table were determined by taking the average speed of all vehicles detected on a lane basis in a video. The three-category classification system successfully categorized traffic states. In video 1, the SORT algorithm successfully tracked 46 vehicles in the left lane and 38 in the right lane, yielding average speeds of 55.76 km/h and 68.59 km/h, respectively. Based on the established threshold criteria (speed $>$ 55 km/h), the queue detection result was classified as free flow. Video 2 demonstrated clear congestion conditions with average speeds of 18.42 km/h (right lane) and 18.9 km/h (left lane), correctly classified as heavy traffic for both lanes. Video 11 showed mixed conditions with the left lane experiencing heavy traffic (12.04 km/h) while the right lane exhibited stable flow (29.62 km/h). Multiple videos (video 1, video 3, video 4, video 5, video 6, video 7, video 8, video 10) exhibited high-speed conditions with average speeds exceeding 55 km/h, correctly classified as free flow. Notable examples include video 3 with exceptionally high speeds (62.24 km/h right, 93.65 km/h left) and video 10 (72.87 km/h right, 85.63 km/h left). Video 9 demonstrated intermediate traffic conditions with speeds of 52.41 km/h (right) and 54.66 km/h (left), appropriately classified as stable flow for both lanes.


Vehicle Speed Distribution and Queue Detection Analysis

a. Overall Speed Distribution

b. Speed Distribution by Lane

c. Queue Detection Results (Lane-wise Distribution)

d. Average Speed by Video

Figure 3: Vehicle Speed Distribution and Queue Detection Analysis

4 Conclusion

This study presents a validated, comprehensive real-time vehicle queue state detection system that integrates state-of-the-art deep learning methods with practical traffic management requirements. The proposed approach, which combines YOLOv5 for object detection with the SORT tracking algorithm, demonstrates exceptional performance across diverse traffic scenarios while maintaining the computational efficiency required for real-world deployment. The primary contributions of this work span several critical aspects of modern traffic monitoring systems. First, we introduce an integrated deep learning framework that seamlessly combines object detection and tracking components, specifically optimized for traffic queue analysis. Second, we propose a robust queue state classification methodology that accurately categorizes traffic conditions based on vehicle speed dynamics. Comprehensive experimental validation was conducted across multiple datasets, encompassing both high-resolution (1080×1920) and low-resolution (480×848) video streams, demonstrating the system's versatility and reliability under varying input conditions. The processing pipeline achieves efficient resource utilization: YOLOv5-based detection averages 44.2 ms per frame, while SORT tracking requires only 0.52 ms per frame. This results in a computational distribution where 98.8% of processing time is devoted to detection and just 1.2% to tracking, highlighting the efficiency of the overall algorithm design. Memory usage averaged 1050 MB across all test scenarios, confirming the system's feasibility for deployment in edge computing environments with limited computational resources. These characteristics make the proposed system well-suited for scalable, real-time traffic monitoring in urban and highway settings.

While the current implementation demonstrates robust performance across the evaluated scenarios, several areas for future studies remain. Although the dataset is comprehensive within its scope, it could be expanded to include adverse weather conditions, varying lighting environments, and diverse intersection geometries to further enhance the system's robustness and generalizability. Incorporating multiple camera viewpoints represents a promising direction, as it could enable more comprehensive spatial coverage and improve detection and tracking accuracy in complex, occlusion-prone traffic environments. Future studies should also focus on developing lightweight model architectures—such as through model pruning, quantization, or efficient network design—to reduce computational demands without compromising detection accuracy, thereby enabling deployment on lower-power edge devices. Additionally, integrating vehicle trajectory prediction and behavioral modeling could significantly enhance the system's intelligence, allowing it to anticipate traffic dynamics and support proactive traffic signal optimization. Such advancements would move the system beyond reactive monitoring toward adaptive, predictive traffic management.

Acknowledgments

This study was supported by the Research Fund of Yalova University (Project Number: 2023/AP/0002).

Authors' Contributions

In this study, AB, YEA, and AT proposed the study idea together. AB and YEA carried out the experimental studies, simulations, and wrote the initial draft. AT contributed to the study by supervising and interpreting. AB, YEA, and AT wrote the final version of the paper. All authors read and approved the final version of the paper.

Competing Interests

The authors declare that they have no conflict of interest.

References

[1] S. Lee, K. Xie, D. Ngoduy, and M. Keyvan-Ekbatani, "An advanced deep learning approach to real-time estimation of lane-based queue lengths at a signalized junction," *Transportation research part C: emerging technologies*, vol. 109, pp. 117–136, 2019.

[2] R. Rahman and S. Hasan, "Real-time signal queue length prediction using long short-term memory neural network," *Neural Computing and Applications*, vol. 33, pp. 3311–3324, 2021.

[3] M. Umair, M. U. Farooq, R. H. Raza, Q. Chen, and B. Abdulhai, "Efficient video-based vehicle queue length estimation using computer vision and deep learning for an urban traffic scenario," *Processes*, vol. 9, no. 10, p. 1786, 2021.

[4] J. Wu, H. Xu, Y. Zhang, Y. Tian, and X. Song, "Real-time queue length detection with roadside lidar data," *Sensors*, vol. 20, no. 8, p. 2342, 2020.

[5] Y. Zhao, J. Zheng, W. Wong, X. Wang, Y. Meng, and H. X. Liu, "Various methods for queue length and traffic volume estimation using probe vehicle trajectories," *Transportation Research Part C: Emerging Technologies*, vol. 107, pp. 70–91, 2019.

[6] G. Comert, T. Amdeberhan, N. Begashaw, N. G. Medhin, and M. Chowdhury, "Simple analytical models for estimating the queue lengths from probe vehicles at traffic signals: A combinatorial approach for nonparametric models," *Expert Systems with Applications*, vol. 252, p. 124076, 2024.

[7] Q. Zhou, R. Mohammadi, W. Zhao, K. Zhang, L. Zhang, Y. Wang, C. Roncoli, and S. Hu, "Queue profile identification at signalized intersections with high-resolution data from drones," in *2021 7th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, pp. 1–6, IEEE, 2021.

[8] S. Jayatilleke, V. Wickramasinghe, and N. Amarasingha, "Introduction of a simple estimation method for lane-based queue lengths with lane-changing movements," *Journal of The Institution of Engineers (India): Series A*, vol. 104, no. 1, pp. 143–153, 2023.

[9] M. Y. Arafat, A. S. M. Khairuddin, U. Khairuddin, and R. Paramesran, "Systematic review on vehicular licence plate recognition framework in intelligent transport systems," *IET intelligent transport systems*, vol. 13, no. 5, pp. 745–755, 2019.

[10] D. Liu, C. An, M. Yasir, J. Lu, and J. Xia, "A machine learning based method for real-time queue length estimation using license plate recognition and gps trajectory data," *KSCE Journal of Civil Engineering*, vol. 26, no. 5, pp. 2408–2419, 2022.

[11] X. Zhan, R. Li, and S. V. Ukkusuri, "Lane-based real-time queue length estimation using license plate recognition data," *Transportation Research Part C: Emerging Technologies*, vol. 57, pp. 85–102, 2015.

[12] P. Pudasaini, A. Karimpour, and Y.-J. Wu, "Real-time queue length estimation for signalized intersections using single-channel advance detector data," *Transportation research record*, vol. 2677, no. 7, pp. 144–156, 2023.

[13] W. Al Okaishi, A. Zaarane, I. Slimani, I. Atouf, and M. Benrabh, "A vehicular queue length measurement system in real-time based on ssd network," *Transport and Telecommunication*, vol. 22, no. 1, pp. 29–38, 2021.

[14] Ultralytics, "Yolov5." https://github.com/ultralytics/yolov5?tab=readme-ov-file. (Accessed: 2023-02-10).

[15] A. Battal, Y. Avci, and A. Tuncer, "Vehicle detection and counting in traffic videos using deep learning," *ICENTE'23*, p. 272, 2023.

[16] M. Boneh, "Vehicle detection." https://github.com/MaryamBoneh/VehicleDetection/tree/main/Dataset. (Accessed: 2023-02-12).

[17] L. Soetanto, "Vehicle detection dataset." https://www.kaggle.com/datasets/lyensoetanto/vehicle-images-dataset. (Accessed: 2023-02-21).

[18] M. TÁRNOK, "5 vehichles for classification." https://www.kaggle.com/datasets/mrtontrnok/5-vehichles-for-multicategory-classification. (Accessed: 2024-06-06).

[19] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, "Simple online and realtime tracking," in *2016 IEEE international conference on image processing (ICIP)*, pp. 3464–3468, Ieee, 2016.

[20] P. K. Thadagoppula and V. Upadhyaya, "Speed detection using image processing," in *2016 International Conference on Computer, Control, Informatics and its Applications (IC3INA)*, pp. 11–16, IEEE, 2016.

[21] C.-J. Lin, S.-Y. Jeng, and H.-W. Lioa, "A real-time vehicle counting, speed estimation, and classification system based on virtual detection zone and yolo," *Mathematical Problems in Engineering*, vol. 2021, no. 1, p. 1577614, 2021.

[22] A. M. Santos, C. J. Bastos-Filho, A. M. Maciel, and E. Lima, "Counting vehicle with high-precision in brazilian roads using yolov3 and deep sort," in *2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pp. 69–76, IEEE, 2020.

[23] Y. Zhao, X. Zhou, X. Xu, Z. Jiang, F. Cheng, J. Tang, and Y. Shen, "A novel vehicle tracking id switches algorithm for driving recording sensors," *Sensors*, vol. 20, no. 13, p. 3638, 2020.

El Cezeri
Journal of Science and Engineering
ISSN: 2148-3736

Research Article

Effect of Star Structure Versus Linear Polymers on Structure and Shear Rheology

Nese Cakir Yigit[1a]

[1]Department of Polymer Materials Engineering, Yalova University, Yalova, 77100, Türkiye
nese.cakir@yalova.edu.tr

DOI : 10.31202/ecjse.1735047
Received: 05.07.2025 Accepted: 30.09.2025
How to cite this article:
Nese Cakir Yigit, "Effect of Star Structure Versus Linear Polymers on Structure and Shear Rheology", El-Cezeri Journal of Science and Engineering, Vol: 12, Iss: 3, (2025), pp.(365-xx).
ORCID: [a]0000-0002-4714-4488

Abstract Star-shaped polymers have a compact molecular architecture and reduced entanglement density compared to linear analogues. They offer opportunities to tailor solutions and melt properties for various applications. This study examines the influence of polymer architecture on rheological behavior by comparing star-shaped polystyrene (PS) with linear PS. Star polymer was synthesized via atom transfer radical polymerization (ATRP) using a poly(divinylbenzene) (PDVB) core and PS arms. Both star and linear polymers with narrow molecular weight distributions were evaluated alongside a commercially available high-molecular-weight PS. Rheological measurements were performed at 25 °C in the presence of dioctyl adipate (DOA) plasticizer at varying shear rates. Under flow, polymer chains are dispersed, increasing free space and reducing molecular interactions. At higher shear rates, the polymers achieved the infinite shear viscosity plateau ($\eta\infty$). The commercially used linear PS mixture with DOA showed shear rate-independent behavior. Meanwhile, the star polymer had a smaller hydrodynamic volume and gyration radius, resulting in lower viscosity.

Keywords: Linear polymer, Multiarm star polymer, Polymer characterization, Rheology.

I. INTRODUCTION

Star polymers, characterized by multiple arms extending from a central core, represent a class of branched macromolecules with a compact structure, globular shape, and extensive surface area. Incorporating diverse functional groups into star polymers significantly broadens their potential applications, including polymer therapeutics, diagnosis, optical imaging, membranes, coatings, viscosity modifiers, catalysis, separation media, thin films, and numerous other advanced materials [1-3].

Atom transfer radical polymerization (ATRP) is among the most widely used living polymerization techniques for synthesizing star polymers. Two primary synthetic strategies are employed: "core-first" and "arm-first" approaches [4-6]. In the "arm-first" procedure, the addition of the cross-linker at a specific conversion of the monomers occurs in a "one-pot" process. A "two-pot" process conducts the synthesis of star polymers using the purified linear polymers, known as macroinitiators (MIs). These techniques effectively synthesize numerous monomers into star-like polymers, such as methyl methacrylate, styrene, and tert-butyl acrylate. Despite polymerization using "core-first" and "arm-first" methods, all the arms of these star polymers display linear characteristics [7-10].

Linear polymers generally exhibit poor mechanical properties and limited thermal stability; therefore, cross-linking reactions, such as those applied in the construction of star polymers, are commonly employed to generate polymer networks with enhanced elasticity, thermal resistance, and porosity. Additionally, star polymers possess compact structures with reduced entanglement, abundant functional groups, and superior processability. These characteristics define their unique physicochemical properties, including lower glass transition temperatures, reduced bulk and solution viscosities, increased solubility, and enhanced ease of processing and modifications. Star polymers, in particular, have attracted growing interest in gene delivery and nanomedicine, as their unique architecture enables the encapsulation of genetic material within their cores, thereby allowing efficient regulation of gene release into the cytosol. Consequently, star polymers have garnered considerable attention in recent years and are increasingly recognized as promising materials for a wide spectrum of technological and biomedical applications [3, 11-13].

A rheological study ranges from viscoelastic liquids to viscoelastic solids for analyzing Newtonian or non-Newtonian fluids, examining shear rate and shear stress, thereby providing valuable insights into their intrinsic fluid properties [14, 15]. The characteristics of the rheological properties of star-shaped systems are influenced by the dynamics of two components: the linear arms and the core, which move in conjunction with the arms. As a result, star polymers with a low number of arms f have a Gaussian coil conformation; stress relaxation is independent. As the f in the star polymer increases, the compression of the structure leads to a further decrease in the viscosity of the samples. High f values and significant core size lead to a core-shell structure, requiring rheological behavior analysis concerning the compact centers of star macromolecules [16-18]. Moreover, computer simulations demonstrated that the behavior of multiarm star-shaped polymers in suitable solvents resembles that of solid particles and markedly differs from that of classical coil-like macromolecules [19].

Liu et al. investigated the effect of short arms on the rheological behavior of star polymers. Their study showed that as the short arm length increases, the linear viscoelastic rubber plateau decreases, which causes the relaxed short arms to act as a "solvent" and dilute the polymers. They found their work encouraging for guidance on molecular design to tune linear and nonlinear rheology in star polymers [20]. Van Ruymbeke et al. investigated a quantitative assessment of star rheology using a simple model accounting for high-frequency response and low-frequency arm relaxation. They observed deviations and confirmed the universal mechanism of arm relaxation governing star rheology. However, they mentioned that subtle issues must be addressed for accurate quantitative dynamics [21].

Although star-shaped polymers have been extensively studied in academic research, their industrial application remains limited. Key challenges include reducing preparation costs and simplifying synthetic routes and purification processes, particularly for removing linear or star-star coupled products, toxic metal catalysts, and odorous compounds. Addressing these issues is essential for broader industrial adoption of star polymers [22]. Mandal et al. synthesized narrowly dispersed and degradable homoarm and miktoarm star polymers in one pot using the catalytic living ring-opening metathesis polymerization (ROMP) mechanism with vinyl ether chain transfer agents. They argued that this sustainable and environmentally friendly method offers a valuable synthetic tool for polymer engineers in supramolecular, industrial, or biomedical research [23]. Sparnacci et al. investigated a synthetic procedure for core-crosslinked star copolymers suitable for direct industrial implementation. Using the arm-first technique, they synthesized core-cross-linked star (CCS) copolymers via reversible addition-fragmentation chain transfer polymerization (RAFT). They evaluated their performance as viscosity index improvers in hydraulic lubricants. For comparison, a commercial linear PMA-type transmission oil was also tested. The study demonstrates that by adjusting the properties of CCS polymers, these materials can serve as model lubricant formulations for hydraulic applications and fluid production [24].

Moreover, star polymers have been successfully applied in the coating industry as rheological regulators, enhancing material processing and liquidity. Studies on drilling fluid systems based on star polymers have appeared in some papers in recent years. Luo et al. published an acrylamide-based star-shaped copolymer for water-based drilling fluids with better rheological properties and more effective filtration performance at high temperatures that can address the limitations of linear polymeric additives [25]. Jian et al. recently developed a new generation of amphoteric star polymers, which exhibited high molecular weight, branching density, and surface area for drilling fluid systems. The polymer formed aggregates, generated stable emulsions in the study, and maintained plastic viscosity under high-temperature conditions. The star polymer produced thin filter cakes and minimal fluid loss volume in oil-based mud systems, demonstrating excellent fluid control and rheology profiles [26].

Herein, in two steps, we used an “arm-first” strategy to produce a multi-arm star PS of absolute molecular weight using atom transfer radical polymerization (ATRP). We also investigated the nonlinear rheological behavior of multiarm star PS chains and compared the results with the behavior of linear PS polymers with approximately the same molecular weights. Furthermore, all synthesized star and linear polymers were compared with the commercially available PS polymer for rheological properties to see the effect of precise molecular weight control and narrow molecular weight distribution on the architecture.

II. EXPERIMENTAL SECTION

2.1 Materials

Styrene (St, 99% Aldrich) and divinyl benzene (DVB, 80%, Aldrich) were purified by passing through a column filled with basic alumina to remove the inhibitor immediately before use. N, N, N′, N″, N″-Pentamethyldiethylenetriamine (PMDETA, Aldrich) was distilled over NaOH before use. Ethyl α-bromoisobutyrate (EiBr, 98%), CuBr (99.9%), and all solvents were purchased from Aldrich and used as received. PS average M_w ~192,000 was provided by Aldrich and used as received. Polyethylene glycol 400 (PEG) was purchased from Ataman Chemicals (Türkiye). The plasticizers dioctyl terephthalate (DOTP), epoxidized soybean (ESBO), and dioctyl adipate (DOA), all with 99% purity, were obtained from Plastay Chemistry (Türkiye).

2.2 Characterization

Polymers were characterized by gel permeation chromatography (GPC). The GPC measurements were carried out with an Omnisec multi-detector GPC instrument from Malvern Panalytical Ltd. (Worcestershire, U.K.) equipped with four-column PL gel 5 mm Mixed-C and multi-detector including refractive index, UV/visible PDA, viscosimetric, and light scattering detectors to measure the molecular weight distribution and absolute molecular weights of polymers in tetrahydrofurane (THF) as a mobile phase calibrated with a narrow molecular weight distributed PS standard. The flow rate of THF was 0.9 mL/min at 35 ºC. The concentrations of sample solutions were prepared as 1 to 5 mg/mL, and the injection volume was 0.1 mL. A differential refractive index (dn/dc) of 0.185 mL/g was measured for PS.

The Agilent 6890N gas chromatograph (GC), equipped with a 7683 automatic liquid injector, an FID detector, and a 3-meter 180-μm DB-1 capillary column, followed the star formation. The injector and detector temperatures were maintained at 280

°C and 285 °C, respectively. The initial column temperature is 40 ºC, ultimately increasing to 120 ºC at a heating rate of 40 ºC/min.

Fourier transform infrared (FTIR) spectroscopy measurements (FTIR) were obtained using an Agilent Technologies Cary 630 FTIR instrument within the 4000–500 cm^{-1} range.

The hydrodynamic diameters of the polymers were determined using dynamic light scattering (DLS) using a Malvern NanoZSP instrument with a He-Ne laser operating with a 632 nm wavelength at an angle of 173 ºC. Samples were prepared at a 1 mg/mL concentration in THF dispersions.

Differential scanning calorimetry (DSC) was conducted using a TA Instruments Q2000 Series Thermal Analysis System equipped with an autosampler within a temperature range of −80 to 400 °C at a 10 °C min-1 heating rate under a nitrogen atmosphere and using approximately 5 mg of sample.

Rheological measurements were conducted using the TA Instruments Discovery Hybrid Rheometer HR-2, equipped with a 220 mm holder and a parallel steel plate. All measurements were recorded at 25 ± 0.01 ºC.

2.3 Synthesis of linear polymers by ATRP

Linear polystyrene (PS) polymers (P1 and P2) were prepared by ATRP of St using PMDETA ligand and a copper(I) catalyst. Two different monomer: initiator ratios were used: [M]/[I] = 200 for P1 and [M]/[I] = 2000 for P2, with EiBr as initiator. A typical experimental procedure is as follows: St (P1 = 20.0 mL, 175.0 mmol; P2 = 50.0 mL, 436.0 mmol), PMDETA (P1 = 0.183 mL, 0.875 mmol; P2 = 0.045 mL, 0.218 mmol), CuBr (P1 = 0.125 g, 0.875 mmol; P2 = 0.031 g, 0.218 mmol), and EiBr (P1 = 0.128 mL, 0.875 mmol; P2 = 0.032 mL, 0.218 mmol) were added in sequence to a Schlenk flask with a magnetic stirring bar. The flask was sealed with septa. Then the reaction mixture was degassed by three freeze–pump–thaw cycles. After degassing, the Schlenk flask was left in a vacuum and placed in a thermostated oil bath at 110 ºC. Following a 45-minute polymerization of P1 and a 24-hour polymerization of P2, the flask was cooled to room temperature and exposed to oxygen to terminate the polymerization. Subsequently, the polymerization mixture was diluted with THF, passed through a basic alumina column to remove the catalyst, and precipitated in excess methanol. The polymer was dried for 24 h in a vacuum oven at 40 °C (P1 = 5.6 g, 31% yield, P2 = 15.5 g, 34% yield).

2.4 Synthesis of multiarm star polymer by "arm-first" technique

P1 (2.0 g, 0.267 mmol) was dissolved in 11 mL of anisole in a Schlenk flask equipped with a magnetic stir bar, achieving a total macromonomer concentration of 0.023 M in the reaction. DVB (0.570 mL, 4.01 mmol), PMDETA (0.056 mL, 0.267 mmol), and CuBr (0.038 g, 0.267 mmol) were added to the solution. After dissolution, an aliquot (~ 10 mL) was quickly removed from the reaction mixture for GC analysis. The mixture in the Schlenk flask was degassed by three freeze–vacuum–thaw cycles. Then, filled with argon gas, sealed, and immersed in a thermostated oil bath at 110 ºC. Samples were taken from the reaction mixture periodically to screen DVB conversion. After 24 h at 95 % conversion, the reaction was stopped, diluted in THF, and passed through an alumina column. The entire reaction mixture was dried on a rotary evaporator, redissolved in CHCl3, and precipitated in a methanol/diethyl ether (1/3; v/v) mixture, recovered by suction filtration, and dried in a vacuum oven at 40 ºC for overnight.

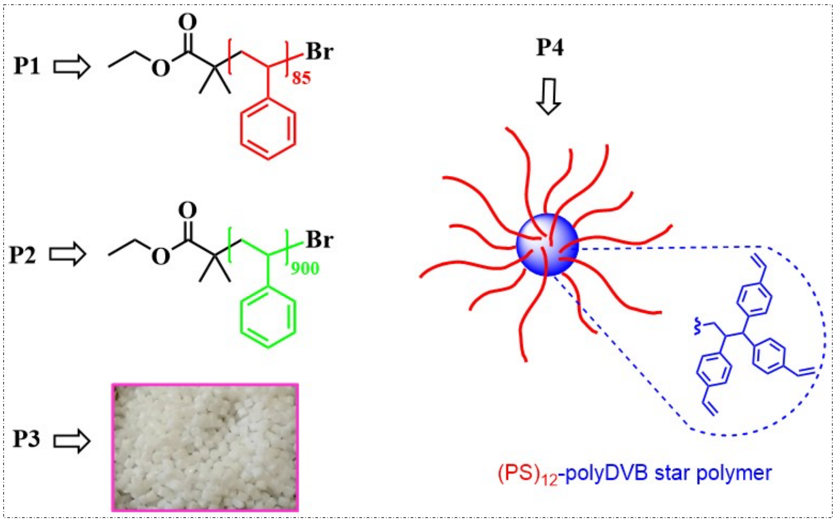
III. RESULTS AND DISCUSSION

The present study involves the rheological comparison of the synthesized low and high-molecular-weight linear PS polymers, a multiarm star PS polymer consisting of PDVB core and short PS homopolymer arms, and commercially available linear PS with high molecular weight distribution. Scheme 1 illustrates all the P1-P4 polymers investigated in this study.

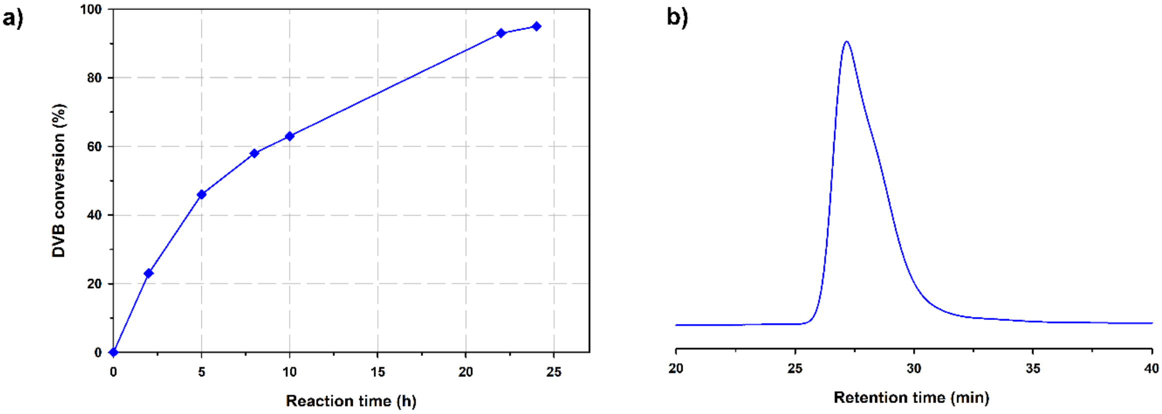
First of all, we employed the ATRP reaction of styrene monomer to obtain linear low-dispersity PS samples using EiBr as an initiator and CuBr/ PMDETA as a catalyst in bulk polymerization at 110 ºC, as described earlier [27]. Linear polymers possessing 2-bromoisobutyryl end-groups were synthesized with narrow dispersity values using a feed ratio of St: EiBr/200:1 and 2000:1 for P1 and P2, respectively. The initiating end group: catalyst: PMDETA ratio was 1:1:1 in all cases. The polymers were purified by passing their solution in THF through a column filled with neutral aluminum oxide, followed by evaporation of the solvent and precipitation into methanol. Results of molecular weight values (M_n and M_w) and dispersity (PDI) are depicted in Table 1.

In addition to these synthesized polymers, high molecular weight PS (P3) was purchased commercially from Aldrich to see whether the rheological properties of linear or star polymers with narrow molecular weights were superior to broad distribution polymers. The molecular weight values, polydispersities, and intrinsic viscosities of all polymers for which rheology measurements were made are shown in Table 1.

Multiarm star polymer (P4) was prepared by ATRP, and the “arm-first” method using P1 as a macroinitiator. The P1 was crosslinked with DVB to form a multiarm $(PS)_n$-polyDVB star polymer (P4). The star formation was monitored by the GC instrument. The samples were diluted with $CHCl_3$ to an appropriate concentration for the GC. The time-dependent conversion of DVB was calculated from the ratio of the peak area of DVB to the peak area of the internal standard anisole. Until 50% DVB conversion, the plot in Figure 1a was linear. The curvature at higher DVB conversion may have been caused by certain alkyl bromide initiating sites within the highly cross-linked polyDVB core becoming less accessible and less likely to be activated and react with DVB. After 24 hours, the reaction was stopped at 95% DVB conversion, and the mixture was purified by fractional precipitation from THF into a methanol/diethyl ether mixture to eliminate unconverted linear PS. Figure 1b shows the relevant GPC chromatogram of the purified multiarm star polymer with monomodal distribution. As envisaged, PS homopolymer was used as a macroinitiator for synthesizing a multiarm star polymer, and subsequent purification steps were effectively carried out.



(1)



(2)

Figure 1. (1) Structures of star and linear polymers used in this study. Low molecular weight linear PS (P1) is highlighted in red, high molecular weight linear PS (P2) is highlighted in green, and commercially obtained high molecular weight linear PS (P3) is framed in pink. The DVB core of multiarm star polymer (P4) is highlighted in blue with linear red arms (P1). (2)a) Kinetic diagram of the DVB conversion, obtained via GC, depending on the star polymer synthesis reaction time. b) Corresponding RI-trace of purified multiarm $(PS)_n$-polyDVB star polymer (P4), obtained via GPC in THF at 35 °C.

All of the obtained polymers, P1-P4, exhibited monomodal GPC traces, as shown in Figure 2a. The commercially available P3 polymer had a relatively broader molecular mass distribution than the ATRP systems used in synthesizing other polymers (P1, P2, and P4), indicating that an uncontrolled polymerization was used.

Figure 2b shows the FTIR spectra of all polymers, P1-P4. The C–H aromatic stretching vibrations were observed at 3060 and 3025 cm^{-1}, while the C–H stretching vibration occurred at 2925 cm^{-1}. Additionally, the phenyl ring stretching vibrations were noted at 1600, 1496, 1460, 750, and 700 cm^{-1}, indicating the presence of typical absorption bands for PS across all polymers. As all the polymers (P1-P4) bear structural similarities, they showed similar spectral properties. Moreover, due to the appearance of the intense characteristic bands of PS arms, star polymer (P4) showed similar bands to its homopolymer (P1) signals.

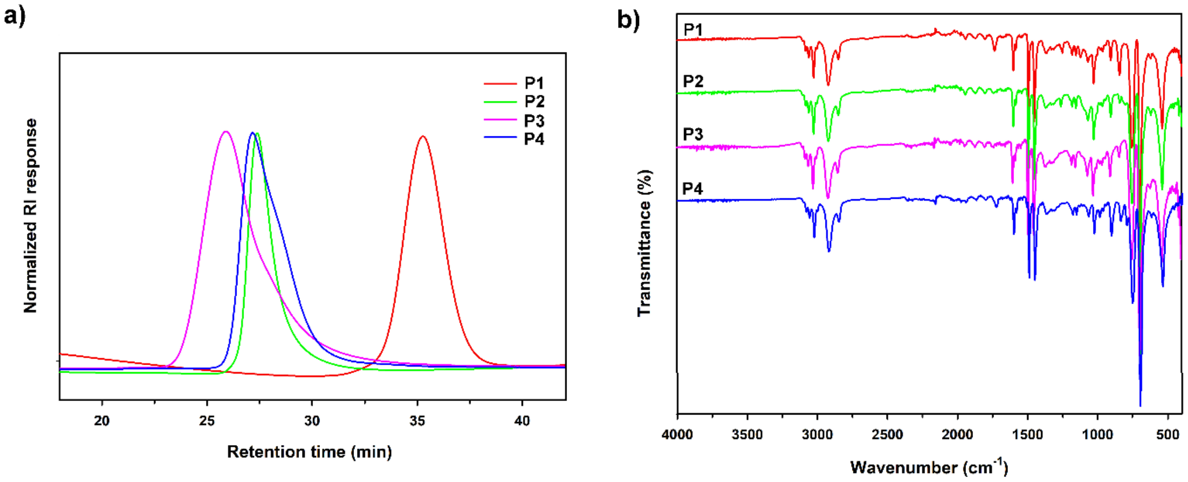


Figure 2. a) GPC chromatograms of all P1-P4, in THF at 35 ºC. b) FTIR spectrum of all polymers, P1-P4.

Table 1 shows the actual molecular weight (M_w = 129000) of $(PS)_n$-polyDVB multiarm star polymer obtained by multi-detector GPC, with the introduction of the d*n*/d*c* = 0.185 mL/g of linear PS standard. Since the star polymer was obtained by the "arm first" technique, the average number of arms *f* per star macromolecule was calculated by using the molecular weight of the entire star together with the weight of the macroinitiator (MI) under the assumption that the crosslinking agent used was incorporated into the polymer core, with the following equation (1):

$$f = \frac{M_{w,star}}{M_{w,arm} + [DVB]/[PS]\; x\; M_{DVB}\; x\; conv_{DVB}} \tag{1}$$

Here, $M_{w,star}$, $M_{w,arm}$ are the weight-average molecular weights of the star polymer and macroinitiator, respectively, and [DVB]/[PS], M_{DVB}, $conv_{DVB}$ are the feed molar ratio of the crosslinker to macroinitiator, the molecular weight, and conversion of crosslinker. As given in Table 1, *f* was calculated to be 12 using equation (1).

The intrinsic viscosities ([η]) of all polymers, P1-P4, were determined by multi-detector GPC in THF at 35 ºC, as shown in Table 1. For polymers P1, P2, and P3, [η] increased with molecular weight. A low viscosity value was measured for the star polymer, P4. As expected, star polymers have a compact structure, which leads to a decrease in the viscosity of the samples. Moreover, with the increase in the number of arms, a further compaction of the structure and a further reduction in the viscosity of the samples are observed [28]. It is generally recognized that the most expedient technique for elucidating the structure of star polymers is to compare the intrinsic viscosity of the star polymer $[\eta]_{star}$ and its linear analog with the same molecular weight, $[\eta]_{linear}$, where g' is the contraction factor, as given in equation (2) [29].

$$g' = [\eta]_{star} / [\eta]_{linear} \quad (M = \text{constant}) \tag{2}$$

The intrinsic viscosity for linear PS, $[\eta]_{linear}$ with comparative molecular weight was obtained from the Mark–Houwink–Sakurada (MHS) equation, as given in equation 3. The MHS values of $K = 1.44 \times 10^{-4}$ dL/g and $\alpha = 0.707$ were obtained from the literature for linear narrow PS in THF at 35 ºC [8]. Then, using these parameters $[\eta]_{linear}$ was calculated to be 0.59 dL/g for a specified molecular weight (M_w = 129000 g/mol) of linear PS. Moreover, the $[\eta]_{star}$ of star polymer was measured to be 0.174 dL/g by the viscometer detector in multi-detector GPC.

$$[\eta] = K \; x \; M^{\alpha} \tag{3}$$

It is also demonstrated that the number of arms, *f*, is inversely proportional to g' in regular star polymers with equal arm lengths, as shown in equation 4. Then, *f* was calculated as 13, which closely agrees with the result obtained from equation 1.

$$\log g' = 0.36 - 0.8 \log f \tag{4}$$

Table 1 also shows the hydrodynamic radius (R_h) of all polymers, P1-P4. As expected, the increased molecular weights increased the R_h values of polymers.

Table 1. Multi-detector GPC data of polymers P1-P4

Polymer	Architecture	f^{e}	M_n (g/mol)	M_w (g/mol)	*PDI*	[η] (dL/g)	R_h (nm)
P1	Linear [a]	-	7500	8800	1.17	0.082	2.19
P2	Linear [b]	-	80000	94000	1.17	0.46	6.72
P3	Linear [c]	-	104100	198000	1.91	0.73	9.23
P4	Star [d]	12(13) [f]	98500	129000	1.31	0.174	7.51

[a] $[M]_0$:$[I]_0$:[CuBr]:[PMDETA] = 200:1:1:1; ATRP reaction at 110 ºC for 45 min.
[b] $[M]_0$:$[I]_0$:[CuBr]:[PMDETA] = 2000:1:1:1; ATRP reaction at 110 ºC for 24 h.
[c] Obtained commercially from Aldrich.
[d] $[DVB]_0/15 = [P1]_0 = [CuBr]_0 = [PMDETA]_0$= 0.023 M in anisole at 110 °C for 24.
[e] Calculated according to equation (1).
[f] Calculated according to equations (2), (3), and (4).

Particle size measurements were established for P1-P4 in THF suspension by DLS. Figure 3 revealed that DLS data weighted by numbers (Figure 3a) and volume (Figure 3b) of P1-P4 obtained. The number-weighted DLS measurement of the samples and the volume-weighted exhibited one peak with unimodal particle size distributions for all samples analyzed. The hydrodynamic diameter (D_h) value indicated an increase from 3.73 nm to 10.82 nm (number-average) after crosslinking the linear polymer (P1) to form the star polymer (P4). Comparing the D_h values of linear PS polymers, P1-P3, there was a linear relationship between the intensity of the light scattered and the increase in molecular weight, resulting in a larger size.

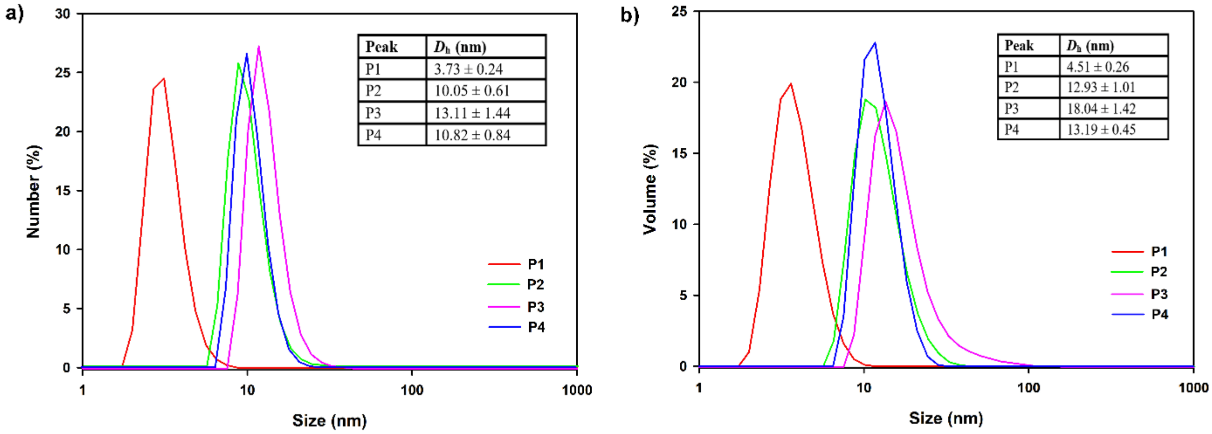


Figure 3. Dynamic light scattering (DLS) measurements of the size dispersion of P1-P4 samples represent data weighted by number (a) and volume (b) of particles.

The thermal behaviour of polymers, P1-P4, was studied by DSC. As shown in Figure 4a, DSC measurements showed the glass transition temperatures (T_g) of P1-P4 to be 95.6 °C, 108.1 °C, 103.8 °C, and 105 °C, respectively. Generally, the T_g of polymers was shifted to higher temperatures (from 95 to 105 ºC) with increased molecular weights. As expected, no crystalline fraction was observed in all the polymers. For the multiarm star polymer (P4), broad glass transitions were observed by DSC. Interestingly, the T_g of the PS arm cannot be detected. This can be due to its partial miscibility with the polyDVB core, which may lead to a plasticizing effect and thus significant broadening of the T_g of the star polymer. In addition, the unique structure of star-shaped copolymers hinders the movement of segments.

As shown in Figure 4b, the crystallization peak of the PS segments was evident for all polymers. Moreover, the crystallization peak of PS in the multiarm star polymer (P4) was more prominent. This can be because the PS chains are directly connected to the star core molecule and give better mobility to the PS chains, which is advantageous for crystallization.

For rheological measurements, the miscibility of the four most commonly used plasticizers with PS was tested by selecting the P3 polymer as a reference. P3 polymer solutions containing different plasticizers were prepared at a concentration of 20% w/w. Concentration was calculated as c(% w/w) = (m_{PS}/(m_{PS}+ $m_{plasticizer}$)) x 100. Chloroform ($CHCl_3$) was chosen as a cosolvent; P3 was dissolved in $CHCl_3$, and plasticizers were added to the polymer solutions individually. After that, $CHCl_3$ was evaporated during stirring under ambient conditions for up to 2 days until the sample weight stabilized. As depicted in Figure 5, phase separation occurred in PS polymer blends containing PEG, DOTP, or ESBO after solvent evaporation.

However, the PS mixture containing DOA was transparent and viscous, so we used the DOA-containing P1, P2, P3, and P4 blends.

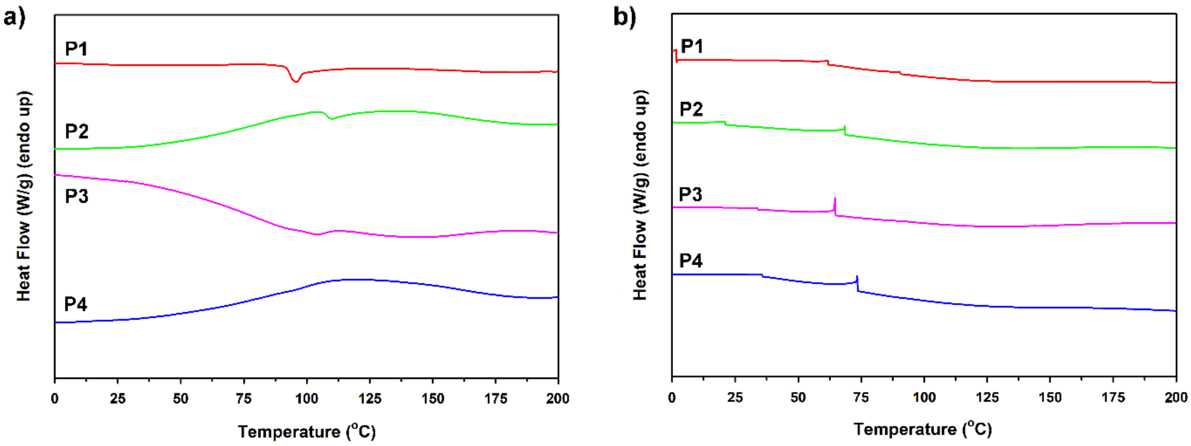


Figure 4. DSC curves for polymers P1- P4 with scanning rates of 10 °C min^{-1} (a) first heating run and (b) first cooling run.

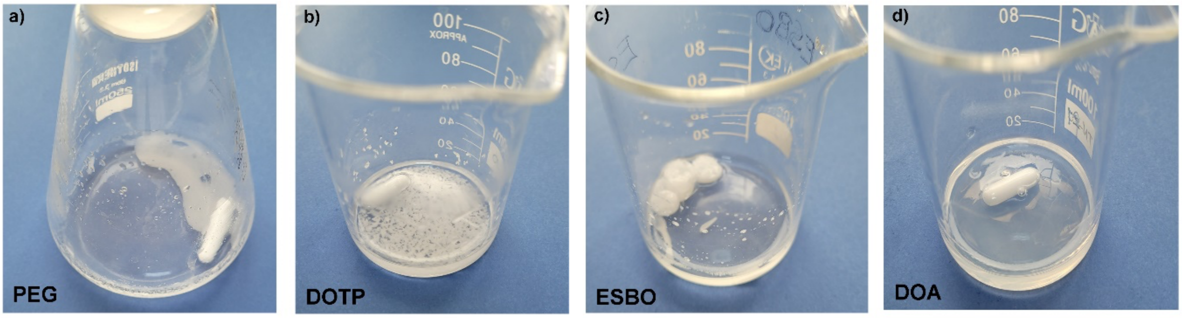


Figure 5. Images of blends containing P3 with (a) PEG, (b) DOTP, (c)ESBO, and (d) DOA

A parallel steel plate performed rheological measurements. Since the samples are PS blends containing plasticizers, parallel plate measurements are more suitable for this high-viscosity sample, as they provide good contact between the rheometer plate surfaces and the polymeric liquid. Slip is also prevalent when employing rheometers to test highly concentrated polymer solutions. The measured viscosity can be substantially lower than the actual viscosity of the sample when leakage occurs. To address these limitations, slip tests were conducted on samples with varying gap heights and geometries. It was determined that slip was not occurring within the measured range of shear and strain.

A shear rate sweep test ranging from 10^{-2} to 10^{3} s^{-1} was performed to examine the rheological properties of the polymer solutions. Figure 6 illustrates the flow curves for all DOA-containing PS samples at a concentration of 20% w/w (P1-P4 blends). Polymers in solution showed a shear rate-dependent viscosity behavior below a critical shear rate γ_{crit}, termed the non-Newtonian power-law region, where the viscosity decreases as the shear rate increases. One possible explanation for the relation between shear rate and viscosity is that molecular alignment under shear facilitates the flow of polymer chains, reducing shear viscosity as the shear rate increases. The other observation was the relation between molecular weight and viscosity. The dependence of viscosity on shear rate is initiated at lower γ levels with an increase in molecular weight. Above this γ_{crit}, the solution viscosity remained constant regardless of shear rate, referred to as the zero-shear viscosity, η_0. The transition from shear-rate-dependent to shear-rate-independent viscosity occurred within a comparatively narrow region due to the P1, P2, and P4 samples with a narrow molecular weight distribution. Additionally, Figure 6a indicates that assessing molecular weight degradation caused by shear forces presents challenges in the non-Newtonian region, as samples with varying molar masses exhibit constant viscosity at a specific shear rate. DOA-containing P3 blends showed totally shear-rate independent behavior because they are commercially available with high PDI.

Figure 6b presents the shear stress versus the shear rate flow curves of P1-P4 blends. All samples exhibited shear-thinning behaviour, particularly in lower molecular weight PS blends.

The viscosity of the samples decreased with increasing shear rate and exhibited an infinite shear viscosity (η_∞) at high shear rates. The flow behavior of all samples demonstrated shear-thinning Non-Newtonian characteristics, which intensified with the increasing molecular weight of PS in the DOA-containing PS samples, ultimately transitioning to a more Newtonian behavior.

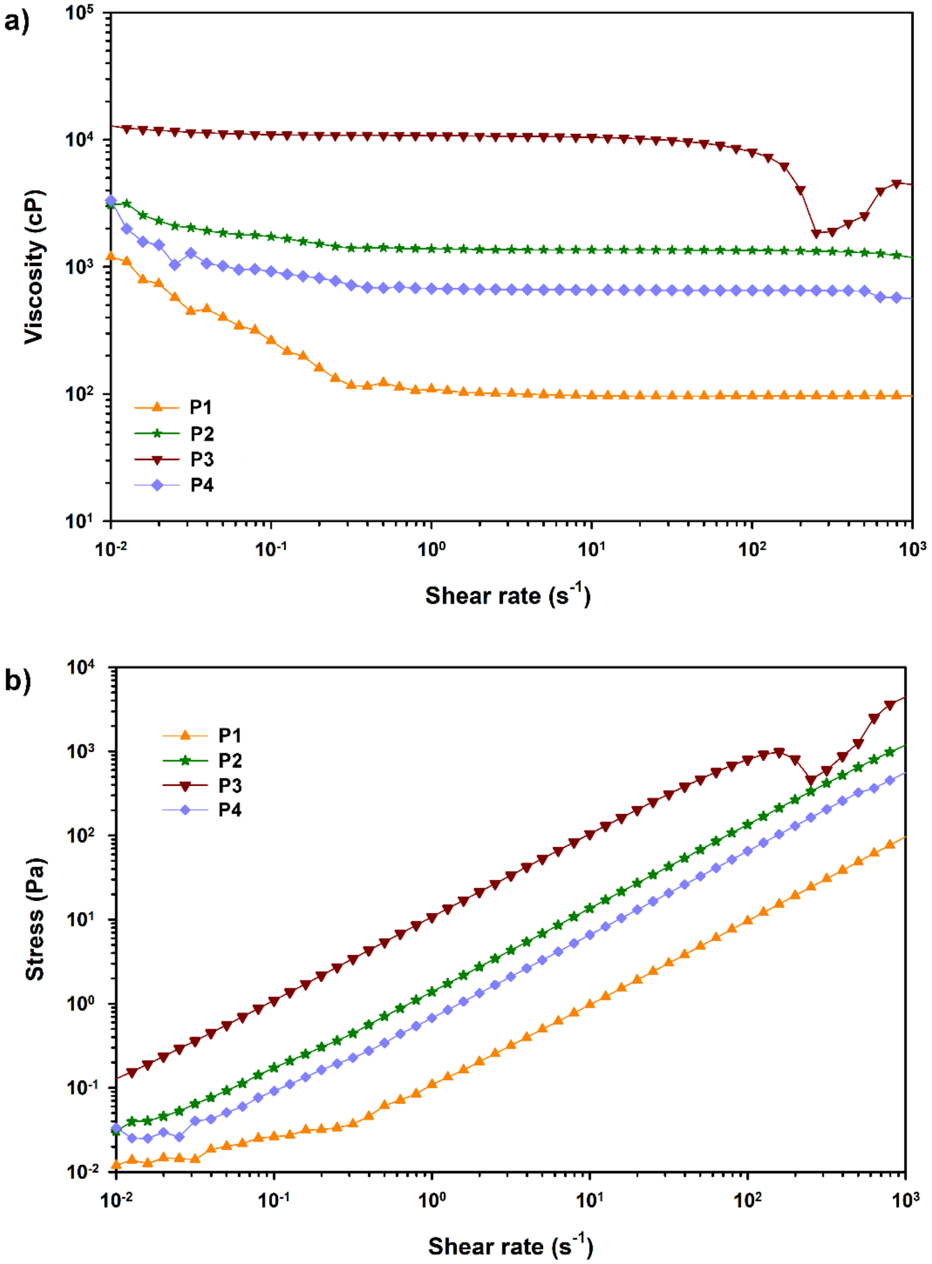


Figure 6. a) Dependency of the shear viscosity on the shear rate for P1-P4 samples, b) Shear stress versus shear rate for P1-P4 samples at 25 °C.

Table 2. Zero-rate and infinite-rate viscosities and coefficient values fitting with Carreau and Cross Models

Samples	*η_0 (mPa.s)* Zero-Rate Viscosity	*η_∞ (mPa.s)* Infinite-Rate Viscosity	*R^2* Carreau Model	*R^2* Cross Model
DOA only	499.511	13.7095	0.9832	0.9832
P1:DOA	1240.35	97.0102	0.9931	0.9939
P2:DOA	3070.90	1334.02	0.9830	0.9830
P3:DOA	10283.4	2686.23	0.9021	0.9001
P4:DOA	3258.37	570.746	0.8599	0.9253

Furthermore, an increase in the molecular weight of PS resulted in a slight increase in the parameter η_∞, indicating a rise in viscosity, as demonstrated in Table 2. The Carreau and Cross model was employed regarding the overall error in estimates of the non-Newtonian viscosity-shear rate relationship, and the results indicated a comparable shear-thinning flow behavior.

IV. CONCLUSIONS

This study reports how variations in polymer architecture affect rheological properties by systematically comparing star-shaped polystyrene (PS) to its linear counterparts of similar molecular weight and to high-molecular-weight, broad-distribution linear PS. Research begins with synthesizing a star-shaped polymer with a poly(divinylbenzene) (PDVB) core and PS arms using ATRP polymerization. This star-branched PS displays relatively narrow molecular weight distributions due to well-ordered arms. Linear PS samples with approximately the same molecular weight are prepared by conventional ATRP to compare the effects of architecture on rheology directly. Additionally, high-molecular-weight PS with a broad molecular weight distribution was purchased from Aldrich to evaluate whether linear or star polymers with narrow molecular weight distributions exhibit superior rheological properties compared to those with broad distributions.

Rheology tests were performed using a 20% w/w polymer solution, including dioctyl adipate (DOA) plasticizer. The viscosity and shear stress of the sample solutions were measured at various shear rates at 25 °C. The entangled and randomly oriented polymer chains become dispersed during flow, reducing molecular interactions and creating more free space, decreasing viscosity in polymer solutions. The polymers were completely disentangled at elevated shear rates, resulting in the system's viscosity attaining the infinite shear viscosity plateau (η_∞). The commercially used linear PS mixture containing DOA exhibited completely shear rate-independent behavior because it contained polymer chains of different lengths with a broad molecular weight distribution. The star polymer exhibited a smaller hydrodynamic volume and radius of gyration, and thus lower viscosity, compared to its linear analogs of similar molar mass.

These findings demonstrate that polymer topology significantly influences rheological behavior, indicating that star-shaped architectures offer advantages for applications requiring lower viscosity and improved processability. Future research should investigate a broader range of functional star polymers, assess their performance under conditions relevant to industry, and clarify structure–property relationships to inform the design of materials with specific rheological characteristics.

Acknowledgments
We thank Prof. Dr. Aydan Dag for generously allocating the Schlenk Line used in this work.

Authors' Contributions

Author wrote the manuscript.

Competing Interests

The authors declare that they have no conflict of interest.

References

[1] J. M.Ren, T. G.McKenzie, Q.Fu, , E. H.Wong, J. Xu, Z.An, S. Shanmugam, T. P. Davis, C. Boyer, and G. G. Qiao, Star polymers, *Chemical Reviews*, *116* (12), 6743-6836, (2016).
[2] P. Tikhonov, N. Vasilenko, and A.Muzafarov, Multiarm star polymers. Fundamental aspects. A review, Doklady Chemistry: *Springer*, Vol. 496, pp 1-17,(2021).
[3] W.Wu, W. Wang, and J. Li, Star polymers: Advances in biomedical applications, *Progress in Polymer Science*, *46*, 55-85, (2015).
[4] J.Du, and Y. Chen, Preparation of poly (ethylene oxide) star polymers and poly (ethylene oxide)–polystyrene heteroarm star polymers by atom transfer radical polymerization, *Journal of Polymer Science Part A: Polymer Chemistry*, *42* (9), 2263-2271, (2004).
[5] A. Herberg, X Yu,. and D.Kuckling, End group stability of atom transfer radical polymerization (ATRP)-synthesized poly (N-isopropylacrylamide): Perspectives for diblock copolymer synthesis, *Polymers*, *11* (4), 678, (2019).
[6] J. Xia, X.Zhang, and K.Matyjaszewski, Synthesis of star-shaped polystyrene by atom transfer radical polymerization using an "arm first" approach, *Macromolecules*, *32* (13), 4482-4484, (1999).
[7] M. Spiniello, A. Blencowe, and G. G.Qiao, Synthesis and characterization of fluorescently labeled core cross-linked star polymers, *Journal of Polymer Science Part A: Polymer Chemistry*, *46* (7), 2422-2432, (2008).
[8] N. Cakir Yigit, G.Hizal, and U.Tunca, A powerful tool for preparing peripherally post-functionalized multiarm star block copolymer, *Polymer Bulletin*, *75*, 3523-3538, (2018).
[9] D. S. Esen, Cakir N.Yigit, U. Tunca, G. Hizal, and N.Arsu, Synthesis and characterization of multiarm (Benzoin-PS) m-polyDVB star polymer as a polymeric photoinitiator for polymerization of acrylates and methacrylates, *Journal of Polymer Science*, *59* (18), 2082-2093, (2021).
[10] Y. Dai, H. Sun, S. Pal, Y. Zhang, S.Park, C. P. Kabb, W. D. Wei, and B. S. Sumerlin, Near-IR-induced dissociation of thermally-sensitive star polymers, *Chemical Science*, *8* (3), 1815-1821, (2017).
[11] T. K. Goh, K. D. Coventry, A.Blencowe, and G. G. Qiao, Rheology of core cross-linked star polymers, *Polymer*, *49* (23), 5095-5104, (2008).
[12] D. Truzzolillo, D. Vlassopoulos, and M. Gauthier, Osmotic interactions, rheology, and arrested phase separation of

star–linear polymer mixtures, *Macromolecules*, *44* (12), 5043-5052, (2011).

[13] C. Wang, W. He, F. Wang, H.Yong, T. Bo, D. Yao, and M. Li, Recent progress of non-linear topological structure polymers: synthesis, and gene delivery. *Journal of Nanobiotechnology*, *22*(1), 40, (2024)

[14] F. Navaie, E.Esmaeilnezhad, and H.-J. Choi, Effect of rheological properties of polymer solution on polymer flooding characteristics, *Polymers*, *14* (24), 5555, (2022).

[15] L. Gury, M. Gauthier, M. Cloitre, and D. Vlassopoulos, Colloidal jamming in multiarm star polymer melts, *Macromolecules*, *52* (12), 4617-4623, (2019).

[16] T. Ebrahimi, S. G.Hatzikiriakos, and P. Mehrkhodavandi, Synthesis and Rheological Characterization of Star-Shaped and Linear Poly (hydroxybutyrate), *Macromolecules*, *48* (18), 6672-6681, (2015).

[17] F. Snijkers, H. Y. Cho, A. Nese, K. Matyjaszewski, W. Pyckhout-Hintzen, and D.Vlassopoulos, Effects of core microstructure on structure and dynamics of star polymer melts: from polymeric to colloidal response, *Macromolecules*, *47* (15), 5347-5356, (2014).

[18] A. Miros, D.Vlassopoulos, A. Likhtman, and J. Roovers, Linear rheology of multiarm star polymers diluted with short linear chains, *Journal of Rheology*, *47* (1), 163-176, (2003).

[19] P. Polanowski, K. Hałagan, and A.Sikorski, Star polymers vs. dendrimers: Studies of the synthesis based on computer simulations, *Polymers*, *14* (13), 2522, (2022).

[20] S. Liu, J. Zhang, L. R. Hutchings, L. Peng, X. Huang, and Q.Huang, The “solvent” effect of short arms on linear and nonlinear shear rheology of entangled asymmetric star polymers, *Polymer*, *281*, 126125, (2023).

[21] E. Van Ruymbeke, E. B. Muliawan, D.Vlassopoulos, H. Gao, and K.Matyjaszewski, Melt rheology of star polymers with large number of small arms, prepared by crosslinking poly (n-butyl acrylate) macromonomers via ATRP, *European Polymer Journal*, *47* (4), 746-751. (2011).

[22] J. M. Ren, T. G. McKenzie, Q. Fu, E. H. Wong, J. Xu, Z. An, and G. G. Qiao, Star polymers. *Chemical reviews*, *116*(12), 6743-6836. (2016).

[23] A.Mandal, and A. F. Kilbinger, Catalytic Living ROMP: Synthesis of Degradable Star Polymers. *ACS Macro Letters*, *12*(10), 1372-1378. (2023).

[24] K. Sparnacci, T. Frison, E. Podda, D. Antonioli, M. Laus, M. Notari, and R. Po, Core-crosslinked star copolymers as viscosity index improvers for lubricants. *ACS Applied Polymer Materials*, *4*(12), 8722-8730. (2022).

[25] Z. Luo, L. Wang, J. Pei, P. Yu, and B. Xia, A novel star-shaped copolymer as a rheology modifier in water-based drilling fluids, *Journal of Petroleum Science Engineering*, *168*, 98-106. (2018).

[26] G. Jian, , A. Santra, C. Mbuncha, and G. Ross, A novel star polymer for regulating fluid loss in oil-based mud under high temperature conditions, *Journal of Molecular Liquids*, *389*, 122815. (2023).

[27] N. Cakir Yigit, , G. Hızal, and U. Tunca, Synthesis of multiarm star block copolymer based on host-guest inclusion complexation, *Journal of innovative engineering and natural science (Online)*, *2* (1), 1-16. (2022).

[28] D. Srivastva, and A. Nikoubashman, (Flow behavior of chain and star polymers and their mixtures, *Polymers*, *10* (6), 599. (2018).

[29] P. Kratochvíl, and M. Netopilík, On the contraction factors of long-chain branched macromolecules, *European Polymer Journal*, *51*, 177-181. (2014).

El Cezeri
Journal of Science and Engineering
ISSN: 2148-3736

Research Article

Fuzzy Fine-Kinney Occupational Health and Safety Risk Analysis in Offices: An Application in a Public Institution in Istanbul

Ertan Gündoğdu[1a] and Ümran Şengül[2b]
[1]*Social Security Institution, Diyarbakır, 21100, Türkiye*
[2]Department *Department of Business, Faculty of Political Sciences, Çanakkale Onsekiz Mart University, Çanakkale,17000,* Türkiye

umransengul@yahoo.com

DOI : 10.31202/ecjse.1641624
Received: 17.02.2025 Accepted: 16.04.2025
How to cite this article:
Ertan Gündoğdu and Ümran Şengül, " Fuzzy Fine-Kinney Occupational Health and Safety Risk Analysis in Offices: An Application in a Public Institution in Istanbul", El-Cezeri Journal of Science and Engineering, Vol: 12, Iss: 3, (2025), pp.(375-394).
ORCID: [a]0009-0002-6024-1626; [b]0000-0001-5867-863X.

Abstract In offices, there are numerous risks concerning the health and safety of employees, making it challenging to eliminate all risk factors at once. Risk analyses are conducted to prioritize risks based on urgency and to mitigate their danger levels. An occupational health and safety risk analysis was performed using the Fine-Kinney method for 330 hazard sources in a public institution conducting office activities in Istanbul. The uncertainty of parameters in the Fine-Kinney method complicates obtaining objective results; therefore, in the second phase, the parameters of probability, frequency, severity, and risk magnitude were processed according to the principles of fuzzy logic to achieve more objective and realistic outcomes. For the fuzzy analysis, two separate fuzzy models were developed using the same rule base. The difference between the models lies in the selection of membership functions; thus, the impact of membership function choice on the risk score has been explored. Lastly, by comparing Fine-Kinney and fuzzy model analyses, it was concluded that fuzzy models are more successful than the method, and the second fuzzy model, which uses triangular membership functions, was found to be more successful than the first fuzzy model which uses pi membership functions.

Keywords: Fuzzy Fine-Kinney, Risk analysis, Occupational health and safety, Office workplaces.

I. INTRODUCTION

Occupational safety is defined as systematic efforts to cleanse workplaces of hazards arising from work execution and conditions that may harm health, thereby providing a better working environment [1]. The World Health Organization (WHO) defines a work accident as 'an unplanned event, often leading to injuries, damage to machinery and equipment, or a temporary halt in production,' while the International Labor Organization (ILO) describes it as 'an unexpected, unplanned event causing specific damage or injury' [2]. The ILO estimates that approximately 2.3 million people globally succumb annually to work-related accidents or illnesses, equating to over 6000 deaths daily. Around 340 million occupational accidents occur worldwide each year, with 160 million people suffering from work-related diseases. While the ILO periodically updates these estimates, an increase in accidents and health issues over time has been observed [3]. According to the European Statistical Office (EUROSTAT), the EU experiences more than 5500 fatal work accidents annually, with 6.9 million workers suffering accidents and 23 million reporting work-related health problems [4]. In Turkey, according to the Social Security Institution's 2021 data [5], 1382 insured individuals died due to work accidents, and 35 due to occupational diseases. The difficulty in identifying occupational diseases is one of the most problematic areas in the field of Occupational Health and Safety (OHS) in Turkey. Reliable statistics on how many workers in Turkey contract occupational diseases and how many of these lead to chronic problems or deaths are unattainable [6].

Risk, or hazard, is defined as the danger or probability of loss or damage [7]. Risk analysis involves first identifying hazard sources and calculating and ranking the associated risks. In the second step, control measures are determined based on risk scores, and the resulting new situation is re-analyzed to complete the second step. The risk assessment process encompasses these two steps. In the context of occupational health and safety, risk assessment involves identifying potential hazards in the workplace or from external sources, assessing the harm they could cause to employees, the workplace, and the environment, and determining the necessary precautions [1]. Risk assessment methods can be categorized into three main types: quantitative (e.g., Risk Assessment Decision Matrix (RADM), John-Ridley risk assessment method, Failure Mode and Effects Analysis (FMEA), Fine-Kinney risk assessment method, etc.), qualitative (e.g., Preliminary Hazard Analysis (PHA), Primary Risk Analysis (PRA) using Checklists, What if Analysis, Job Safety Analysis, Hazard and Operability Study (HAZOP)), and mixed methods. There are over 150 risk assessment methods in the literature [8].

Offices are places where organizational and managerial activities are carried out. The increasing number of office workers today, along with working conditions, shows that ergonomics and an ergonomic approach significantly affect employee productivity [9]. Office work is not limited to a specific area of the business but spreads across all areas; therefore, defining an office as a place or certain activities is insufficient; it should rather be defined according to its functions [10]. In a business, an office can be a place where different operational activities like production, storage, logistics, etc., are carried out, as well as where all activities in sectors such as banking and insurance are conducted. Risks in offices can be divided into three main groups: physical risk factors, psychosocial risk factors, and ergonomic risk factors [11]. Office occupational diseases can be categorized into four groups: musculoskeletal system diseases, circulatory system diseases, allergic diseases, and psychological diseases; these are a group of multisystemic diseases caused by factors such as prolonged sitting, repetitive movements, computer work, excessive strain on wrists and fingers, ambient temperature, humidity, light, and other non-ergonomic conditions [12].

In this study, occupational health and safety risk analysis was conducted in a public institution engaged in office activities in Istanbul. The data for the study were collected through observations at the workplace. Hazard sources found in the literature were also included in the analysis to evaluate potential but unidentified hazards at the workplace. Initially, risk assessment for a total of 330 hazard sources was conducted using the Fine-Kinney method. The data were then re-analyzed using the Fuzzy Fine-Kinney method. The results obtained are discussed.

In the study, occupational health and safety risk analysis was conducted in a public institution operating in the province of Istanbul. The data of the study were collected by observation at the workplace. To evaluate the hazards that could not be detected at the workplace but could potentially be present, the hazard sources found in the literature were also included in the analysis. In the first stage, a risk assessment was made according to the Fine-Kinney method for a total of 330 hazard sources. Then, the risks were re-analyzed with Fuzzy Fine-Kinney. To conduct risk analysis using fuzzy logic, the probability, frequency, severity, and risk magnitude parameters of the Fine-Kinney method were fuzzified, and two separate fuzzy models were developed. These models employed the same rule base but were differentiated by the selection of membership function types. Subsequently, the analysis results obtained from both the Fine-Kinney and Fuzzy Fine-Kinney models were compared.

In the literature, there are limited studies on risk assessment in office workplaces; these studies are conducted in small offices with a limited number of hazard sources. Furthermore, no studies employing the Fuzzy Fine-Kinney method approach for risk analysis in offices have been found. Thus, this study is expected to contribute to the literature and provide a practical example for occupational health and safety (OHS) professionals working in the field.

The subsequent sections of the study include a review of the literature on risk analysis studies integrating the Fine-Kinney method with other methods, the methodology used in this study, the findings of the risk analysis method applied, and the results and discussions obtained.

II. LITERATURE REVIEW

Some of the studies integrating the Fine-Kinney method with different approaches in the literature include: Özfırat (2016) conducted a risk assessment for the longwall production method, one of the most used methods in underground coal mining. His risk assessment emphasized the need for controlled operation in Grizzly coal mines to prevent the dangerous concentration of methane during production [13]. Kokangül (2017) identified hazard sources in a large machinery manufacturing plant in Turkey and categorized these hazards into main and sub-categories. He prioritized each category using the AHP method and compared the risk scores obtained with the Fine-Kinney method after normalization. His study proposed a new approach that the class intervals of the Fine-Kinney method could be used for the results obtained with the AHP method [14]. Gul et al. (2017) analyzed the maintenance procedures of a ballast tank on a ship. The probability, frequency, and severity parameters of the Fine-Kinney method were weighted using fuzzy AHP, and the risk score was calculated using fuzzy VIKOR, identifying the highest risks as "high temperature inside the tank," "head injury at the manhole entrance of the tank," and "falling from height during welding or cutting operations in the ballast tank" [15]. Oturakçı and Dağsuyu (2017) evaluated 10 risks at a construction site using the Fine-Kinney method, then fuzzified the method's scales with triangular membership functions to create a fuzzy model. They found that the priority levels of 5 out of 10 risks changed after comparing the risk levels obtained from the Fine-Kinney and fuzzy models, arguing that the fuzzy model provided more precise results [16]. Gönen (2018) rescaled the musculoskeletal system discomfort survey developed by Cornell University, inspired by the Fine-Kinney method, and applied it to a company in the automotive sector producing cables, identifying the musculoskeletal discomforts of assembly line workers [17]. Gul (2018) applied the fuzzy analytic hierarchy process to weigh the parameters of the Fine-Kinney method and then used the fuzzy VIKOR method to prioritize the hazards. In his study, OHS experts in Turkey weighted risk parameters for an onshore wind turbine and evaluated the agreed rankings of hazards, revealing the most significant dangers during the construction and operation phases of the wind turbine. The most critical hazards during the construction of the turbine were identified as the lack of safety belts, falling from heights, panic in emergencies, and inability to respond quickly in emergencies. The risks during the operation of the wind turbine were identified as vehicle accidents due to damaged and uneven roads and the shock risk from unauthorized digging. The study concluded with control measures and potential corrective and preventive activities for the risks [18]. Ilbahar (2018) used a new integrated approach, the Pythagorean Fuzzy Proportional Risk Assessment (PFPRA), combining Fine-Kinney, Pythagorean Fuzzy Analytic Hierarchy Process, and Fuzzy Inference System for risk assessment during excavation operations at a construction site. The results identified the most critical factor as the undefined routes of excavation vehicles on-site and the least critical as the unlicensed drivers of excavation vehicles. Landslide risk due to the lack of proper slope or support was found to be a significant risk, while other critical factors were negligible. The proposed method was compared with Pythagorean fuzzy failure modes and effects analysis, showing that the proposed method provided more informative, reliable, and consistent results for decision-makers regarding uncertainty [19]. Weizhong (2018) proposed a new risk assessment approach based on the Fine-Kinney method, combining triangular fuzzy numbers, the MULTIMOORA method, and the Choquet integral. The study selected the maintenance of a ballast tank as a case to verify the effectiveness of the proposed risk assessment model through comparison and sensitivity analysis [20]. Taranushina and Popova (2019] assessed the risks involved in organizing loading and unloading operations using 20/5 ton capacity bridge cranes with the Fine-Kinney method [21]. Baç and Ekmekci (2020) assessed the psychosocial risks faced by employees in a metal processing plant using the COPSOQII questionnaire, with the data obtained from the COPSOQII survey used as input for Fine-Kinney analysis. Due to the uncertainty in the Fine-Kinney method, this was overcome using the ANFIS (adaptive neuro-fuzzy inference system) module. The model was compared with the actual risks, demonstrating that the ANFIS results predicted the risk scores with high accuracy [22]. Erdebilli and Gür (2020) evaluated a total of 8 risks, both man-made and natural disasters, for a dam in eastern Turkey using both the Fine-Kinney and fuzzy Fine-Kinney methods and compared the results. They found that the importance level of 5 out of the 8 risks changed in the analysis [23]. Aboubakar et al. (2021) identified significant risk factors in Chinese hydroelectric projects in Cameroon and ranked these factors based on their impact on project success, deriving from literature and experts, under four categories. They integrated the Fine-Kinney and SWARA-TOPSIS methods. As a result, the highest risk categories were identified as social and environmental risks; moreover, the highest risk factor was found to be social acceptance by the local population, while restrictions related to environmental issues were the lowest risk factors [24]. Güney and Kahraman (2022) conducted a risk assessment in environmental research laboratories using the Analytical Hierarchy Process (AHP) and Fine-Kinney method. They determined that the cost of implementing safety measures for each risk was approximately 10,000.00 EUR [25]. Tatar et al. (2023) proposed an approach integrating the Fine-Kinney and spherical fuzzy AHP-TOPSIS methods to assess the risks of musculoskeletal disorders in workers harvesting tea. In the first stage, the Fine-Kinney parameters were weighted using spherical fuzzy AHP (SF-AHP). Subsequently, the spherical fuzzy TOPSIS (SF-TOPSIS) method was used to rank the hazards. In the final stage, the proposed model was compared with the SF-CODAS (spherical fuzzy composite distance-based assessment) method based on Fine-Kinney. The results obtained from the SF-TOPSIS and SF-CODAS methods based on Fine-Kinney were analyzed using Spearman's rank correlation coefficient, and the similarity rate in risk rankings was determined to be 92% [26].

III. METHODOLOGY

"In this section, the Fine-Kinney method, fuzzy logic, fuzzy sets, membership functions, and fuzzy model design in MATLAB are described.

3.1. Fine-Kinney method

The Fine-Kinney method was developed by William T. Fine in 1971 for the American Naval Ordnance Laboratory in Maryland. The method was presented to management through a technical report titled 'Mathematical Evaluations for Controlling Hazards.' In the Fine-Kinney method, the Risk Magnitude (RM) is calculated using three parameters: the probability of the accident (P), the potential consequences of an accident (severity-S), and exposure (frequencies of the actions-F) [27].

$$RM = PxSxF \tag{1}$$

In Equation (1), P represents the probability that, once a hazardous event occurs, the complete sequence of events will follow with the necessary timing and coincidence to lead to an accident and its consequences [27]. These values are categorical, and their numerical magnitudes are selected from standard tables as shown in Table 1. a [28]. In Equation (1), S represents the most probable outcomes of a potential accident, including injuries and material damage. This is based on an assessment of the entire situation surrounding the hazard and on accident experience [27]. S signifies the possible outcomes of an accident that may occur, such as an electric shock leading to injury or death while a worker is operating an electrical panel. In the Kinney method, a verbal value such as death, injury, first aid, property damage, etc., is assigned to the severity parameter according to the characteristics of the possible accident. The S scale is shown in Table 1. b [16]. In Equation (1), F represents the frequency of occurrence of the hazardous event [27]. The F scale indicates that the more frequent the exposure to a potentially hazardous situation, the greater the risk. Therefore, when creating the F scale, a value of 1 is assigned for an exposure occurring once a year, and a value of 10 for continuous exposure, with interpolation providing the intermediate values between these two [28]. The frequency scale is displayed in Table 1. c, with reference points highlighted in bold [16].

Table 1. (a) Linguistic scale for probability; (b) Linguistic scale for severity; (c) Linguistic scale for frequency of occurrence

(a)

Description	Rating
Most likely	**10**
Quite possible	6
Unusual	3
Remotely possible	**1**
Extremely remote	0,5
Practically impossible	0,2
Almost impossible	**0,1**

(b)

Description	Rating
Catastrophe	**100**
Multiple fatalities	40
Fatality	15
Extremely serious	7
Disabling injuries	3
Minor cuts	**1**

(c)

Linguistic terms	Rating
Continuously	**10**
Frequently	6
Occasionally	3
Unusually	2
Rarely	**1**
Very rarely	0,5

In the Fine-Kinney method, the results obtained based on the probability, frequency, and severity parameters are evaluated using the linguistic scale provided in Table 2 [28].

Table 2. Linguistic scale for risk magnitude

Risk level	**Description**	**Risk value**
Critical	Immediate action	R>=400
Major Risk	Immediate improvement	200<=R<400
Medium Risk	Measures to be taken	70<=R<200
Minor Risk	Monitoring	20<=R<70
Negligible Risk	Acceptable risk	R<20

3.2. Fuzzy Logic

Traditional two-valued logic falls short in providing an appropriate conceptual framework for defining the meanings of fuzzy concepts like approximate and intense. Hence, the concept of fuzzy logic was developed. Fuzzy logic was first introduced by Lotfi A. Zadeh in 1965 in an article titled "Fuzzy Sets" published in the journal Information and Control [29,30]. Fuzzy logic offers some advantages in modeling complex processes where logic rules are expressed using linguistic variables, and information is often subjective, incomplete, or unreliable, with the problem typically being non-linear.

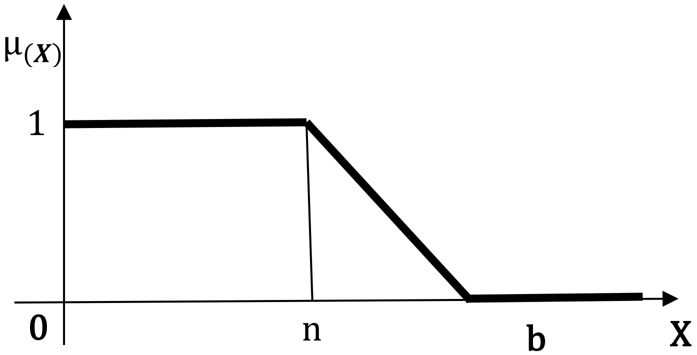


Figure 1. Linear Z Membership Function (Jevscek, 2016)

For example, in risk assessment, fuzzy logic holds good potential, especially in studies where probability assessment largely relies on expert opinion [31]. Fuzzy logic does not reject the concept of membership used in classical set theory; instead, it expands it by moving from two-valued to multi-valued membership [32, 33]. In fuzzy sets, there are many membership functions such as triangular fuzzy numbers, trapezoidal fuzzy numbers, Gaussian fuzzy numbers, and sigmoidal fuzzy numbers [34]. In this study, linear z, pi, triangular, and linear s membership functions have been used. The mathematical definitions of these membership functions are shown below.

Linear Z Membership Function: The linear Z membership function can be expressed as in equation (2) with two variables n and b [35]. Figure 1 shows the linear Z membership function.

$$\mu(\boldsymbol{x}) = \mu A(x; n, b) = \begin{cases} 1, & if\ x < n \\ \dfrac{n-x}{n-b}, & if\ n \le x \le b \\ 0, & if\ x > b \end{cases} \tag{2}$$

Triangular Membership Function: A triangular fuzzy number can be expressed with three variables a_1, a_2, and a_3 as in equation (3) [36]. Figure 2 shows the triangular membership function.

$$\mu_{\hat{A}} = \begin{cases} \dfrac{x-a_1}{a_2-a_1}, & if\ a_1 \le x \le a_3 \\ \dfrac{x-a_2}{a_3-a_2}, & if\ a_2 \le x \le a_3 \\ 0\ , & if\ x < a_1 veya\ x > a_3 \end{cases} \tag{3}$$

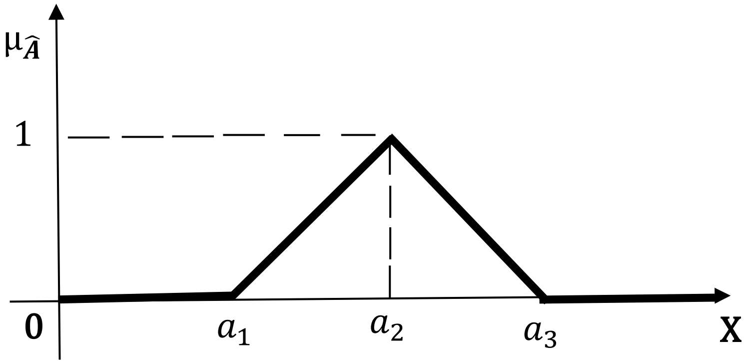


Figure 2. Triangular membership function (Guttorp, 1990).

Pi Membership Function: The fuzzy number Pi can be expressed with variables a, b, c, and d as in equation (4) [37]. Figure 3 shows the pi membership function for a=2, b=4, c=6, d=8.

$$\mu X = \begin{cases} \begin{cases} 0 & , \quad if\ x \le a\ veya\ x \ge d \\ 2\left(\dfrac{x-a}{b-a}\right)^2 & , \quad if\ a \le x \le \dfrac{a+b}{2} \\ 1-2\left(\dfrac{x-b}{b-a}\right)^2 & , \quad if\ \dfrac{a+b}{2} \le x \le b \end{cases} \\ \qquad 1 & , \quad if\ b \le x \le c \\ \begin{cases} 1-2\left(\dfrac{x-c}{d-c}\right)^2 & , \quad if\ c \le x \le \dfrac{c+d}{2} \\ 2\left(\dfrac{x-d}{d-c}\right)^2 & , \quad if\ \dfrac{c+d}{2} \le x \le d \end{cases} \end{cases} \tag{4}$$

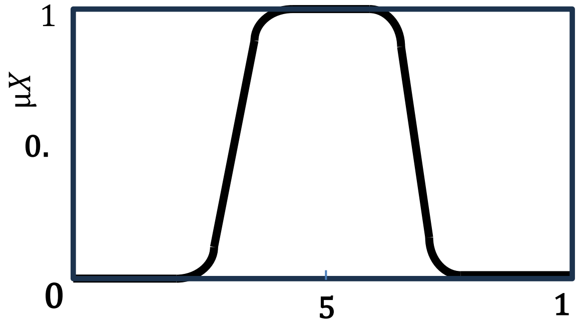


Figure 3. Pi membership function

Linear S Membership Function: The linear s membership function can be expressed with two variables a and m as in equation 5 [35]. Figure 4 shows the Linear s membership function.

$$\mu A(x) = \mu A(x; a, m) = \begin{cases} 0, & \text{eğer } x < a \\ \dfrac{x-a}{m-a}, & \text{eğer } a \le x \le m \\ 1, & \text{eğer } x > m \end{cases} \qquad (5)$$

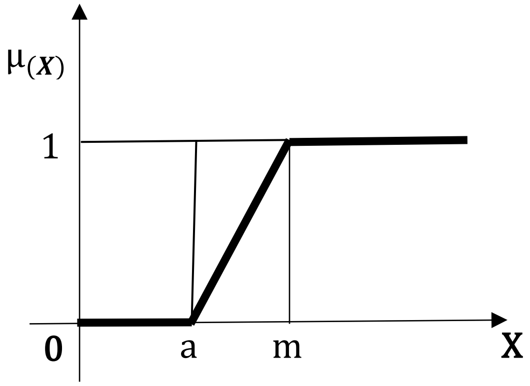


Figure 4. Linear S membership function [35].

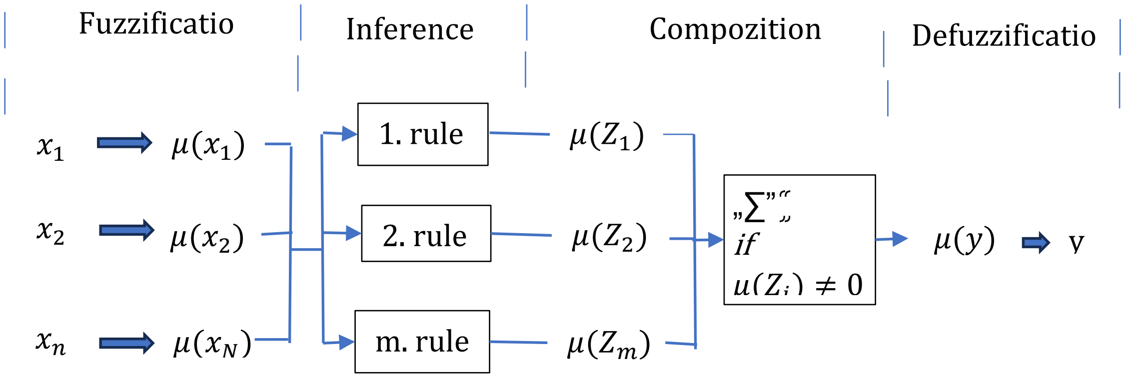


Figure 5. Fuzzy logic decision process [35]

3.3. Fuzzy Inference System

Fuzzy logic-based systems are a combination of four sub-processes: fuzzification, inference, aggregation, and defuzzification. The system uses several rules simultaneously in these processes. Figure 5 illustrates the operational stages of a fuzzy inference system [38].

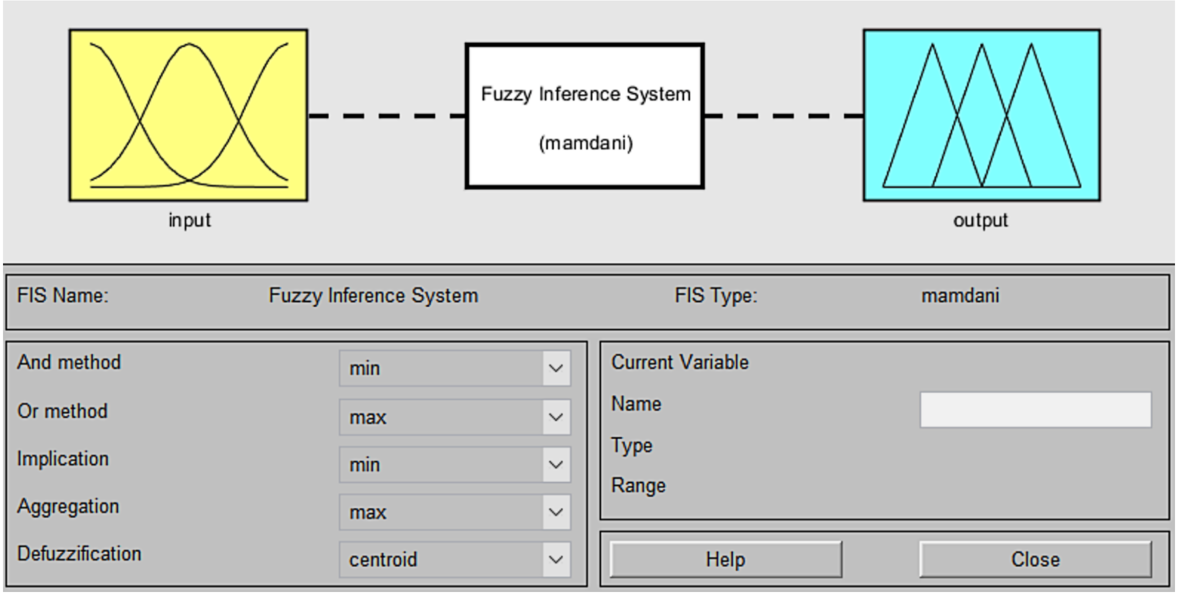
In a fuzzy-based system, the first stage of the decision process is fuzzification, where the membership degrees of each input value, such as $x_1, x_2, \ldots x_n$, are calculated over fuzzy sets, denoted as $\mu(x_1)$, $\mu(x_2)$, … $\mu(x_n)$, thus completing the fuzzification process. In the second stage of the process, the system calculates the degree of truth (triggering power of the rule) for each membership degree expressed as; $\mu(x_1)$, $\mu(x_2)$, … $\mu(x_n)$ according to each rule in the rule base. The triggering power of each rule by the membership degrees is represented as $\mu(z_1)$, $\mu(z_2)$, … $\mu(z_n)$. In the third stage, the system aggregates the fuzzy set outputs; $\mu(z_1)$, $\mu(z_2)$, … $\mu(z_n)$ based on the principle of union (aggregation), hence the output of the third stage is not a clear value but a set of values. In the final stage, the defuzzification process is performed to obtain a crisp (precise) numerical value [38]. Among the fuzzy inference system (FIS) types, Mamdani and Sugeno, Mamdani is one of the first fuzzy control systems in the literature and is one of the most widely used inference systems [19].

IV. IMPLEMENTATION

In this study, risk analysis was conducted as part of the risk assessment process. A public institution in Istanbul, conducting office activities, was selected for the risk analysis. The workplace building consists of eight floors, including a basement and an attic. Additionally, there is a two-story annex building that houses a cafeteria and conference hall integrated with the main building. The workplace is situated within a large garden, featuring parking and green areas, located between a major public hospital and a university campus. Approximately 200 employees work at the site. Interviews with employees revealed that each worker receives an average of 10 visitors daily, leading to an estimate of around 2000 people visiting the office building daily for services. In Turkey, workplaces are classified into three hazard categories—less hazardous, hazardous, and highly hazardous—according to the "Workplace Hazard Classes Notification" published by the Ministry of Labor. Workplaces operating in the office, insurance, and banking sectors are considered less hazardous; however, especially in large offices, risks such as elevator accidents, slips and falls, and electrical shocks among employees are present, which can lead to injuries and even deaths, similar to those in highly hazardous workplaces. Therefore, like in other workplaces, risk analysis in office workplaces must be conducted with precision.

4.1. Fuzzy Pattern Design

The fuzzy model design was carried out using the fuzzy logic toolbox of the MATLAB program. The general view of the system is shown in Figure 6. In Figure 6, the input represents the entry point of the system; in this study, three inputs have been defined, one each for the parameters of severity, frequency, and probability. The output is the exit point of the system; in this study, one output has been defined to represent the Risk Priority Number (RPN). The "Mamdani Inference System" was used as the inference system in the design of the fuzzy model. For the Mamdani inference method, the operators selected are "min" for the and method, "max" for the Or method, "min" for Implication, "max" for Aggregation, and "centroid" for Defuzzification.



FIS Name: Fuzzy Inference System
FIS Type: mamdani
And method: min
Or method: max
Implication: min
Aggregation: max
Defuzzification: centroid

Figure 6. Fuzzy model overview

Table 3. First fuzzy model membership functions

Parameter	*Scale of Parameter*	*Label Given to the Function*	*Membership Function*
Probability	**Almost Impossible**	AI	zmf [0 0.2]
	Practically Impossible	PI	pimf [0.1 0.2 0.4 0.5]
	Extremely Remote	ER	pimf [0.4 0.5 0.9 1]
	Remotely Possible	RP	pimf [0.9 1 2.9 3]
	Unusual	U	pimf [2.9 3 5.9 6]
	Quite Possible	QP	pimf [5.9 6 9.9 10]
	Most Likely	ML	smf [9.9 10]
Frequency	Very Rarely	VR	zmf [0.5 1]
	Rarely	R	pimf [0.5 1 1.5 2]
	Unusually	U	pimf [1.5 2 2.5 3]
	Occasionally	O	pimf [2.5 3 5.5 6]
	Frequently	F	pimf [5.5 6 9.5 10]
	Continuously	C	smf [9.5 10]
Severity	**Minor Cuts**	MC	zmf [0 3]
	Disabling Injuries	DI	pimf [1 3 5 7]
	Extremely Serious	ES	pimf [5 7 13 15]
	Fatality	F	pimf [13 15 38 40]
	Multiple Fatalities	MF	pimf [38 40 98 100]
	Catastrophe	C	smf [98 100]
Risk Magnitude	Negligible Risk	NR	zmf [0 30]
	Minor Risk	MNR	pimf [0 30 60 90]
	Medium Risk	MDR	pimf [60 90 180 210]
	Major Risk	MJR	pimf [180 210 290 320]
	Critical	C	smf [290 720]

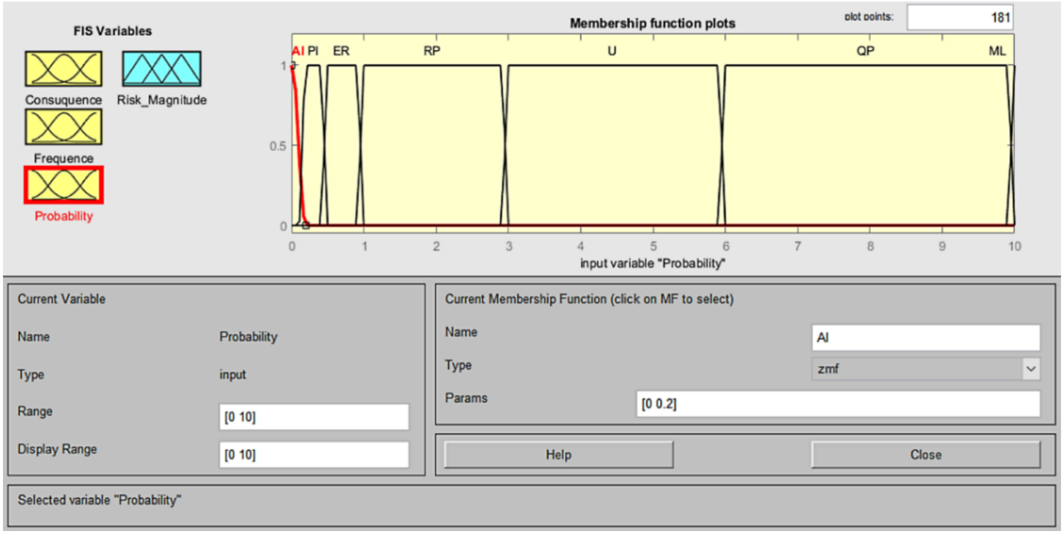
4.2. Rule Base

In the Mamdani inference system, each rule is defined as "if x is A, and if y is B, and … then z is C". The rule base was created in collaboration with an occupational safety specialist engineer. In the Fine-Kinney method, since the input parameters of severity, frequency, and probability can take 6, 6, and 7 different values respectively, a maximum of 6x6x7=252 rules can be defined. To ensure the evaluation of all combinations, all 252 rules have been defined in this study. The same rule base is used in both models; however, different membership functions have been selected.

4.3. First Fuzzy Model

In the first fuzzy model, linear z (linzmf), Pi (pimf), and linear s (linsmf) membership functions were used. Since the highest risk value in the Fine-Kinney analysis is 720, the system's output RPN value was set to 720. The membership functions for the parameters of the first fuzzy model, the labels given to the functions, and the values used in the membership functions are shown in Table 3.

In the designed first model, the graphical representations of the membership functions for the system's input parameters (probability, frequency, severity) and the output parameter RPN are shown in Figure 7, Figure 8, Figure 9, and Figure 10, respectively.



plot points: 181

Current Variable

Name: Probability

Type: input

Range: [0 10]

Display Range: [0 10]

Current Membership Function (click on MF to select)

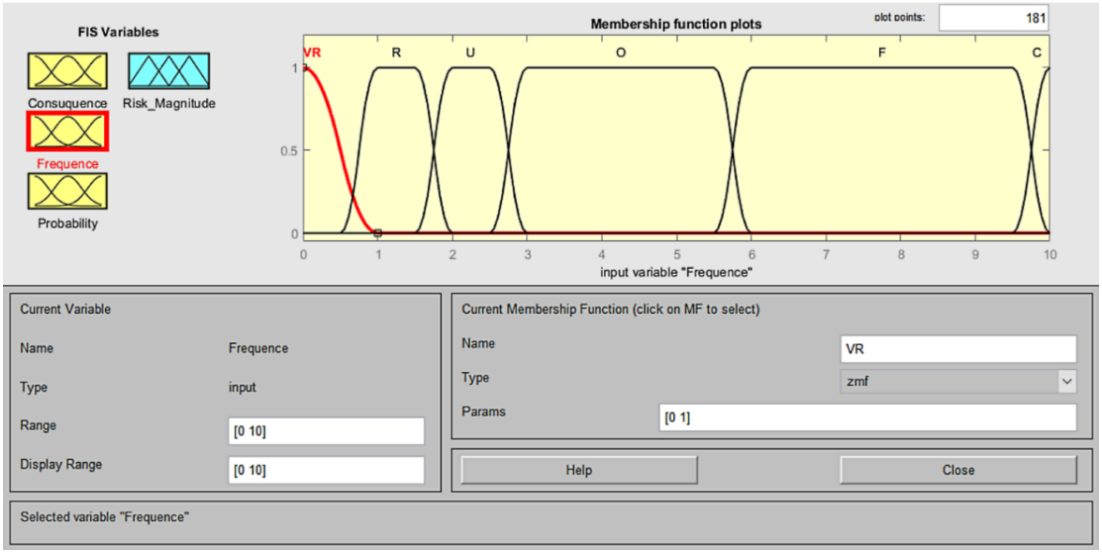
Name: AI

Type: zmf

Params: [0 0.2]

Selected variable "Probability"

Figure 7. Probability membership functions in the first model



plot points: 181

Current Variable

Name: Frequence

Type: input

Range: [0 10]

Display Range: [0 10]

Current Membership Function (click on MF to select)

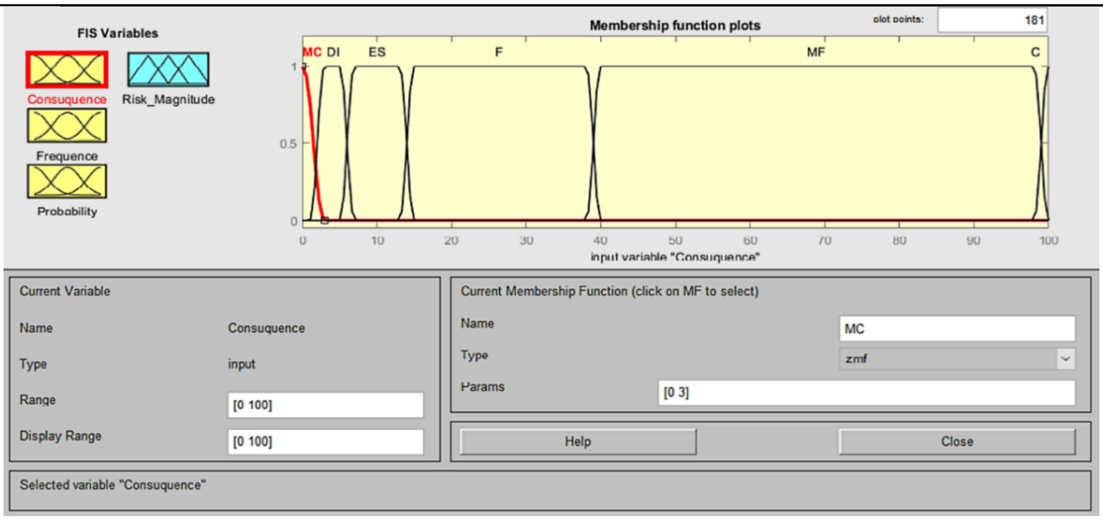
Name: VR

Type: zmf

Params: [0 1]

Selected variable "Frequence"

Figure 8. Frequency membership functions in the first model

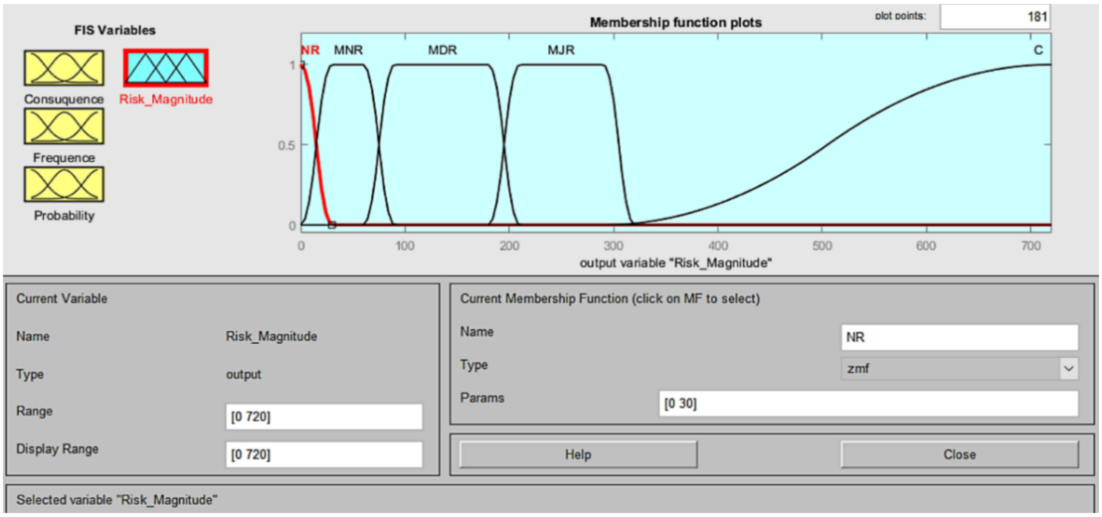


Current Variable
Name: Consuquence
Type: input
Range: [0 100]
Display Range: [0 100]

Current Membership Function (click on MF to select)
Name: MC
Type: zmf
Params: [0 3]

Selected variable "Consuquence"

Figure 9. Severity membership functions in the first model



Current Variable
Name: Risk_Magnitude
Type: output
Range: [0 720]
Display Range: [0 720]

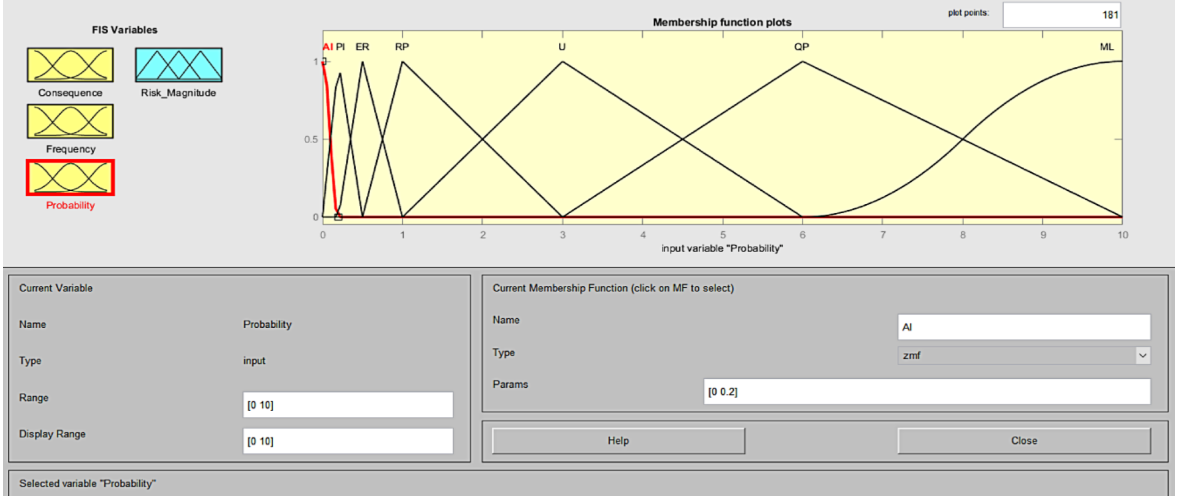
Current Membership Function (click on MF to select)
Name: NR
Type: zmf
Params: [0 30]

Selected variable "Risk_Magnitude"

Figure 10. Risk magnitude membership functions in the first model

4.4. Second Fuzzy Model

In the second fuzzy model, linear z (linzmf), triangular (trimf), and linear s (linsmf) membership functions were used. Since the highest risk value in the Fine-Kinney analysis is 720, the range value for the system's output (RPN) was set to 720. The membership functions for the parameters of the first fuzzy model, the labels given to the functions, and the values used in the membership functions are shown in Table 4.



Current Variable
Name: Probability
Type: input
Range: [0 10]
Display Range: [0 10]

Current Membership Function (click on MF to select)
Name: AI
Type: zmf
Params: [0 0.2]

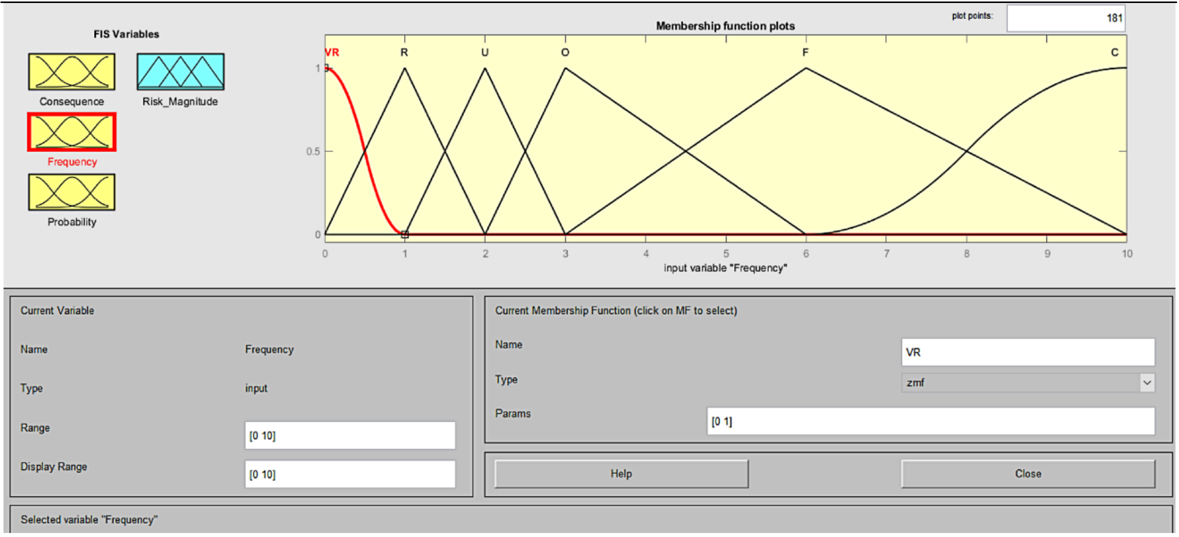
Selected variable "Probability"

Figure 11. Probability membership functions in the second model

In the designed second model, the graphical representations of the membership functions for the system's input parameters (probability, frequency, severity) and the output parameter RPN are shown in Figure 11, Figure 12, Figure 13, and Figure 14, respectively.

Table 4. Second fuzzy model membership functions

Parameter	*Scale of Parameter*	*Label Given to the Function*	*Membership Function*
Probability	**Almost impossible**	AI	zmf [0 0.2]
	Practically impossible	PI	trimf [0.1 0.2 0.5]
	Extremely remote	ER	trimf [0.2 0.5 1]
	Remotely possible	RP	trimf [0.5 1 3]
	Unusual	U	trimf [1 3 6]
	Quite possible	QP	trimf [3 6 10]
	Most likely	ML	smf [6 10]
Frequency	Very rarely	VR	zmf [0 1]
	Rarely	R	trimf [0 1 2]
	Unusually	U	trimf [1 2 3]
	Occasionally	O	trimf [2 3 6]
	Frequently	F	trimf [3 6 10]
	Continuously	C	smf [6 10]
Severity	**Minor cuts**	MC	zmf [0 3]
	Disabling injuries	DI	trimf [0 3 7]
	Extremely serious	ES	trimf [3 7 15]
	Fatality	F	trimf [7 15 40]
	Multiple fatalities	MF	trimf [15 40 100]
	Catastrophe	C	smf [40 100]
Risk Magnitude	Negligible Risk	NR	zmf [0 70]
	Minor Risk	MNR	trimf [0 70 200]
	Medium Risk	MDR	trimf [70 200 300]
	Major Risk	MJR	trimf [200 300 400]
	Critical	C	smf [300 400]



plot points: 181

Current Variable
Name Frequency
Type input
Range [0 10]
Display Range [0 10]

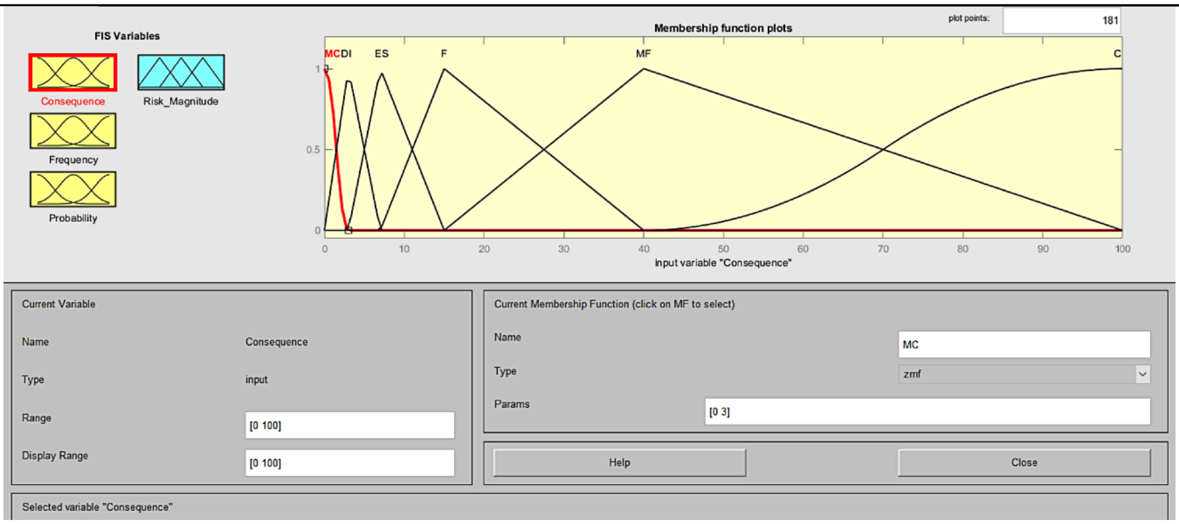
Current Membership Function (click on MF to select)
Name VR
Type zmf
Params [0 1]

Selected variable "Frequency"

Figure 12. Frequency membership functions in the second model

4.5. Analysis in Matlab

The input parameters (probability, frequency, and severity) of the Fine-Kinney method were saved in an Excel file named parameters. The designed fuzzy models were digitally named and saved as the first model and the second model. To perform the fuzzy analyses, the Excel file named inputs parameters containing the input parameters was first imported into the MATLAB screen. Then, fuzzy analysis results were obtained by entering the following code snippet into the command line. The code below uses the fuzzy model named the first model in the analysis. To use the second model in the analysis, the same steps are followed, and only the first model needs to be replaced with the second model. The blue parts starting with // on the right side of the code snippet are not part of the code but are merely explanatory. input=parameters; // It defines the file named parameters as input to Matlab fis=readfis('first_model');// Selects the model to be used (first model or second model) output = evalfis(fis , input)// Evaluates inputs with the selected model

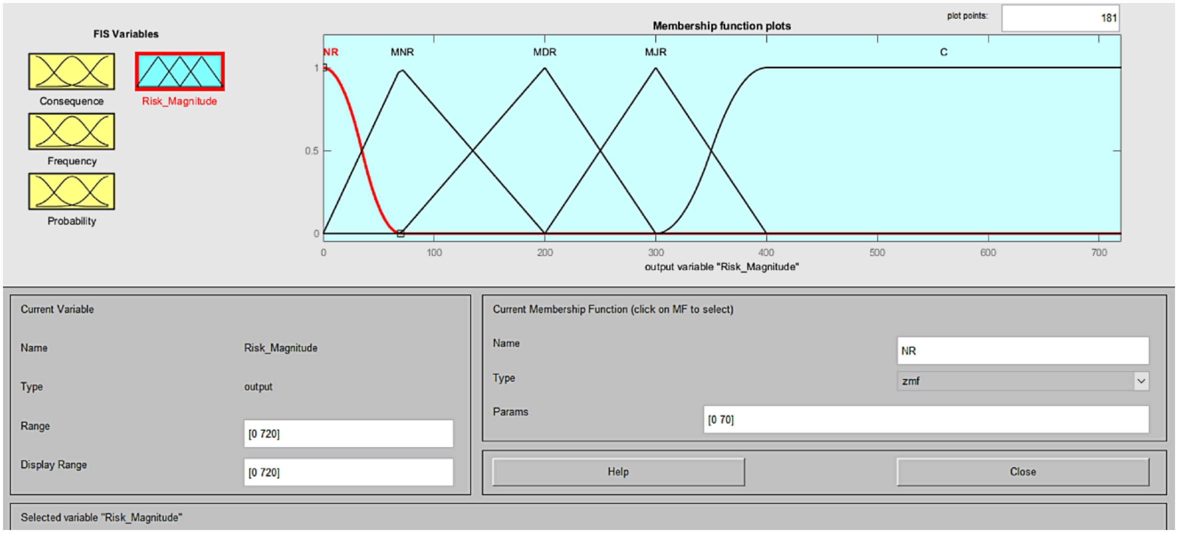


Current Variable
Name Consequence
Type input
Range [0 100]
Display Range [0 100]

Current Membership Function (click on MF to select)
Name MC
Type zmf
Params [0 3]

Selected variable "Consequence"

Figure 13. Severity membership functions in the second model



Current Variable
Name Risk_Magnitude
Type output
Range [0 720]
Display Range [0 720]

Current Membership Function (click on MF to select)
Name NR
Type zmf
Params [0 70]

Selected variable "Risk_Magnitude"

Figure 14. Risk magnitude membership functions in the second model

V. RESULTS

The RM and RPN levels obtained from the analyses for Fine-Kinney and fuzzy Fine-Kinney are comparatively shown in Table 7 attached below. In Table 7, hazard sources are named from H1 to H330, and the probability parameter is abbreviated as P, the frequency parameter as F, and the severity parameter as C. In Table 7, the risk levels vary between 1 and 5. Level 1 is assigned a light green color, level 3 an orange color, and level 5 a dark red color. Levels 2 and 4 are intermediate values, represented by shades of the mentioned colors (a green-orange shade for level 2 and an orange-red shade for level 4).

5.1. Fine-Kinney Analysis

In the analysis conducted with the Fine-Kinney method, risk levels were distributed between 1 and 5. The analysis observed all risk levels from "acceptable risk" to "very high risk".

Table 5. Numerical distribution of risks to levels in Fine-Kinney analysis

Risk Level	Risk Class	Risk Number
1	Acceptable risk	102
2	Possible risk	87
3	Significant risk	49
4	High risk	49
5	Very high risk	43

Most of the level 5 risks arise from electrical hazards. These risks are followed by those related to elevators, ergonomics, lifting loads, and personnel service vehicles, respectively. Table 5 shows the numerical distribution of **risk levels** according to the Fine-Kinney method analysis.

5.2. First Fuzzy Model Analysis

In the analysis conducted with the first fuzzy model, risk levels were distributed between 1 and 5. The analysis observed all risk levels from "acceptable risk" to "very high risk". Most of the level 5 risks arise from electrical hazards. In this analysis, other risks contributing to the density at level 5 were identified as elevator, ergonomics, lifting loads, and personnel service vehicle-related risks, in order.

When the risks evaluated by the Fine-Kinney Method were assessed using the first fuzzy model, it was observed that the level of 76 out of 330 risks (approximately 23%) changed. Risks with at least three levels of change include: The risk of electric shock in elevator number H97 rose from level 1 to level 5. The risk associated with elevator number H98 stopping before reaching the floor and the intervention by individuals who have not received the relevant training dropped from level 5 to level 2. The risk due to insufficient cabin lighting in elevator number H102 rose from level 1 to level 4. The risk of electric shock in the machine room of elevator number H107 dropped from level 5 to level 1. The risk of personnel from authorized companies maintaining elevator number H113 not using appropriate personal protective equipment dropped from level 5 to level 2.

5.3. Second Fuzzy Model Analysis

In the analysis conducted with the second fuzzy model, risk levels were distributed between 1 and 5. The analysis observed all risk levels from "acceptable risk" to "very high risk". A significant portion of the level 5 risks were due to electrical hazards. In this analysis, other predominant risks at level 5 were identified as those related to the boiler room, elevators, ergonomics, and lifting loads, in order.

When the risks evaluated by the Fine-Kinney Method were assessed using the second fuzzy model, it was observed that the level of 196 out of 330 risks (approximately 59%) changed. The risk of electric shock in elevator number H97 rose from level 1 to level 5. The risk due to insufficient cabin lighting in elevator number H102 rose from level 1 to level 4. The risk of electric shock in the machine room of elevator number H107 dropped from level 5 to level 2.

5.4. Comparison of Analysis

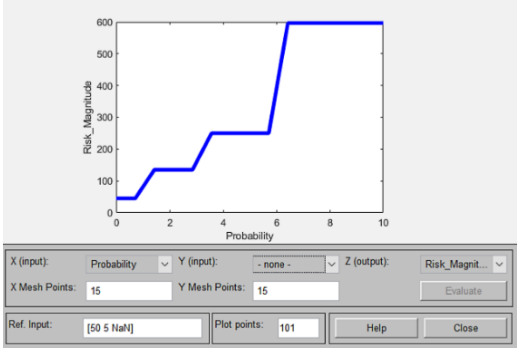
a) It is observed that all three analyses can differentiate risk levels up to the maximum level of 5. From this perspective, it can be said that all three analyses are successful in risk assessment.

b) In the analyses conducted with fuzzy models, it has been observed that the level of some risks identified in the Fine-Kinney analysis increased by 2, 3, or even 4 levels. For example, the risk of electric shock in elevator H97, calculated as level 1 in the Fine-Kinney analysis, was assessed as level 5 in the fuzzy models, moving from an acceptable risk level to a very high-risk level, thereby increasing the urgency for taking control measures for risk H97. Other significant changes from the Fine-Kinney method to fuzzy models include the following: The risk of panic, tripping, falling, and injury due to insufficient lighting in elevator cabin H102 increased from level 1 to level 4. The risk of falling into the elevator shaft, being severely injured, or dying due to the opening of floor and cabin doors of elevator H105 before reaching the floor rose from level 2 to level 4. The risk of overheating, malfunction, fire, and material damage due to inadequate ventilation of the elevator machine room H112 increased from level 3 to level 5. Meanwhile, the levels of some risks decreased. For instance, the risks of workplace accidents, injuries, material damage, and death due to untrained persons intervening in stopped elevator H98 and the risks of workplace accidents, electric shock, falling into the elevator shaft, severe injury, and death due to personnel of authorized companies maintaining elevator H113 not using appropriate personal protective equipment dropped from level 5 to level 2 according to the first fuzzy model, and to level 3 according to the second fuzzy model. The risk of electric shock in the machine room of elevator H107 decreased from level 5 to level 1 according to the first fuzzy model, and to level 2 according to the second fuzzy model.

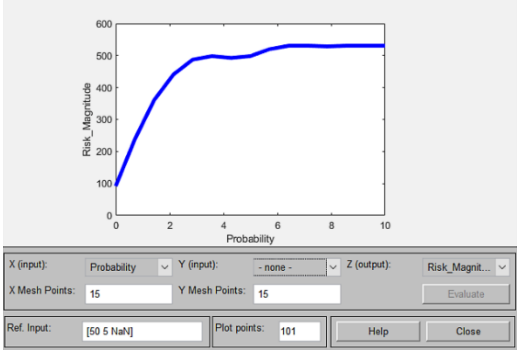
c) Fuzzy models are observed to be more successful than the Fine-Kinney method because they can prevent the possibility of having the same risk levels in different combinations of Fine-Kinney parameters for risks that may have different levels of importance depending on the nature of the workplace. For example, the risk scores for hazards H158, H172, H185, H209, H245, and H280 were calculated with different combinations of probability, frequency, and severity values, resulting in an RPN value of $R=0.5x1x3=1.5$, and these risks were assessed as level 1 risks; however, in the analysis with the second fuzzy model, the risk of neglecting cleanliness and hygiene rules during food production H185 rose to level 2, differentiating it from the other risks.

d) Fuzzy models are considered to be more successful than the Fine-Kinney method because they make subjective evaluations, which can lead to quite different results depending on the experience and opinions of the analyst, are more objective, and allow for more realistic results.

e) The impact of the input parameters on the risk score (RPN) in the designed fuzzy models is shown below in two and three dimensions; on the left for the first fuzzy model, and on the right for the second fuzzy model. In the graphs, the median value of the parameter not shown is used; for example, in Figure 15-26, while plotting the impact of the probability parameter on the RPN, the severity parameter is set to a value of 50, and the frequency parameter to a value of 5, as shown in the Ref Input box, considering these values as references.



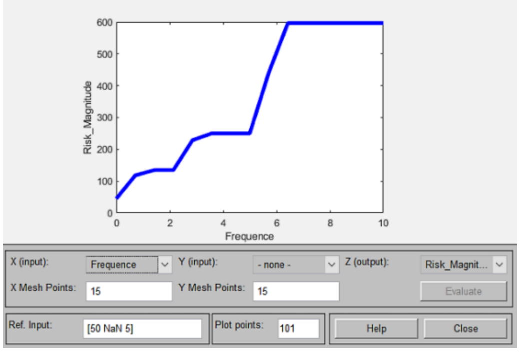
X (input): Probability
Y (input): - none -
Z (output): Risk_Magnit...
X Mesh Points: 15
Y Mesh Points: 15
Ref. Input: [50 5 NaN]
Plot points: 101



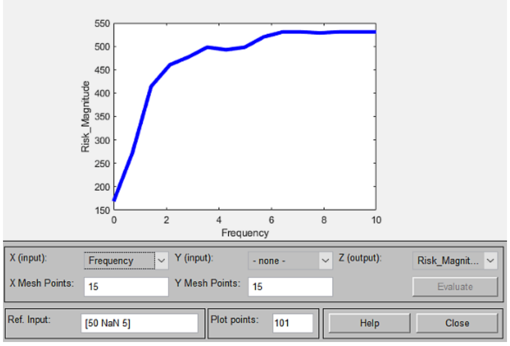
X (input): Probability
Y (input): - none -
Z (output): Risk_Magnit...
X Mesh Points: 15
Y Mesh Points: 15
Ref. Input: [50 5 NaN]
Plot points: 101

Figure 15. The effect of probability on risk magnitude at First Fuzzy Model

Figure 16. The effect of probability on risk magnitude at Second Fuzzy Model



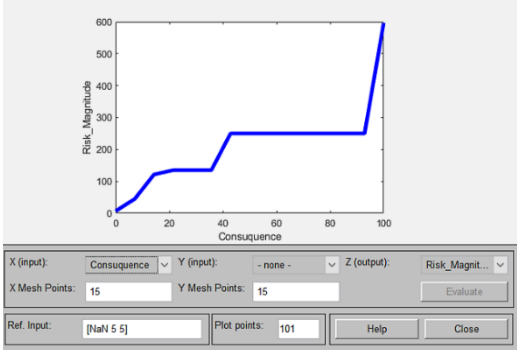
X (input): Frequence
Y (input): - none -
Z (output): Risk_Magnit...
X Mesh Points: 15
Y Mesh Points: 15
Ref. Input: [50 NaN 5]
Plot points: 101



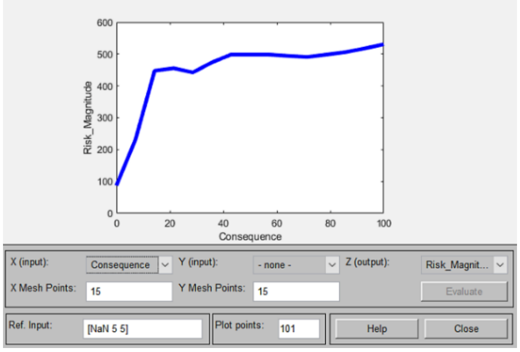
X (input): Frequency
Y (input): - none -
Z (output): Risk_Magnit...
X Mesh Points: 15
Y Mesh Points: 15
Ref. Input: [50 NaN 5]
Plot points: 101

Figure 17. Effect of frequency on risk magnitude at First Fuzzy Model

Figure 18. Effect of frequency on risk magnitude at Second Fuzzy Model



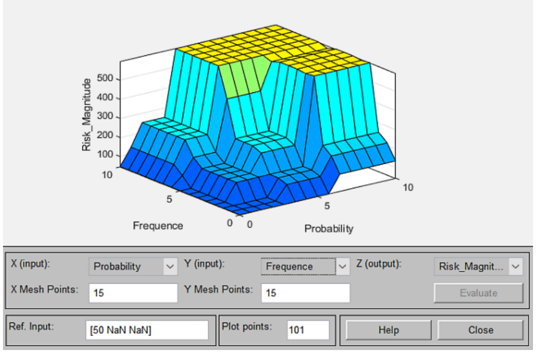
X (input): Consuquence
Y (input): - none -
Z (output): Risk_Magnit...
X Mesh Points: 15
Y Mesh Points: 15
Ref. Input: [NaN 5 5]
Plot points: 101



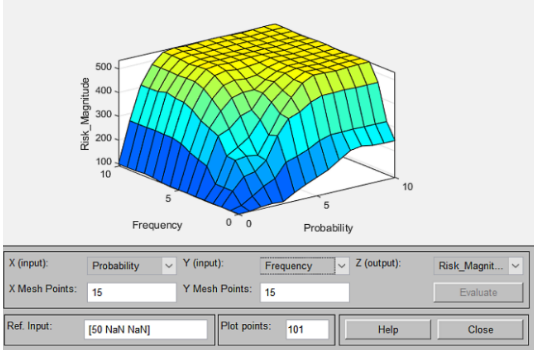
X (input): Consequence
Y (input): - none -
Z (output): Risk_Magnit...
X Mesh Points: 15
Y Mesh Points: 15
Ref. Input: [NaN 5 5]
Plot points: 101

Figure 19. The effect of violence on risk magnitude at First Fuzzy Model

Figure 20. The effect of violence on risk magnitude at Second Fuzzy Model



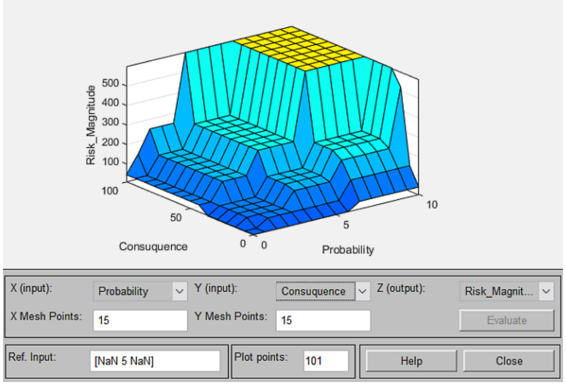
X (input): Probability
Y (input): Frequence
Z (output): Risk_Magnit...
X Mesh Points: 15
Y Mesh Points: 15
Ref. Input: [50 NaN NaN]
Plot points: 101



X (input): Probability
Y (input): Frequency
Z (output): Risk_Magnit...
X Mesh Points: 15
Y Mesh Points: 15
Ref. Input: [50 NaN NaN]
Plot points: 101

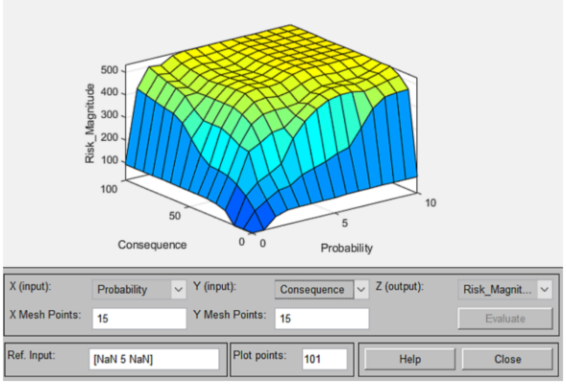
Figure 21. Variation of risk magnitude according to probability and frequency at First Fuzzy Model

Figure 22. Variation of risk magnitude according to probability and frequency at the Second Fuzzy Model



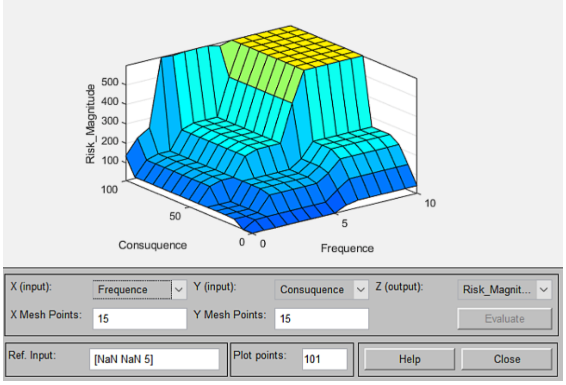
X (input): Probability
Y (input): Consuquence
Z (output): Risk_Magnit...
X Mesh Points: 15
Y Mesh Points: 15
Ref. Input: [NaN 5 NaN]
Plot points: 101

Figure 23. Variation of risk magnitude according to probability and severity at the First Fuzzy Model



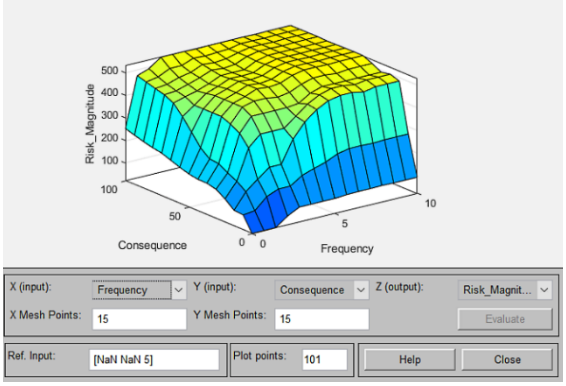
X (input): Probability
Y (input): Consequence
Z (output): Risk_Magnit...
X Mesh Points: 15
Y Mesh Points: 15
Ref. Input: [NaN 5 NaN]
Plot points: 101

Figure 24. Variation of risk magnitude according to probability and severity at the Second Fuzzy Model



X (input): Frequence
Y (input): Consuquence
Z (output): Risk_Magnit...
X Mesh Points: 15
Y Mesh Points: 15
Ref. Input: [NaN NaN 5]
Plot points: 101

Figure 25. Variation of risk magnitude according to frequency and severity at First Fuzzy Model



X (input): Frequency
Y (input): Consequence
Z (output): Risk_Magnit...
X Mesh Points: 15
Y Mesh Points: 15
Ref. Input: [NaN NaN 5]
Plot points: 101

Figure 26. Variation of risk magnitude according to frequency and severity at Second Fuzzy Model

Upon examining the graphs representing the characteristics of the fuzzy models, it is observed that the first fuzzy model (on the left) has sharp transitions, whereas the transitions in the second fuzzy model (on the right) are smoother. This situation is related to the selection of membership functions and threshold values; it has been observed, as seen in Table 6 and Figure 27, resulted in a concentration of risk distribution at the medium level (level 3) and a more balanced distribution of risks across other levels in the second fuzzy model. This distribution is common and realistic; in this respect, the second fuzzy model has produced more successful results than both the Fine-Kinney method and the first fuzzy model.

In natural processes, it is a normal result that the number of 'medium level' situations is higher than the number of 'low level' and 'high level' situations. Therefore, it is an expected result that the risks are concentrated at the medium level in the analysis made with the second fuzzy model. In addition, this result can be considered as a strong sign of the closeness of the analysis made with the second fuzzy model to reality. The reason for reaching this result is that the triangular membership functions used to represent the intermediate levels of the second fuzzy model cover most of the input values of the system as subsets and trigger the triangular membership functions to which these subsets belong, thus affecting the output of the system mostly in favor of the intermediate levels. As a result, the second fuzzy model evaluates most of the risks as medium risk. In contrast, in the first fuzzy model, Pi membership functions were used to represent medium-level risks. The use of Pi Membership Function in fuzzy model design involves some difficulties, namely, Pi membership function is defined with the help of 4 variables, such as a, b, c, and d. These variables determine the boundaries of the right, left, and middle openings. Therefore, it is possible to talk about 3 relationships between the variables. A value assigned to one of the variables will also determine the relationship of this variable with other variables. Although it is possible to determine these relationships, as the number of membership functions used in the system design increases, the operations become longer and more puzzling, and the convenience provided by fuzzy logic loses its importance. However, the Triangle Membership Function can be defined with the help of 3 variables, such as a, b, c. There are only right and left openings here, this shows that it will be sufficient to determine only 2 relationships between the variables. It is easier to determine 2 relations instead of 3; the main reason why the second fuzzy model created using the triangular membership function produces more successful results is this convenience. This situation is because fuzzy logic is closer to human thought and language than classical logic. In the first fuzzy model, although the Pi Membership Functions represent fuzzy sets, they are designed with the design difficulties mentioned above; however, this design has a structure similar to crisp sets as seen in the graphs between Figure 7 and Figure 10; therefore, in the analysis made with the first fuzzy model, relatively similar results were obtained with the classical Fine-Kinney analysis. In the second fuzzy model, the Triangular Membership Functions are easily designed to represent fuzzy sets without the difficulties mentioned above.

Table 6. Comparison of the numerical distribution of risk levels according to the analysis

Risk Level	Number of risks Fine-Kinney	 First fuzzy model	 Second fuzzy model
1	102	74	36
2	87	115	35
3	49	66	171
4	49	38	47
5	43	37	41

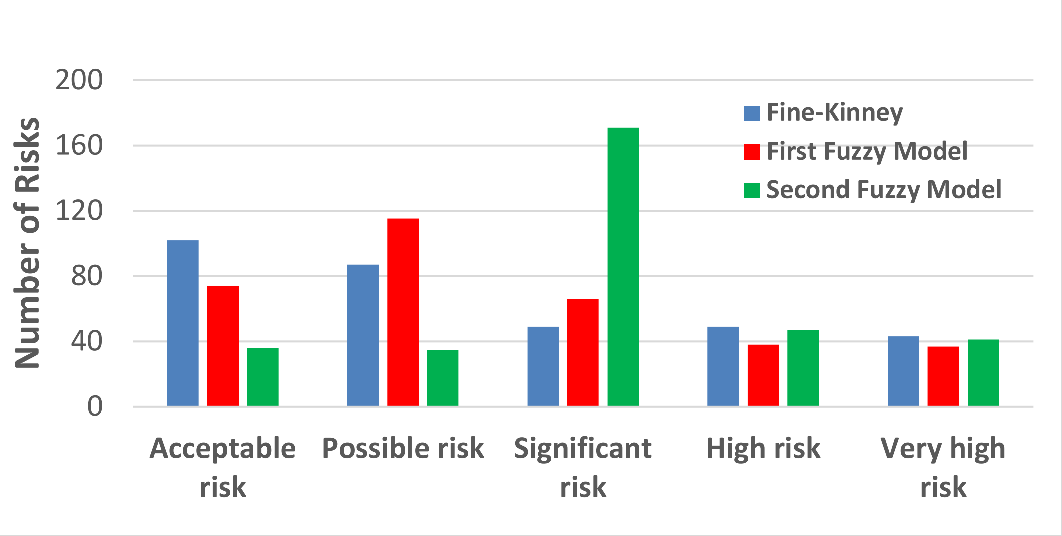


Figure 27. Numerical distribution of risks to levels in analysis

VI. CONCLUSIONS

In this study, an occupational health and safety risk analysis was conducted at a public institution's office engaged in office activities in Istanbul. The analysis found that, similar to construction, mining, shipyard, and other industrial facilities, office workplaces also contain numerous very high-risk (level 5) elements, primarily arising from electrical issues, elevators, boiler rooms, ergonomic incompatibilities, lifting activities, and personnel service vehicles. Level 5 risks require immediate action. Risks due to ergonomic issues can be mitigated over a longer term, but other level 5 risks need urgent measures to be addressed. The study also performed an occupational health and safety risk analysis for office workplaces using the Fine-Kinney method. One of the disadvantages of the Fine-Kinney method, the issue of obtaining quite different (subjective) risk scores depending on the assessor's experience and opinion, was addressed by fuzzifying the method's parameters. Another disadvantage is the possibility of obtaining the same risk score with different combinations of probability, frequency, and severity parameters; this second issue was tackled by creating a rule base that attributed different levels of importance to the parameters. Furthermore, in fuzzy models, the rule base allowed for the reflection of workplace characteristics in the results, considering the relationships between parameters.

It has been shown that fuzzy models based on Fine-Kinney, with a systematically constructed rule base (where the output parameter value increases or at least remains constant in response to increasing values of input parameters) and appropriately selected types and middle-boundary values of membership functions, provide more realistic results than the Fine-Kinney method and can be successfully used in office workplaces. Moreover, the type of membership functions used in fuzzy models and the selection of values assigned to these functions are crucial; this selection should be made based on the scale of the method to be used and the characteristics of the workplace, as evidenced by the better performance of the second fuzzy model compared to the first.

In the literature, there has been no detailed study on the selection of the type of membership functions and the assignment of boundary and middle values for fuzzy risk assessment based on Fine-Kinney. There is a perceived need for further research in this area. Additionally, in subsequent studies in office workplaces, fuzzy logic modeling could be employed to prioritize control measures based on criteria such as feasibility, and benefit/cost analysis.

Authors' Contributions

In this study, authors contributed equally to the study.

Competing Interests

The authors declare that they have no conflict of interest. The authors of this article declare that the materials and methods used in this study do not require ethical committee permission and/or legal-special permission

References

[1] D. Baybora, Aİ. Oral, HN. Gerek, T. Kaplan Senyen, L. Akın, Ö. Ekmekçi, B.Piyal, Occupational Health and Safety. Anadolu University, Eskişehir, Turkey, 2019.

[2] F. Kocabaş, U. Aydın, V. Canbey Özgüler, MN. İlhan, S. Demirkaya, N. Ak, C. Özbaş, Relationship between psychosocial risk factors in the workplace and work related diseaes, occupational disease and work accident. Sosyal Güvence Dergisi. 2018; 7 (14) , 28–62. doi.org/10.21441/sguz.2018.68

[3] ILO world statistic, http://www.ilo.org/moscow/areas-of-work/occupational-safety-and-health/WCMS_249278/lang--en/index.htm (*accessed 03.05.23*).

[4] Eurostat, https://ec.europa.eu/eurostat/web/health/data/database (*accessed 02.04.2023*).

[5] SGK Statistical Year Books, https://www.sgk.gov.tr/Istatistik/Yillik/fcd5e59b-6af9-4d90-a451-ee7500eb1cb4/

[6] B. Arpat, MK. Namal, Occupational Health and Safety. Gazi Kitabevi, Ankara, Turkey, 2020.

[7] Özer MA. Process, performance and risk analysis/management in organizations, Gazi Kitabevi, Ankara, Turkey, 2020.

[8] H. İşsever, M. Güler, MS. Erdoğan, A. Ersoy Yılmaz, T. Ferdi, Occupational health and safety. Istanbul University, Open and Distance Education Faculty, İstanbul, Turkey. 2022.

[9] D. Tengilimoğlu, S. Acar, F. Kahyaoğlu, The importance of ergonomics in design of office furnitures: an empirical study. Ankara Sağlık Hizmetleri Dergisi 7 (2), 23–36. 2008.

[10] Benligiray S. Office management, Anadolu University Open Education Faculty, Eskişehir, Turkey. 2005.

[11] Ş. Yavuz, B. Gür, A. Yavuz, Investigation of the perception of occupational health and safety in employees in manufacturing works. Journal of Social and Humanities Sciences Research 7 (59), 2618–2627. 2020. doi.org/10.26450/jshsr.2030

[12] HF. Ulucan, S. Zeyrek, Occupational health and safety in offices. https://dosyalar.nevsehir.edu.tr/be511600bb819cd8c66a174ca5598500/ofislerde_isg.pdf. 2012.

[13] Özfırat MK, Yetkin ME, Şimşir F, Kahraman B. 2016. Assessment of current hazard sources in longwall production in terms of work safety. Scientific Mining Journal 55 (1), 3–16. 2016.

[14] A. Kokangül, U. Polat, C. Dağsuyu, A new approximation for risk assessment using the AHP and Fine Kinney methodologies. Safety Science 91, 24–32. 2017. doi.org/10.1016/j.ssci.2016.07.015.

[15] M. Gul, E. Celik, E. Akyuz, A hybrid risk-based approach for maritime applications: the case of ballast tank maintenance. Human and Ecological Risk Assessment, 23 (6), 1389–1403.2017. doi.org/10.1080/10807039.2017.1317204.

[16] M. Oturakçı, C. Dağsuyu, Fuzzy Fine - Kinney Approach in Risk Assessment and an Application. Karaelmas Journal of Occupational Health and Safety , 1 (1), 17–25. 2017.

[17] D. Gönen, AD. Karaoglan, MAB.,Ocaktan, A. Oral, H. Atıcı, B. Kaya, A new risk assessment approach for the analysis of musculoskeletal disorders. Gazi University Faculty of Engineering and Architecture Journal, 33(2), 425-440, 2018. doi.org/10.17341/gazimmfd.416351.

[18] M. Gul, AF. Guneri, M. Baskan, An occupational risk assessment approach for construction and operation period of wind turbines. Global Journal of Environmental Science and Management. 4, 281-298 2018. doi:10.22034/gjesm.2018.03.003.

[19] E. Ilbahar, A. Karaşan, S. Cebi, C. Kahraman, A novel approach to risk assessment for occupational health and safety using Pythagorean fuzzy AHP & fuzzy inference system. Safety Science, 103, 124–136. 20108. doi:10.1016/j.ssci.2017.10.025.

[20] W. Wang, X. Liu, Y. Qin, A fuzzy Fine-Kinney-based risk evaluation approach with extended MULTIMOORA method based on Choquet integral. Computers & industrial engineering, 125, 111-123. 2018. doi:10.1016/j.cie.2018.08.019.

[21] II. Taranushina, OV. Popova, Occupational Risk Management at Hazardous Production Facilities Operating Hoisting Mechanisms. Occupational Safety in Industry. 2019. doi:10.24000/0409-2961-2019-11-82-88.

[22] N. Baç, I. Ekmekci, Psychosocial risk assessment by Fine Kinney and ANFIS method: A case study in a metal processing plant, in: Arezes, P.M., Boring, R.L. (Eds.), Advances in Safety Management and Human Performance. Springer International Publishing, Cham, pp. 84–90. 2020.

[23] B. Erdebilli, L. Gür ,Application of risk assessment with fuzzy Fine-Kinney method. Journal of Industrial Engineering, 31819, 75–86. 2020.

[24] BO. Aboubakar, H. Li, YA. Souleymanou, Hydropower projects risk assessment and raking using combined SWARA-TOPSIS and Fine-Kinney methods, in 2021 10th International Conference on Power Science and Engineering (ICPSE). Presented at the 2021 10th International Conference on Power Science and Engineering (ICPSE), IEEE, Istanbul, Turkey, pp. 81–91. 2021.

[25] G. Güney, B.Kahraman, Implementation of the analytic hierarchy process (AHP) and Fine-Kinney method (FKM) against risk factors to determine the total cost of occupational health and safety precautions in environmental research laboratories. International Journal of Occupational Safety and Ergonomics, 28 (4), 2606–2622. 2022. doi: 10.1080/10803548.2021.2010969.

[26] V. Tatar, O. Yazıcıoglu, B. Ayvaz, A novel risk assessment model for work-related musculoskeletal disorders in tea harvesting workers. Journal of Intelligent & Fuzzy Systems, 44 (2), 1-19. 2022. doi:10.3233/jifs-222652

[27] W.T. Fine, Mathematical evaluation for controlling hazards. Journal of Safety Research, 3, 157-166. 1971.

[28] GF. Kinney, AD. Wiruth, Practical risk analysis for safety management (No. NWC TP 5865). Naval Weapons Center, China Lake/California/USA. 1976.

[29] Z. Başkaya, Fuzzy linear programming. Ekin Yayınevi, Bursa, Turkey, 2011.

[30] LA. Zadeh, Fuzzy sets. Information and Control 8, 338–353. 1965.

[31] E. Bozkuş, Ö. Bozkuş, Risk Assessment Approaches Based on Fuzzy Methods in Occupational Health and Safety. Journal of Occupational Health and Safety Academy, 4 (2), 49–64. 2021. doi.org/10.38213/ohsacademy.956021.

[32] E. Kıyak, A. Kahvecioğlu, Fuzzy logic and its application to flight control problem, Aviation and Space Technologies Journal, 1(2), 63–72. 2003.

[33] J. Yen, R. Langari, Fuzzy logic: Intelligence, Control, and Information. Pearson Education, India. 1999.

[34] A. Özdağoğlu Fuzzy operations, defuzzification and verbal thresholds, Detay Yayıncılık, Ankara, Turkey. 2016.

[35] M. Jevscek, 2016. Competencies assessment using fuzzy logic. RUO. Revija za Univerzalno Odlicnost 5, 187.

[36] P. Guttorp, 1990. Fuzzy mathematical models in engineering and management science. Taylor & Francis.

[37] S.K. Meher, B.U. Shankar, A. Ghosh, , 2007. Multispectral remote sensing image classification using wavelet-based features, in: Soft Computing in Image Processing. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 3–34.

[38] L., Pokorádi, Fuzzy logic-based risk assessment. 2002.

Appendix A (Additional Data)

Table 7: Analysis Results

Hazard	Input Parameters			Fine-Kinney		First Fuzzy Model		Second Fuzzy Model		Hazard	Input Parameters			Fine-Kinney		First Fuzzy Model		Second Fuzzy Model	
No	P	F	C	RM	RPN	RM	RPN	RM	RPN	No	P	F	C	RM	RPN	RM	RPN	RM	RPN
H1	3	6	40	720	5	597	5	536	5	H214	0.5	6	15	45	2	45	2	90	3
H68	3	6	40	720	5	597	5	536	5	H251	6	0.5	15	45	2	45	2	147	3
H69	3	6	40	720	5	597	5	536	5	H289	6	0.5	15	45	2	45	2	147	3
H70	3	6	40	720	5	597	5	536	5	H290	6	0.5	15	45	2	45	2	147	3
H74	3	6	40	720	5	597	5	536	5	H298	6	0.5	15	45	2	45	2	147	3
H75	3	6	40	720	5	597	5	536	5	H270	1	6	7	42	2	45	2	90	3
H77	3	6	40	720	5	597	5	536	5	H91	0.5	2	40	40	2	45	2	90	3
H82	3	6	40	720	5	597	5	536	5	H41	6	6	1	36	2	45	2	129	3
H88	6	3	40	720	5	597	5	536	5	H266	6	6	1	36	2	45	2	129	3
H93	3	6	40	720	5	597	5	536	5	H13	0.5	10	7	35	2	45	2	90	3
H94	3	6	40	720	5	597	5	536	5	H26	0.5	10	7	35	2	45	2	90	3
H98	6	3	40	720	5	45	2	93	3	H173	0.5	10	7	35	2	45	2	90	3
H107	3	6	40	720	5	9	1	22	2	H234	10	0.5	7	35	2	45	2	147	3
H117	3	6	40	720	5	597	5	536	5	H159	10	3	1	30	2	45	2	129	3
H130	3	6	40	720	5	597	5	536	5	H190	1	2	15	30	2	45	2	90	3
H228	10	10	7	700	5	597	5	536	5	H194	1	2	15	30	2	45	2	90	3
H232	10	10	7	700	5	597	5	536	5	H195	1	2	15	30	2	45	2	90	3
H233	10	10	7	700	5	597	5	536	5	H196	1	2	15	30	2	45	2	90	3
H96	6	1	100	600	5	597	5	536	5	H197	1	2	15	30	2	45	2	90	3
H120	6	1	100	600	5	597	5	536	5	H203	1	2	15	30	2	45	2	90	3
H121	6	1	100	600	5	597	5	536	5	H217	3	10	1	30	2	45	2	129	3
H161	3	2	100	600	5	597	5	536	5	H225	10	3	1	30	2	45	2	129	3
H325	3	2	100	600	5	597	5	536	5	H246	1	10	3	30	2	135	3	190	3
H326	3	2	100	600	5	597	5	536	5	H250	1	2	15	30	2	45	2	90	3
H329	3	2	100	600	5	597	5	536	5	H150	3	3	3	27	2	45	2	90	3
H59	6	6	15	540	5	597	5	536	5	H151	3	3	3	27	2	45	2	90	3
H236	6	6	15	540	5	597	5	536	5	H34	0.5	0.5	100	25	2	45	2	93	3
H238	6	6	15	540	5	597	5	536	5	H36	0.5	0.5	100	25	2	45	2	93	3
H255	6	6	15	540	5	597	5	536	5	H44	0.5	0.5	100	25	2	45	2	93	3
H256	6	6	15	540	5	597	5	536	5	H45	0.5	0.5	100	25	2	45	2	93	3
H86	0.5	10	100	500	5	597	5	536	5	H46	0.5	0.5	100	25	2	45	2	93	3
H104	0.5	10	100	500	5	135	3	190	3	H50	0.5	0.5	100	25	2	45	2	93	3
H113	6	2	40	480	5	45	2	90	3	H61	0.5	0.5	100	25	2	45	2	93	3
H17	3	10	15	450	5	250	4	300	4	H80	0.5	0.5	100	25	2	45	2	93	3
H43	3	10	15	450	5	250	4	300	4	H83	0.5	0.5	100	25	2	45	2	93	3
H100	3	10	15	450	5	597	5	536	5	H84	0.5	0.5	100	25	2	45	2	93	3
H254	10	3	15	450	5	597	5	536	5	H89	0.5	0.5	100	25	2	45	2	93	3
H230	6	10	7	420	5	250	4	300	4	H116	0.5	0.5	100	25	2	45	2	93	3
H258	6	10	7	420	5	250	4	300	4	H138	0.5	0.5	100	25	2	45	2	93	3
H71	1	10	40	400	5	250	4	300	4	H142	0.5	0.5	100	25	2	45	2	93	3
H73	1	10	40	400	5	250	4	300	4	H52	3	0.5	15	23	2	45	2	93	3
H95	1	10	40	400	5	250	4	300	4	H54	3	0.5	15	23	2	45	2	93	3
H127	10	1	40	400	5	250	4	300	4	H57	3	0.5	15	23	2	45	2	93	3
H76	3	3	40	360	4	597	5	536	5	H66	3	0.5	15	23	2	45	2	93	3
H294	3	3	40	360	4	250	4	300	4	H72	3	0.5	15	23	2	45	2	93	3
H295	3	3	40	360	4	250	4	300	4	H111	3	0.5	15	23	2	45	2	147	3
H296	3	3	40	360	4	250	4	300	4	H212	0.5	3	15	23	2	45	2	90	3
H27	6	0.5	100	300	4	250	4	474	5	H4	0.5	6	7	21	2	45	2	90	3
H31	6	0.5	100	300	4	250	4	474	5	H12	0.5	6	7	21	2	45	2	90	3
H115	3	1	100	300	4	250	4	300	4	H30	6	0.5	7	21	2	45	2	93	3
H125	3	1	100	300	4	250	4	300	4	H67	3	1	7	21	2	45	2	90	3
H126	6	0.5	100	300	4	250	4	474	5	H90	0.5	6	7	21	2	45	2	90	3
H131	3	1	100	300	4	597	5	536	5	H135	0.5	6	7	21	2	45	2	90	3
H132	6	0.5	100	300	4	250	4	474	5	H213	1	3	7	21	2	45	2	90	3
H137	0.5	6	100	300	4	250	4	300	4	H235	3	1	7	21	2	45	2	90	3
H162	6	0.5	100	300	4	597	5	536	5	H267	0.5	6	7	21	2	45	2	90	3
H199	10	2	15	300	4	250	4	300	4	H268	0.5	6	7	21	2	45	2	90	3
H313	0.5	6	100	300	4	250	4	300	4	H275	6	0.5	7	21	2	45	2	93	3
H314	0.5	6	100	300	4	250	4	300	4	H279	6	0.5	7	21	2	45	2	93	3
H315	0.5	6	100	300	4	250	4	300	4	H122	1	0.5	40	20	2	45	2	93	3
H328	3	1	100	300	4	250	4	300	4	H128	0.5	1	40	20	2	45	2	90	3
H20	3	6	15	270	4	250	4	300	4	H141	1	0.5	40	20	2	45	2	93	3
H109	3	6	15	270	4	597	5	536	5	H160	10	2	1	20	2	7	1	72	3
H134	3	6	15	270	4	250	4	300	4	H273	6	3	1	18	1	45	2	91	3
H193	6	3	15	270	4	250	4	300	4	H307	6	3	1	18	1	45	2	91	3
H210	3	6	15	270	4	250	4	300	4	H308	6	3	1	18	1	45	2	91	3
H241	3	6	15	270	4	250	4	300	4	H330	3	2	3	18	1	7	1	18	1
H252	6	3	15	270	4	250	4	300	4	H2	0.5	10	3	15	1	45	2	90	3
H253	6	3	15	270	4	250	4	300	4	H157	0.5	2	15	15	1	7	1	18	1
H287	6	3	15	270	4	250	4	300	4	H170	0.5	2	15	15	1	7	1	18	1
H260	6	6	7	252	4	250	4	300	4	H242	0.5	2	15	15	1	7	1	18	1
H262	6	6	7	252	4	250	4	300	4	H243	0.5	2	15	15	1	7	1	18	1
H311	6	6	7	252	4	250	4	300	4	H249	0.5	2	15	15	1	7	1	18	1
H324	3	2	40	240	4	135	3	190	3	H143	1	2	7	14	1	7	1	18	1
H25	3	10	7	210	4	135	3	190	3	H198	1	2	7	14	1	7	1	18	1
H224	10	3	7	210	4	135	3	190	3	H152	3	0.5	7	11	1	9	1	83	3
H237	3	10	7	210	4	135	3	190	3	H226	0.5	3	7	11	1	7	1	18	1

No	P	F	C	RM	RPN	RM	RPN	RM	RPN	No	P	F	C	RM	RPN	RM	RPN	RM	RPN
H239	3	10	7	210	4	135	3	190	3	H227	0.5	3	7	11	1	7	1	18	1
H282	10	3	7	210	4	135	3	190	3	H278	3	0.5	7	11	1	9	1	83	3
H283	3	10	7	210	4	135	3	190	3	H92	0.5	0.5	40	10	1	45	2	93	3
H14	0.5	10	40	200	4	135	3	190	3	H97	0.2	0.5	100	10	1	597	5	536	5
H19	0.5	10	40	200	4	135	3	190	3	H108	0.5	0.5	40	10	1	45	2	90	3
H23	0.5	10	40	200	4	135	3	190	3	H133	0.5	0.5	40	10	1	45	2	93	3
H78	10	0.5	40	200	4	135	3	238	4	H139	0.5	0.5	40	10	1	45	2	93	3
H79	0.5	10	40	200	4	135	3	190	3	H140	0.5	0.5	40	10	1	45	2	93	3
H85	0.5	10	40	200	4	135	3	190	3	H216	0.5	0.5	40	10	1	45	2	93	3
H136	10	0.5	40	200	4	135	3	238	4	H220	0.5	0.5	40	10	1	45	2	93	3
H297	1	2	100	200	4	250	4	300	4	H221	0.5	0.5	40	10	1	45	2	93	3
H299	0.5	10	40	200	4	135	3	190	3	H222	0.5	0.5	40	10	1	45	2	93	3
H304	10	0.5	40	200	4	135	3	238	4	H223	0.5	0.5	40	10	1	45	2	93	3
H305	10	0.5	40	200	4	135	3	238	4	H261	0.5	0.5	40	10	1	45	2	93	3
H306	10	0.5	40	200	4	135	3	238	4	H49	6	0.5	3	9	1	9	1	22	2
H301	6	2	15	180	3	250	4	300	4	H174	1	3	3	9	1	7	1	18	1
H35	3	0.5	100	150	3	135	3	238	4	H63	1	0.5	15	8	1	45	2	83	3
H37	3	0.5	100	150	3	135	3	238	4	H145	1	0.5	15	8	1	45	2	83	3
H55	3	0.5	100	150	3	135	3	238	4	H181	1	0.5	15	8	1	45	2	83	3
H56	3	0.5	100	150	3	135	3	238	4	H191	1	0.5	15	8	1	45	2	83	3
H112	1	10	15	150	3	597	5	536	5	H192	1	0.5	15	8	1	45	2	83	3
H155	1	10	15	150	3	135	3	190	3	H204	0.5	1	15	8	1	7	1	18	1
H319	0.5	3	100	150	3	135	3	190	3	H205	1	0.5	15	8	1	45	2	83	3
H322	3	0.5	100	150	3	135	3	238	4	H163	0.5	2	7	7	1	7	1	18	1
H323	3	0.5	100	150	3	135	3	238	4	H164	0.5	2	7	7	1	7	1	18	1
H284	3	3	15	135	3	135	3	190	3	H165	0.5	2	7	7	1	7	1	18	1
H285	3	3	15	135	3	135	3	190	3	H166	0.5	2	7	7	1	7	1	18	1
H286	3	3	15	135	3	135	3	190	3	H167	0.5	2	7	7	1	7	1	18	1
H276	6	3	7	126	3	135	3	190	3	H168	0.5	2	7	7	1	7	1	18	1
H302	3	6	7	126	3	135	3	190	3	H184	0.5	2	7	7	1	7	1	18	1
H310	3	6	7	126	3	135	3	190	3	H240	0.5	2	7	7	1	7	1	18	1
H28	3	1	40	120	3	135	3	190	3	H272	0.5	2	7	7	1	7	1	18	1
H42	0.5	6	40	120	3	135	3	190	3	H47	3	0.5	3	5	1	9	1	22	2
H53	3	1	40	120	3	135	3	190	3	H48	3	0.5	3	5	1	9	1	22	2
H81	0.5	6	40	120	3	135	3	190	3	H264	0.5	3	3	5	1	7	1	18	1
H110	3	1	40	120	3	45	2	93	3	H5	0.5	1	7	4	1	7	1	18	1
H124	3	1	40	120	3	135	3	190	3	H11	0.5	0.5	15	4	1	45	2	83	3
H129	0.5	6	40	120	3	135	3	190	3	H33	0.5	0.5	15	4	1	45	2	83	3
H292	6	0.5	40	120	3	135	3	238	4	H169	0.5	0.5	15	4	1	45	2	83	3
H293	6	0.5	40	120	3	135	3	238	4	H175	1	0.5	7	4	1	9	1	22	2
H316	0.5	2	100	100	3	135	3	190	3	H188	0.5	0.5	15	4	1	45	2	83	3
H9	1	6	15	90	3	135	3	190	3	H189	0.5	0.5	15	4	1	45	2	83	3
H10	6	1	15	90	3	135	3	190	3	H206	0.5	0.5	15	4	1	45	2	83	3
H16	3	10	3	90	3	135	3	190	3	H269	1	0.5	7	4	1	9	1	22	2
H18	3	10	3	90	3	135	3	190	3	H277	1	0.5	7	4	1	9	1	22	2
H99	3	10	3	90	3	45	2	93	3	H309	1	0.5	7	4	1	9	1	22	2
H101	10	3	3	90	3	45	2	90	3	H156	1	3	1	3	1	7	1	20	1
H171	3	2	15	90	3	135	3	190	3	H177	0.5	6	1	3	1	7	1	20	1
H215	3	2	15	90	3	135	3	190	3	H263	1	3	1	3	1	7	1	20	1
H218	6	1	15	90	3	135	3	190	3	H317	6	0.5	1	3	1	9	1	22	2
H219	6	1	15	90	3	135	3	190	3	H22	0.5	0.5	7	2	1	9	1	22	2
H257	1	6	15	90	3	135	3	190	3	H123	0.1	0.5	40	2	1	45	2	93	3
H312	1	6	15	90	3	135	3	190	3	H144	0.5	0.5	7	2	1	9	1	22	2
H321	3	2	15	90	3	135	3	190	3	H148	0.5	0.5	7	2	1	9	1	22	2
H327	1	6	15	90	3	135	3	190	3	H149	0.5	0.5	7	2	1	9	1	22	2
H200	6	2	7	84	3	135	3	190	3	H158	0.5	3	1	2	1	7	1	20	1
H281	6	2	7	84	3	135	3	190	3	H172	0.5	3	1	2	1	7	1	20	1
H60	0.5	10	15	75	3	45	2	90	3	H182	0.5	0.5	7	2	1	9	1	22	2
H146	10	0.5	15	75	3	135	3	188	3	H185	1	0.5	3	2	1	9	1	22	2
H147	10	0.5	15	75	3	135	3	188	3	H186	0.5	0.5	7	2	1	9	1	22	2
H179	10	0.5	15	75	3	135	3	188	3	H187	0.5	0.5	7	2	1	9	1	22	2
H180	10	0.5	15	75	3	135	3	188	3	H207	0.5	0.5	7	2	1	9	1	22	2
H3	1	10	7	70	3	135	3	190	3	H208	0.5	1	3	2	1	7	1	18	1
H274	1	10	7	70	3	135	3	190	3	H209	0.5	1	3	2	1	7	1	18	1
H259	3	3	7	63	2	45	2	90	3	H229	0.5	0.5	7	2	1	9	1	22	2
H87	3	0.5	40	60	2	45	2	147	3	H231	0.5	0.5	7	2	1	9	1	22	2
H103	10	2	3	60	2	135	3	190	3	H244	0.5	0.5	7	2	1	9	1	22	2
H105	3	0.5	40	60	2	250	4	300	4	H245	3	0.5	1	2	1	9	1	22	2
H114	0.5	3	40	60	2	135	3	190	3	H248	0.5	0.5	7	2	1	9	1	22	2
H300	3	0.5	40	60	2	45	2	147	3	H265	1	2	1	2	1	7	1	20	1
H303	3	0.5	40	60	2	45	2	147	3	H271	0.5	0.5	7	2	1	9	1	22	2
H51	1	0.5	100	50	2	45	2	147	3	H280	3	0.5	1	2	1	9	1	22	2
H106	0.5	1	100	50	2	135	3	190	3	H288	0.5	0.5	7	2	1	9	1	22	2
H118	1	0.5	100	50	2	45	2	147	3	H6	0.5	0.5	3	1	1	9	1	22	2
H119	0.5	1	100	50	2	45	2	90	3	H7	0.5	0.5	3	1	1	9	1	22	2
H320	0.5	1	100	50	2	45	2	90	3	H24	0.1	10	1	1	1	9	1	22	2
H8	1	3	15	45	2	45	2	90	3	H32	0.1	0.5	15	1	1	9	1	22	2
H15	0.5	6	15	45	2	45	2	90	3	H40	0.5	1	1	1	1	7	1	20	1
H29	6	0.5	15	45	2	45	2	147	3	H102	0.5	0.5	3	1	1	250	4	300	4
H38	6	0.5	15	45	2	45	2	147	3	H153	0.5	2	1	1	1	7	1	20	1
H39	6	0.5	15	45	2	45	2	147	3	H154	0.5	2	1	1	1	7	1	20	1
H58	0.5	6	15	45	2	45	2	90	3	H176	1	0.5	1	1	1	9	1	22	2

No	P	F	C	RM	RPN	RM	RPN	RM	RPN	No	P	F	C	RM	RPN	RM	RPN	RM	RPN
H62	6	0.5	15	45	2	45	2	147	3	H183	0.5	2	1	1	1	7	1	20	1
H64	6	0.5	15	45	2	45	2	147	3	H247	0.5	0.5	3	1	1	9	1	22	2
H178	6	0.5	15	45	2	45	2	147	3	H291	1	0.5	1	1	1	9	1	22	2
H201	6	0.5	15	45	2	45	2	147	3	H318	0.5	2	1	1	1	7	1	20	1
H202	6	0.5	15	45	2	45	2	147	3	H21	0.1	0.5	1	0	1	9	1	22	2
H211	1	3	15	45	2	45	2	90	3	H65	0.5	0.5	1	0	1	9	1	22	2

El Cezeri
Journal of Science and Engineering
ISSN: 2148-3736

Research Article

Audit Opinion Prediction with Data Mining Methods: Evidence from Borsa Istanbul (BIST)

Zafer Kardeş[1a], Tuğrul Kandemir[2b]

[1] Emirdag Vocational School, Afyon Kocatepe University, Afyonkarahisar, 03600, Türkiye
[2]Faculty of Economics and Administrative Sciences, Afyon Kocatepe University, Afyonkarahisar, 03030, Türkiye

zaferkardes@aku.edu.tr

DOI : 10.31202/ecjse.1679862
Received: 19.04.2025 Accepted: 30.06.2025
How to cite this article:
Zafer Kardeş and Tuğrul Kandemir, “Audit Opinion Prediction with Data Mining Methods: Evidence From Borsa Istanbul (BIST)”, El-Cezeri Journal of Science and Engineering, Vol: 12, Iss: 3, (2025), pp.(395-409).
ORCID: 0000-0002-5719-8551, 0000-0002-3544-7422.

Abstract This study compares the performance of 12 different data mining methods in predicting audit opinions for companies’ financial statements. The study dataset consists of 2,093 firm-year observations from 161 companies listed on Borsa Istanbul for 2010-2022. The independent audit opinion types were classified using a set of 28 independent financial and non-financial variables. The following prediction models were used in the study: Bayesian Networks, Naive Bayes, Logistic Regression, Artificial Neural Networks, Radial Basis Function, Support Vector Machines, K-Nearest Neighbor, AdaBoost.M1 Algorithm, Decision Trees (J48), Random Forest, Decision Stump, and Classification and Regression Tree (CART). According to the results of the analysis, the Random Forest model demonstrated the best performance with a prediction accuracy rate of 96.68% for predicting audit opinions. The statistical results of the models were compared based on prediction accuracy, confusion matrix, detailed accuracy results, Type I error rate, Type II error rate, and performance measures. This study pioneers developing models capable of accurately predicting audit opinions using 2,093 firm-year observations based on financial and non-financial variables. This contributes likely to the audit literature by predicting audit opinion types through data mining classification methods. The framework design used in the study is anticipated to serve as a decision support tool for internal and external auditors, accountants, shareholders, company executives, tax authorities, other public institutions, individual and institutional investors, stock exchanges, law firms, financial analysts, credit rating agencies, and the banking system when making decisions.

Keywords: Audit Opinion Prediction, Auditing, Classification, Data Mining, Machine Learning.

1. INTRODUCTION

According to International Auditing Standards, auditors are required to consider a range of factors related to the risk of material misstatements in their clients’ financial statements when designing and implementing audit procedures. Specifically, auditors must “discuss the susceptibility of the entity’s financial statements to material misstatements” and “identify and assess the risks of material misstatement (a) at the financial statement level, and (b) at the level of classes of transactions, account balances, and disclosures” [1]. In this context, auditors must develop and use models that predict the type of audit opinion and provide timely and reliable assessments of the risk of material misstatements, given the data derived from the financial statements, the audited entity, and other available company information. By utilizing such models, auditors can scan their client portfolios and focus on clients who are more likely to receive adverse audit opinions. This approach enables auditors to save time and resources and helps mitigate potential risks.

The fundamental concept of auditing is to verify the information presented in financial statements, as this information serves as the basis for decision-making by various groups such as shareholders, potential investors, intermediaries, managers, financial advisors, financial analysts, creditors, and government authorities. From these users’ perspective, an audit is effective when auditors confirm that financial statements are free from errors, omissions, fraud, and manipulation. The development and application of new technologies across various disciplines also impact the accounting and auditing fields. One way to enhance the effectiveness of audits is to use audit opinion prediction methods and employ a model tailored to audit conditions and the environment [2].

Data mining is an emerging method that is gaining increasing attention because of its ability to enhance the effectiveness of financial statement audits. Audit opinion models developed by researchers using data mining analytics enable auditors to utilize large datasets during the audit process. Traditional data analysis methods assume that the data are already structured, and that the reliability of the data analysis process has been verified. Therefore, these methods cannot handle large datasets obtained from diverse sources. However, data mining makes this possible in the audit process by automatically extracting patterns from the data and testing hypotheses [3].

Data mining has become widely applied in various business practices in the fields of accounting and finance. Recently, [4] highlighted that 82% of data mining applications in accounting are focused on predictive studies. Additionally, they noted that the most commonly used data mining techniques in accounting are artificial neural networks, regression, decision trees, and support vector machines.

The primary aim of this study is to explore the capabilities of data mining by comparing the predictive performance of 12 different data mining classification methods for predicting audit opinions. In this context, the study problem, that is, the main study question to be answered, is as follows:

- Which data mining method will demonstrate the best performance in predicting independent audit opinions in Turkey?

In this study, 12 different data mining methods were analyzed, and models were developed using WEKA software based on data related to the type of independent audit opinion. The performance of the models was presented based on the prediction accuracy, confusion matrix, detailed accuracy metrics (TP Rate, FP Rate, Precision, Recall, F-Measure, and ROC Area), Type I and Type II error rates, and performance metrics (Kappa statistic, MAE, RMSE, RAE, and RRSE).

The study sample included 161 companies from 19 sectors. The dataset comprises 2,093 firm-year observations for the period 2010 and 2022. In the models, 24 financial variables and four non-financial variables were used. The results indicate that the independent variables have high explanatory power in explaining the type of audit opinion.

Numerous studies [5], [6], [7] have examined the prediction of audit opinion types. The varying findings of these studies have motivated subsequent study. Most of these studies have been conducted in developed markets, with relatively few analyzing the prediction of audit opinions in emerging markets. This study, which explores data mining capabilities by comparing 12 different classification methods using a Turkey sample, is expected to contribute to the relevant literature. The findings of this study may be valuable for auditors and researchers, and could be a decision-support tool for predicting audit opinions.

Section 2 reviews previous studies on audit opinion prediction and data mining. Section 3 outlines the study's methodology. The findings of this study are discussed in Section 4. Finally, Section 5 present the main conclusions, limitations, and recommendations for future study.

2. LITERATURE REVIEW

In recent years, data mining methods have gained significant attention in the international literature, particularly regarding decision-making processes within various accounting, management, finance, and auditing applications to improve audit opinion prediction and decision support. For instance, auditors have employed data mining techniques to predict audit opinions [5], [6], [7], [8]. Researchers have also used data mining to identify factors influencing audit opinions [9], [10], [11], predict fraudulent financial reporting [12], [13], [14], predict management fraud [15], predict bankruptcy [16], [17], detect financial statement restatements [18], assist in auditor selection [19], evaluate risk [20], apply analytical review procedures [21], and perform continuous auditing [22].

Data mining methods are also gaining attention in Turkey, especially in decision-making processes within various accounting, finance, and auditing applications. For example, data mining methods have been applied to predict audit opinions [23], [24], [25]. Additionally, studies conducted with Turkey samples have used data mining techniques to identify factors influencing audit opinions [26], [27], [28], [29], predict fraudulent financial reporting [30], [31], [32], [33], [34], [35], [36], detect manipulations in financial statements [37], assess audit quality [38], and detect fraud in vehicle insurance claims [39]. Studies that use data mining methods for audit opinion prediction are summarized chronologically in Table 1.

Table 1 presents a comparative overview of the studies based on the periods they covered, the countries in which they were conducted, the size of the datasets used, the independent variables employed, and the data mining methods applied.

Early studies primarily relied on classical statistical modelling techniques, such as probit and logistic regression models [9]. However, more recent studies have shown an increasing preference for modern machine learning techniques, such as decision trees and artificial neural networks [5], [6], [7], [23]. Artificial neural networks appear to be the most frequently used data mining methods in the literature. These are followed by logistic regression, decision trees, discriminant analysis, support vector machines, K-nearest neighbor, naive Bayes, and probit regression. Some studies [9], [40], [41], [43], [44], [49] have focused on a single method, whereas others have conducted comparative analyzes of multiple methods. This study presents a comparative analysis of 12 data mining methods.

Studies present varying levels of accuracy depending on the methodology used, sample characteristics, and country of application. The primary motivation for conducting this study is to compare the performance of data mining methods in predicting audit opinions within a Turkey sample using financial and non-financial ratios that are crucial for the audit process.

Table 1. Summary of Studies on Audit Opinion Prediction

Author(s)	Data Set	Method(s) Used	Findings
[9]	Period: 1969-1980 275 modified and 441 unmodified opinions Country: ABD	Probit Regression	A probit regression model is developed using financial and market variables to predict qualified audit opinions.
[40]	Period: 1992-1994 eight modified and 103 unmodified opinions Country: Finland	Kruskal-Wallis Test and Logistic Regression	The study's results suggest that an effective model can explain the qualifications in the audit reports of publicly traded Finnish companies. The model achieved a 62% prediction accuracy.
[41]	Period: 1997-1999 50 modified and 50 unmodified opinions Country: Greece	Ordinary Least Squares (OLS) and Logistic Regression	The developed model achieved a 78% prediction accuracy.
[42]	Period: 1997-1999 50 modified and 50 unmodified opinions Country: Greece	UTADIS, Discriminant Analysis, and Logistic Regression	The developed UTADIS model achieved a 78.83% prediction accuracy.
[43]	Period: 1998-2003 859 modified and 5.189 unmodified opinions Countries: England and Ireland	Support Vector Machines	The results demonstrate a correct classification performance for 73% of the companies in the sample, with "Credit Rating" highlighted as the most crucial variable.
[44]	Period: 2001 162 modified and 23 unmodified opinions Country: Greece	Ordinary Least Squares (OLS) and Logistic Regression	The developed model achieved a 90% prediction accuracy.
[45]	Period: 1997-2004 264 modified and 3.069 unmodified opinions Country: England	Artificial Neural Network (ANN), Probabilistic Neural Network (PNN), and Logistic Regression (LR)	The probabilistic neural network outperformed traditional artificial neural networks and logistic regression with an 84% classification accuracy.
[46]	Period: 1998-2003 980 modified and 4.296 unmodified opinions Country: England	K-Nearest Neighbor (K-NN), Discriminant Analysis, and Logistic Regression	The K-NN model's overall classification accuracy (66-72.2%) was higher than that of discriminant analysis (60.4-62.4%) and logistic regression (61.1-65.9%).
[8]	Period: 1995-2004 225 modified and 225 unmodified opinions Countries: England and Ireland	Artificial Neural Network, C4.5 Decision Trees, and Bayesian Belief Network	The highest classification performance was shown by Bayesian belief networks at 82.22% followed by the artificial neural network model at 81.11% and the decision tree model at 77.69%.
[6]	Period: 2001-2007 708 modified and 310 unmodified opinions Country: Iran	Support Vector Machines and Decision Trees	A new method combining support vector machines and decision trees showed better performance.
[5]	Period: 2001-2007 347 modified and 671 unmodified opinions Country: Iran	Probabilistic Neural Network (PNN), Multilayer Perceptron (MLP), Radial Basis Function (RBF), and Logistic Regression (LR)	The results show that MLP achieved 88% accuracy in predicting audit opinions and demonstrated the best performance among the methods considered.
[23]	Period: 2010-2013 55 modified and 55 unmodified opinions Country: Turkey	C5.0 Decision Trees, Discriminant Analysis, and Logistic Regression	The classification results of the models revealed that the C5.0 decision tree algorithm had the highest classification accuracy at 98.2% outperforming discriminant and logistic regression models in explaining audit opinions.
[24]	Period: 2011-2014 58 modified and 58 unmodified opinions Country: Turkey	GRI Algorithm, C5.0, and CART Decision Trees	The C5.0 decision tree model showed the highest performance in predicting audit opinion types with an overall correct classification rate of 88.8%.
[7]	Period: 2008-2010 78 modified and 369 unmodified opinions Country: Spain	Probabilistic Neural Network (PNN) and Multilayer Perceptron (MLP)	The study showed accuracy performances of 98.10% and 81.70% for the probabilistic neural network and the multilayer perceptron, respectively, on the test sample.

Author(s)	Data Set	Method(s) Used	Findings
[47]	Period: 2010-2013 4.421 modified and 9.140 unmodified opinions Country: Serbia	C5.0, Random Forest, Regularized Random Forest, Stochastic Gradient Boosting, Extreme Gradient Boosting, K-Nearest Neighbor, Multilayer Perceptron, Support Vector Machine, Linear Discriminant Analysis, Logistic Regression, Probit Regression, and Mixed-Effects Logistic Regression	The results show that several methods from both fields achieved comparable prediction performance at approximately 89% in the first scenario. However, in the second scenario, machine learning algorithms, particularly tree-based ones like Random Forest, performed significantly better reaching 79%.
[25]	Period: 2009-2019 868 modified and 6.259 unmodified opinions Country: Turkey	K-Nearest Neighbor, Decision Trees, Logistic Regression, Discriminant Analysis, Multilayer Perceptron, Naive Bayes, Random Gradient, Random Forest, AdaBoost, and XGBoost	The Random Forest (94.7%) and XGBoost (95.0%) algorithms provided the best results in the study. It is stated that information such as financial ratios, previous period opinion type, previous period independent audit firm class, and audit report delays effectively predict future period results.
[3]	Period: Five annual data 7.199 modified and 23.000 unmodified opinions Countries: England and Ireland	Artificial Neural Network, Support Vector Machines, and K-Nearest Neighbor	The empirical results reveal that the support vector machine and artificial neural network models achieved a higher average accuracy rate (95.9%) and outperformed the K-Nearest Neighbor method (93.8%).
[48]	Period: 2018-2019 9.363 modified and 16.644 unmodified opinions Country: England and Ireland	Linear Discriminant Analysis, Logistic Regression, Naive Bayes, and Decision Trees	The results show that the decision tree model achieved a correct classification performance of 96.3%. The decision tree model outperformed the logistic regression, linear discriminant, and naive Bayes network models on the 2018 and 2019 datasets.
[49]	Period: 2012-2016 251 modified and 229 unmodified opinions Country: Iran	Probit Model	The model shows a 72.9% accuracy in correctly classifying the overall sample to predict the type of auditor's opinion.
[50]	Period: 2018 year 25 modified and 898 unmodified opinions Country: China	BP Neural Network and Adam Optimizer Algorithm	The results illustrate that the Adam optimizer model achieved the highest prediction accuracy at 94.05%.
[51]	Period: 2001-2017 25.943 unqualified, 10.217 unqualified with explanatory language and 1.165 going-concern opinions Country: ABD	Support Vector Machines, Decision Trees, K-Nearest Neighbor (K-NN), and Rough Clusters	This study compares the performance of four data mining techniques in predicting audit opinions on companies' financial statements. Support Vector Machines demonstrated the highest overall prediction accuracy (75.6%), Type I, and Type II error rates.

3. METHODOLOGY

4.1. Study Purpose

This study aims to compare the prediction performance of data mining classification methods using these methods to predict audit opinions.

4.2. Data Set

The statistical population comprises all the companies listed on the BIST between 2009 and 2022. The sample for this study included companies that met the following criteria:

- Companies continuously traded on BIST from 2009 to 2022 were included. Independent audit reports must be available for year t-1 and year t, and all companies must have their financial statements available for year t. Company data has been accessible on the KAP (Public Disclosure Platform) website since 2009. The independent variables used in this study include the opinion of the previous year's audit. Therefore, the 2009 audit report data were used for the 2010 data. Based on this, the starting year for the dataset was set to 2010.
- Banks, financial institutions, investment firms, financial intermediaries, and holding companies were excluded from the sample because of their distinct reporting structures.

After these restrictions were applied, the annual sample size was 161 companies. Table 2 presents the selection of companies included in the study sample.

The selection of sample companies and applied restrictions is presented in Table 2. During the data collection process, the total number of companies operating on BIST was 651. A total of 250 companies whose financial statements and independent auditor reports were unavailable between 2009 and 2022 were excluded from the sample. Additionally, 240 companies comprising financial institutions, brokerage firms, banks, holding companies, and investment firms were excluded because of their distinct reporting structures.

Table 2. Selection of Companies in the Study Sample

Explanation	Number of Companies
Total Number of Companies Listed on BIST	651
Exclusions:	
1. Companies without Financial Statements and Independent Auditor Reports for 2009-2022	(250)
2. Financial Institutions, Brokerage Firms, Banks, Holdings, and Investment Firms	(240)
Final Sample Size	161

Companies listed under these restrictions were excluded from the sample to ensure more consistent and meaningful results. This approach helps prevent study findings from being influenced by the structural differences between companies. The final sample comprised of 161 companies.

The dataset used in the study consists of 13 years of data (2010-2022) from 161 companies comprising 2,093 firm-year observations. Table 3 presents the distribution of audit opinion types for the 2,093 observations.

Table 3. Distribution of Dataset by Audit Opinion Types

Total Observations	Unmodified Opinion Observations (n)	Modified Opinion Observations (n)
2.093	1.855	238

Table 3 illustrates the distribution of the dataset by audit opinion types. Modified audit opinions include qualified, disclaimer, and adverse opinion types. Table 4 shows the breakdown of modified audit opinions which accounted for 11.37% (238 observations) of the dataset.

Table 4. Distribution of Modified Audit Opinions

Qualified Opinion (n)	Disclaimer of Opinion (n)	Adverse Opinion (n)
223	15	-

When examining the distribution of modified audit opinions, 93.70% (223 observations) are qualified opinions, while 6.30% (15 observations) are disclaimers. There were no audit reports with adverse opinions from the observations included in the study.

4.3. Dependent and Independent Variables

This study utilized both financial and non-financial variables to predict audit opinions. The independent variables consist of a new set of variables inspired by those commonly used in the literature and are thought to influence audit opinions. The Stockeys Pro database was used to construct a dataset for financial variables. The variable set includes 28 independent financial and non-financial variables derived from each company's annual financial statements and audit reports, along with a dependent variable representing modified and unmodified audit opinions. Information on the dependent variable, the four non-financial independent variables, and the 24 financial independent variables used in the study are presented in Table 5.

Table 5 presents the variables used in this study. The set of independent variables consists of both financial and non-financial variables. Financial variables are divided into six categories: (1) liquidity ratios, (2) profitability ratios, (3) valuation ratios, (4) growth ratios, (5) financial structure ratios, and (6) activity ratios. The non-financial variables include auditor change, auditor size, previous year's audit opinion, and audit report delay. Among the independent variables used in the study, three variables - auditor change, auditor size, and the previous year's audit opinion (X1-X3) - are nominal and binary (1/0). The remaining 25 independent variables (X4-X28) are numeric.

4.4. Data Preprocessing Process

The dataset prepared for this study contained 38 attributes and 2,094 rows containing 79,572 data cells to predict audit opinions. The final dataset included 29 attributes consisting of one dependent variable and 28 independent variables. Among the financial variables, 446 instances of missing data were missing from the company-year observations. Missing values in the dataset were replaced with the mean values of the respective attributes.

The models used in this study were analyzed using WEKA software. After loading the dataset into WEKA, the "Interquartile Range" filter is applied using a filtering method. The "Interquartile Range" filter detects outliers and extreme values based on interquartile ranges in the observations. Interquartile Range, a filter for detecting outliers and extreme values based on the interquartile range. The filter skips the class attribute.

Outliers: Q3 + OF*IQR < x <= Q3 + EVF*IQR or Q1 - EVF*IQR <= x < Q1 - OF*IQR

Extreme values: x > Q3 + EVF*IQR or x < Q1 - EVF*IQR

Key:

Q1 = 25% quartile

Q3 = 75% quartile
IQR = Interquartile Range, difference between Q1 and Q3
OF = Outlier Factor
EVF = Extreme Value Factor

Following the application of this filter, two nominal variables, “Outlier” and “Extreme Value”, were added to the dataset. Table 6 provides the information on these variables.

Table 5. Variables Used in the Study

Variable Group	Variables
Dependent Variable:	
Y	Audit Opinion Type (1: Unmodified audit opinion; 0: Modified audit opinion)
Non-Financial Independent Variables:	
X_1	Did the Auditor Change? (1: Auditor has not changed 0: Auditor has changed)
X_2	Auditor Size (1: Big Four Audit Firms; 0: Big Four Non-audit Firms)
X_3	Prior Year Audit Opinion (1: Unmodified audit opinion; 0: Modified audit opinion)
X_4	Audit Report Delay (Number of days between the end of the fiscal year and the audit report announcement date)
Financial Independent Variables:	
X_5 Liquidity Ratios	Current Ratio
X_6 Liquidity Ratios	Liquid Ratio
X_7 Liquidity Ratios	Cash Rate
X_8 Profitability Ratios	Return on Assets (ROA) (%)
X_9 Profitability Ratios	Gross Margin (%)
X_{10} Profitability Ratios	Operating Margin (%)
X_{11} Profitability Ratios	EBITDA Margin (%)
X_{12} Profitability Ratios	Net Profit Margin (%)
X_{13} Profitability Ratios	Return on Equity (ROE) (%)
X_{14} Profitability Ratios	EBIT Margin (%)
X_{15} Valuation Rates	Market Value/Book Value
X_{16} Growth Rates	Active Growth (%)
X_{17} Growth Rates	Net Sales Growth (%)
X_{18} Growth Rates	Equity Growth (%)
X_{19} Financial Structure	Debt to Capital Ratio (%)
X_{20} Financial Structure	Fixed Assets / Assets (%)
X_{21} Financial Structure	Short Term Debt / Assets (%)
X_{22} Financial Structure	Short Term Debt / Total Debt (%)
X_{23} Financial Structure	Equity / Assets (%)
X_{24} Financial Structure	Debt to Equity Ratio (%)
X_{25} Activity Ratios	Asset Turnover Ratio
X_{26} Activity Ratios	Receivables Turnover Ratio
X_{27} Activity Ratios	Stock Turnover Ratio
X_{28} Activity Ratios	Trade Payables Turnover Ratio

Table 6. Information on Outliers and Extreme Values

Variable Name	Label	Number	Percent (%)
Outliers	No	1.487	71.05
Outliers	Yes	606	28.95
Extreme Values	No	1.625	77.64
Extreme Values	Yes	468	22.36

In the dataset, if an observation is classified as an outliers or extreme value, it is labelled with “Yes.” If it is not, it is labelled with “No.” In Table 6, regarding the outliers variable, 71.05% (1,487 observations) of the 2,093 observations in the dataset are labelled as “No,” while 28.95% (606 observations) are labelled as “Yes.” For the extreme values variable, 77.64% (1,625 observations) are labelled as “No,” and 22.36% (468 observations) are labelled as “Yes.” When considering the outliers and extreme values variables in general, it was observed that there were not too many outliers or extreme values.

Table 7. Dataset Information after Removing Outliers and Extreme Values

Variable Name	Label	Number	Percent (%)
Outliers	No	1.286	100.00
Outliers	Yes	-	-
Extreme Values	No	1.286	-
Extreme Values	Yes	-	-
Audit Opinion Type	1	1.189	92.46
Audit Opinion Type	0	97	7.54

In total, 606 observations are labelled “Yes” for the outliers variable, and 468 observations are labelled “Yes” for the extreme values variable. These values were excluded from the dataset to ensure meaningful results. The “Remove With Values” filter removes the outliers and extreme values from the dataset. After this process, information regarding the dataset, including outliers, extreme values, and dependent variables, is presented in Table 7.

After removing outliers and extreme value, no observations are labelled “Yes.” That is, none of the 1,286 remaining observations contained outliers or extreme values. Following the removal of these values, the dependent variable consisted of 92.46% (1,189 observations) unmodified audit opinions and 7.54% (97 observations) modified audit opinions.

To address the issue of dataset imbalance, we apply the SMOTE (Synthetic Minority Over-sampling Technique) method increasing the number of non-unqualified audit opinion observations from 97 to 1,188. After the SMOTE application, the distribution of the dependent variable became 1,189 (50.02%) unmodified audit opinions and 1,188 (49.98%) modified audit opinions. We resolved the data imbalance issue by applying the SMOTE method, achieving a more even distribution between the two categories.

4.5. Data Mining Classification Methods Used in the Study

We utilized 12 different data mining classification methods to predict independent audit opinions. Table 8 lists the methods used in this study.

Table 8. Data Mining Classification Methods Used in the Study

Serial No.	Method
1	Bayesian Belief Network (BBN)
2	Naive Bayes (NB)
3	Logistic Regression (LR)
4	Artificial Neural Networks (ANN)
5	Radial Basis Function (RBF)
6	Support Vector Machines (SVM)
7	K-Nearest Neighbor (K-NN)
8	AdaBoost.M1 Algorithm
9	Decision Trees J48 (DT)
10	Random Forests (RF)
11	Decision Stump (DS)
12	Classification and Regression Tree (CART)

The 12 different data mining classification methods listed in Table 8 were used to establish the models and perform the analyses. To ensure that the results obtained are reliable and generalizable, 12 different methods were preferred instead of just a few methods. The methods considered were diversified and included probabilistic, statistical, rule-based, machine learning-based, and ensemble methods. This made it possible to objectively compare which model would perform better. As the methods analyzed are common data mining and machine learning algorithms, this study also offered the possibility of comparison with other studies in the literature.

The data mining classification methods used in the study were analyzed using WEKA software. The default parameter settings of the algorithms were used in all analyses. This made the model comparisons repeatable and more objective. The default parameters of the relevant algorithms are accessible via the WEKA written interface.

4. FINDINGS

This section presents the results of the models that were established to predict independent audit opinions. First, we provided descriptive statistics for the variables, followed by the model results displayed in separate tables based on the classification matrix and detailed accuracy metrics.

4.1. Descriptive Statistics on Variables

Table 9 presents the descriptive statistics.

Table 9. Descriptive Statistics for Study Variables

Panel A. Descriptive Statistics for Nominal Data

Variable Name	Label	Explanation	Number	Percent (%)
Opinion Type	1	Unmodified Opinion Type	1.189	50.02
	0	Modified Opinion Type	1.188	49.98
Did the Auditor Change?	1	Auditor Unchanged	1.930	81.19
	0	Auditor Changed	447	18.81
Auditor Size	1	Big Four Audit Firms	1.376	57.89
	0	Big Four Non-audit Firms	1.001	42.11
Prior Year Audit Opinion	1	Unmodified Opinion Type	1.343	56.50
	0	Modified Opinion Type	1.034	43.50

Panel B. Descriptive Statistics for Numerical Data

Variable Name	Minimum	Maximum	Mean	StdDev
Audit Report Delay	25	104	65.819	13.651
Current Ratio	0.15	4.93	1.58	0.777
Liquid Ratio	0.04	3.54	0.959	0.573
Return on Assets (ROA) (%)	0	1.671	0.229	0.283
Gross Margin (%)	-0.27	0.407	0.033	0.077
Operating Margin (%)	-0.128	0.575	0.194	0.119
EBITDA Margin (%)	-0.255	0.412	0.055	0.097
Net Profit Margin (%)	-0.203	0.49	0.097	0.094
Return on Equity (ROE) (%)	-0.345	0.42	0.026	0.113
EBIT Margin (%)	-0.622	0.836	0.073	0.171
Market Value/Book Value	-0.261	0.732	0.135	0.111
Active Growth (%)	0.17	11.08	1.942	1.503
Net Sales Growth (%)	-0.507	1.304	0.18	0.222
Equity Growth (%)	-0.835	1.455	0.193	0.285
Debt to Capital Ratio (%)	-0.833	1.552	0.169	0.275
Fixed Assets / Assets (%)	0.052	1.733	0.553	0.239
Short Term Debt / Assets (%)	0.002	0.979	0.477	0.194
Short Term Debt / Total Debt (%)	0.042	1.257	0.394	0.188
Equity / Assets (%)	0.121	0.997	0.717	0.167
Debt to Equity Ratio (%)	-0.736	0.948	0.434	0.236
Asset Turnover Ratio	-5.371	8.248	1.37	1.697
Receivables Turnover Ratio	0.03	3.44	0.9	0.501
Stock Turnover Ratio	0.36	21.05	6.041	3.652
Trade Payables Turnover Ratio	0.19	36.08	6.327	6.806

Panel C. Descriptive Statistics for Outliers and Extreme Value

Variable Name	Label	Data Type	Number	Percent (%)
Outlier	No	Nominal	2.377	100.00
	Yes	Nominal	-	-
Extreme Value	No	Nominal	2.377	100.00
	Yes	Nominal	-	-

Descriptive statistics for the study variables are presented in three panels: nominal data in Panel A, numerical data (including minimum, maximum, mean, and standard deviation values) in Panel B, and outliers and extreme values statistics in Panel C. Panel A, which contains descriptive statistics for nominal data, shows the distributions of variables, such as audit opinion, auditor change, auditor size, and the previous year's audit opinion. Panel B, which presents the descriptive statistics for the numerical data, displays the numerical variables' minimum, maximum, mean, and standard deviation values. Panel C provides descriptive statistics for outliers and extreme values, showing the distributions of these variables.

A regression model was established to determine which of the variables had a significant influence on the dependent variable. The result of the regression model showed that the variables trade payables turnover, stock turnover, prior year audit opinion, liquid ratio, ebitda margin, debt to capital ratio, receivables turnover, return on assets, and asset turnover were more important than the other variables. The R-squared value, which was calculated to test the significance of the model created, was found to be 0.9058. This value is quite high and shows that the role of the variables in explaining the dependent variable is strong.

4.2. Comparison of Data Mining Classification Methods Based on Prediction Performances

This study analyzes data related to the independent audit opinion type using 12 different data mining methods in the WEKA software, and models were developed. This section presents a comparison of the statistical results of the models based on prediction accuracy, confusion matrix, detailed accuracy metrics (TP Rate, FP Rate, Precision, Recall, F-Measure, and ROC Area), Type I and Type II error rates, and performance metrics (Kappa statistic, MAE, RMSE, RAE, and RRSE).

The classification models used in this study were divided into five categories: Bayes, Functions, Lazy, Meta, and Trees. Bayesian Networks and Naive Bayes were used for the Bayes category. For the functions category, Logistic Regression, Artificial Neural Networks, Radial Basis Function, and Support Vector Machines were applied. K-Nearest Neighbor was employed in the Lazy category, while AdaBoost.M1 was used for the Meta category. For the Trees category, the J48 Decision Tree, Random Forest, Classification and Regression Tree (CART), and Decision Stump were used. Each of these classifiers has its own advantages and disadvantages. This study aims to determine the method that performs better in the context of audit opinion prediction. The prediction accuracy, performance, and confusion matrices of the models are listed in Table 10.

Table 10 presents the prediction accuracy rates and confusion matrices of the data mining models. The BBN, NB, LR, ANN, RBF, SVM, K-NN, AdaBoost.M1, DT, RF, DS, and CART models show overall prediction performances of 90.66%, 85.36%, 91.21%, 96.05%, 91.88%, 91.33%, 96.42%, 91.17%, 95.62%, 96.68%, 91.33%, and 94.11%, respectively. As seen in Table 10, the best performance in building the prediction model was achieved by Random Forest (RF), with an accuracy rate of 96.68%. K-Nearest Neighbor (K-NN) was the second-best model, with an accuracy of 96.42%. Naive Bayes (NB) shows the lowest performance among the models, with an accuracy rate of 85.36%. This result suggests that the NB model is weaker than the other models in predicting the correct audit opinion type. When comparing the models in terms of prediction accuracy, it is clear that the RF model outperformed the others. In summary, the RF model is more effective in predicting audit opinion types than the other models. Some studies in the literature [25], [47] also indicate that the Random Forest model performs better than other models.

When examining the confusion matrices in Table 10, the support vector machines (SVM) and decision stump (DS) models show the highest performance in predicting unmodified opinions, with an accuracy of 97.81%. The model with the lowest performance in predicting unmodified opinions was naive Bayes (NB), with 84.44%. For predicting modified opinions, the K-nearest neighbor (K-NN) demonstrated the highest performance at 97.31%, followed by the Random Forest (RF) model at 96.55%. The models with the lowest performance in predicting modified opinions were support vector machines (SVM) and decision stump (DS), at 84.85%.

Table 10. Data Mining Classification Methods Used in the Study

Model	Prediction Accuracy (%)	Confusion Matrix		
BBN	90.66	a	b	
		1073	116	a=1
		106	1082	b=0
NB	85.36	a	b	
		1004	185	a=1
		163	1025	b=0
LR	91.21	a	b	
		1133	56	a=1
		153	1035	b=0
ANN	96.05	a	b	
		1137	52	a=1
		42	1146	b=0
RBF	91.88	a	b	
		1140	49	a=1
		144	1044	b=0
SVM	91.33	a	b	
		1163	26	a=1
		180	1008	b=0
K-NN	96.42	a	b	
		1136	53	a=1
		32	1156	b=0
AdaBoost.M1	91.17	a	b	
		1154	35	a=1
		175	1013	b=0
DT J48	95.62	a	b	
		1133	56	a=1
		48	1140	b=0
RF	**96.68**	a	b	
		1151	38	a=1
		41	1147	b=0
DS	91.33	a	b	
		1163	26	a=1
		180	1008	b=0
CART	94.11	a	b	
		1119	70	a=1
		70	1118	b=0

Figure 1 shows the prediction accuracy rates of the models arranged from the highest to the lowest.

When examining the prediction accuracies of the models, it is evident that they are effective in accurately predicting audit opinion type. Therefore, the results suggest that the development of an accurate classification model for audit opinion prediction in Turkey could significantly contribute to the field.

Detailed accuracy results based on the weighted averages of the models are presented in Table 11. These results include the weighted averages for the TP Rate, FP Rate, precision, recall, F-measure, and ROC area for each model.

In Table 11, the detailed accuracy results based on the weighted averages of all models are compared. The TP rate represents the ratio of correctly predicted positive instances. The TP rate aligns with the prediction accuracy results shown in Table 10; the higher the TP rate, the better the model performance. The TP rates for the models ranged between 85.4% and 96.7%, indicating that the models can accurately classify audit opinions. The FP rate represents the ratio of incorrectly predicted negative instances. The lower the FP rate, the better the model performance. FP rates for the models ranged from 14.6% to 3.3%. The naive Bayes (NB) model had the highest FP rate at 14.6%, whereas the Random Forest (RF) model had the lowest FP rate at 3.3%. The precision rates of the models, similar to the TP rates, range from 85.4% to 96.7%. Because recall measures the percentage of the entire dataset, the recall results are parallel to the TP rates of the models.

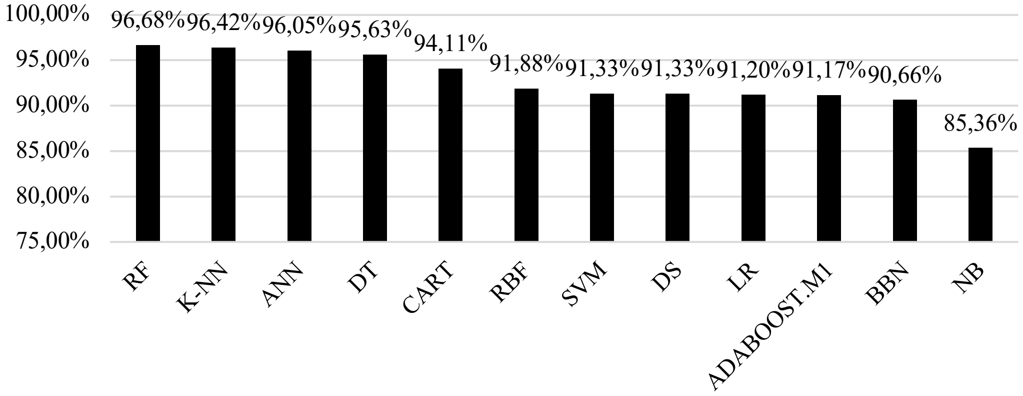


Figure1. Prediction accuracy graph of models

Table 11. Detailed Accuracy Results Based on Weighted Averages of Models

Model	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
BBN	0.907	0.093	0.907	0.907	0.907	0.961
NB	0.854	0.146	0.854	0.854	0.854	0.919
LR	0.912	0.088	0.915	0.912	0.912	0.967
ANN	0.960	0.040	0.960	0.960	0.960	0.984
RBF	0.919	0.081	0.921	0.919	0.919	0.967
SVM	0.913	0.087	0.920	0.913	0.913	0.913
K-NN	0.964	0.036	0.964	0.964	0.964	0.965
AdaBoost.M1	0.912	0.088	0.917	0.912	0.911	0.964
DT J48	0.956	0.044	0.956	0.956	0.956	0.951
RF	**0.967**	**0.033**	**0.967**	**0.967**	**0.967**	**0.992**
DS	0.913	0.087	0.920	0.913	0.913	0.903
CART	0.941	0.059	0.941	0.941	0.941	0.955

The F-measure addresses the issue of using a single metric instead of two by combining precision and recall into a single measurement. Specifically, the F-measure is the weighted harmonic mean of precision and recall [52]. The F-measure results for the models range from 85.4% to 96.7%, consistent with the overall prediction accuracy results.

The final evaluation metric presented in Table 11 was the ROC area. Based on the ROC area results obtained using the WEKA software, the RF model demonstrated the best performance of 99.2%. The ANN model followed this trend with 98.4%, the RBF and LR models with 96.7%, the K-NN model with 96.5%, the AdaBoost.M1 model with 96.4%, the BBN model with 96.1%, the CART model with 95.5%, the DT (J48) model with 95.1%, the NB model with 91.9%, the SVM model with 91.3%, and the DS model with 90.3%. The RF model showed the best prediction performance for the ROC area values.

Overall, when assessing the detailed accuracy results of the models, the Random Forest (RF) model consistently delivered the most meaningful results in terms of TP rate, FP rate, precision, recall, F-measure, and ROC area. The FP rate is the most significant at its lowest value, whereas the other metrics are most meaningful at their highest values. The RF model achieved the lowest FP rate at 3.3%, whereas its TP rate, precision, recall, and F-measure were 96.7%, and it achieved the highest ROC area of 99.2%. Among the tested models, the RF model outperformed the others by a considerable margin.

The models were also evaluated based on the Type I and Type II error rates, and the results are presented in Table 12. The Type I and Type II error rates were critical when evaluating the performance of the methods. A Type I error occurs when a modified opinion is classified as unmodified. Conversely, a Type II error occurs when an unmodified opinion is classified as modified. The costs of Type I and Type II errors differed. Classifying a modified opinion as an unmodified opinion can lead to incorrect decisions and severe economic losses, whereas classifying an unmodified opinion as a modified opinion can waste time on additional investigations. Although each model aims to reduce both the Type I and Type II error rates, a model with a Type I error rate lower than the Type II error rate is generally preferred [12]. The K-NN model exhibited the lowest Type I error rate of 2.7%. This was followed by the RF and ANN models, with a Type I error rate of 3.5%. The SVM and DS models showed the lowest Type II error rate of 2.2%, followed by the AdaBoost.M1 model with 2.9%. Overall, the results indicate that the Type I and Type II error classification rates are below 15.2% and 15.6%, respectively, across all models. The RF model had the lowest overall error rate of 3.3%, followed by the K-NN model at 3.6% and the ANN model at 4%.

Table 12. Type I and Type II Error Rates of Models

Model	Type I Error Rate	Type II Error Rate	Overall Error Rate
BBN	0.089	0.098	0.093
NB	0.137	0.156	0.146
LR	0.129	0.047	0.088
ANN	0.035	0.044	0.040
RBF	0.121	0.041	0.081
SVM	0.152	0.022	0.087
K-NN	**0.027**	0.045	0.036
AdaBoost.M1	0.147	0.029	0.088
DT J48	0.040	0.047	0.044
RF	0.035	0.032	**0.033**
DS	0.152	**0.022**	0.087
CART	0.059	0.059	0.059

Although the K-NN classification model performed best with respect to the Type I error rate, and the SVM and DS classification models performed best in terms of the Type II error rate, the RF classification model demonstrated the best overall classification result regarding the total error rate.

In the study, the SMOTE method was applied due to the unbalanced classes, and a balanced data set was obtained. As the model was trained on a dataset after removing outliers and extreme values, it was ensured to be more robust and generalizable than the synthetic data created with SMOTE. The high accuracy rates of the RF, K-NN, and ANN models indicate a possible overfitting of the model to the training data. To evaluate this possibility, not only the accuracy rates but also the type I and type II error rates were analyzed. The obtained low error rates indicate that the models are successful and acceptable in terms of generalizability. Additionally, the comprehensive evaluation using metrics such as Precision, Recall, F-Measure, and ROC Area supports that the models do not overfit to the training data.

The models are evaluated based on their performance results presented in Table 13.

Table 13. Performance Results of Models

Model	Kappa Statistics	MAE	RMSE	RAE (%)	RRSE (%)
BBN	0.813	0.105	0.277	21.00	55.37
NB	0.707	0.164	0.345	32.80	66.03
LR	0.824	0.125	0.253	25.01	50.57
ANN	0.921	0.043	0.187	8.69	37.42
RBF	0.838	0.141	0.250	28.19	50.00
SVM	**0.934**	0.103	0.179	20.54	35.89
K-NN	0.929	**0.036**	0.189	**7.24**	37.80
AdaBoost.M1	0.823	0.128	0.252	25.69	50.32
DT J48	0.913	0.057	0.205	11.42	40.94
RF	**0.934**	0.103	**0.179**	20.54	**35.89**
DS	0.827	0.153	0.276	30.50	55.25
CART	0.882	0.081	0.231	16.19	46.27

Table 13 provides the kappa statistic, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Relative Absolute Error (RAE), and Root Relative Squared Error (RRSE) results for the models. Regarding the Kappa statistic, the RF and SVM models demonstrated the highest performance with a value of 0.934. This was followed by K-NN model at 0.929, ANN model at 0.921, DT (J48) model at 0.913, CART model at 0.882, RBF model at 0.838, DS model at 0.827, LR model at 0.824, AdaBoost.M1 model at 0.823, BBN model at 0.813, and NB model at 0.707. Landis and Koch provided ranges for interpreting the kappa statistic presented in Table 14 [53]. These ranges assist in understanding the performance levels of the models based on their kappa values.

Table 14. Interpretation of Kappa Statistics Results

Kappa Statistics Value	Strength of Agreement
<0.00	Poor
0.00-0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-1.00	Almost Perfect

Based on the values provided in the table, when interpreting the results, it is observed that the BBN model demonstrates good performance, while the other 11 models perform very well. A review of the literature reveals that, apart from the study by [47], this study has yet to consider the kappa statistic. This study also considered the kappa statistic when presenting the results.

The MAE, RMSE, RAE, and RRSE values indicate that the closer they are to zero, the better the performance of the model. According to the MAE values, the K-NN model performed best at 3.6%. Based on the RMSE values, the RF and SVM models exhibited the best performances (17.9%). Regarding the RAE values, the K-NN model performed best at 7.24%. Finally, for the RRSE values, the RF and SVM models performed the best, with 35.89%.

5. CONCLUSION

The primary objective of this study is to predict the type of audit opinion using 12 different data mining methods and to compare the performance of these methods. During the data collection process, 651 companies listed on BIST were considered. A total of 250 companies without financial statements and independent audit reports from 2009 to 2022, and 240 companies comprising financial institutions, brokerage firms, banks, holding companies, and investment firms were excluded from the sample to obtain more meaningful results because of their distinct reporting structures. The final sample includes 161 companies from 19 sectors. The dataset consists of 2,093 firm-year observations from companies listed on BIST from 2010 to 2022, including 1,855 unmodified and 238 modified opinions. Among the modified audit opinions, 223 were qualified opinions and 15 were disclaimers of opinion, with no audit reports containing adverse opinions.

Data were collected from the Public Disclosure Platform website, company websites, financial reports, audit reports, and the Finnet Stokeys Pro database. A total of 28 independent financial and non-financial variables were used. The development and application of the classification model were performed using WEKA software. The training and testing phases were conducted using 10-fold cross-validation, in which the dataset was split into layers. In each iteration, ten training data samples were selected, trained with the chosen classifier, and tested on the selected test data.

The “Interquartile Range” filter was applied to detect outliers and extreme values in the dataset. As a result, 606 outliers and 468 extreme values were removed from the dataset, resulting in 1,286 firm-year observations consisting of 1,189 unmodified opinions and 97 modified opinions. In the subsequent phase, a synthetic minority oversampling technique (SMOTE) filter was applied to address the data imbalance issue. After applying this filter, the number of modified opinion observations increased to 1,188, resulting in a final dataset of 2,377 firm-year observations, consisting of 1,189 unmodified and 1,188 modified opinions.

In this study, 12 different data mining classification methods were used: Bayesian networks, naive Bayes, logistic regression, artificial neural networks, radial basis function, support vector machines, K-nearest neighbor, AdaBoost.M1 algorithm, decision trees (J48), random forest, decision stump, and classification and regression tree (CART). The statistical results of the models were compared based on prediction accuracy, confusion matrix, detailed accuracy results (TP Rate, FP Rate, Precision, Recall, F-Measure, and ROC Area), Type I and Type II error rates, and performance metrics (kappa statistic, MAE, RMSE, RAE, and RRSE).

In terms of prediction accuracy, the Random Forest model demonstrated the best performance at 96.68%, followed by the K-nearest neighbor at 96.42%, Artificial Neural Networks at 96.05%, Decision Trees (J48) at 95.62%, Classification and Regression Tree (CART) at 94.11%, Radial Basis Function at 91.88%, Support Vector Machines at 91.33%, Decision Stump at 91.33%, Logistic Regression at 91.21%, AdaBoost.M1 Algorithm at 91.17%, Bayesian Networks at 90.66%, and Naive Bayes at 85.36%. All 12 models showed a prediction accuracy of over 85%. These results indicate the potential of using public financial statements and independent audit report data to accurately classify audit opinions.

Among the methods reported in the literature for predicting audit opinions, the highest prediction accuracy has been achieved by Decision Trees ([23] (98.2%); [24] (88.8%); [48] (96.3%)), Support Vector Machines ([3] (95.9%); [43] (73%); [51] (75.6%)), Logistic Regression ([40] (62%); [41] (78%); [44] (90%)), and Neural Networks ([5] (88%); [7] (98.1%); [45] (84%)).

According to the findings obtained in this study, the Random Forest method, which belongs to the decision tree category, showed the highest prediction performance for audit opinion classification at 96.68%. This result aligns with the studies by [23], [24], [48], who also found that decision tree-based methods perform best. Based on these results, similar to many previous studies, it is evident that the 12 models used in this study can successfully predict audit opinions.

When reviewing the performance of previous models used to predict audit opinions ([8] (82.22%); [40] (62%); [41] (78%); [42] (78.83%); [43] (73%); [45] (84%); [46] (72.20%); [49] (72.90%); [51] (75.60%)), most of these models have achieved classification accuracy below 85%. In contrast, the models used in this study provided overall classification performances ranging from 85.36% to 96.68% showing that they surpassed the 85% threshold. This high classification accuracy further enhances the practical utility of the models used in this study for predicting audit opinion types.

The models developed in this study can be used as a decision support tool for auditors, as expressed by [54], when assessing potential clients and evaluating potential audit risks. For instance, the model can be applied to a potential client’s financial statements to evaluate the likelihood of receiving a non-unqualified audit opinion. A potential audit opinion can be considered part of the client assessment process. The model can also be used to identify current clients likely to receive a non-unqualified audit opinion during the planning phase of the auditor's work. It can also be utilized in the examination of company audit reports by institutions such as BIST, the Capital Markets Board (SPK), and the Public Oversight Accounting and Auditing Standards Authority (KGK). Companies that receive an audit opinion different from what the model predicts could be targeted for fair reporting investigations. In other words, the models can be used as a complementary "supervisory tool" by regulatory authorities in conducting "audits of audits".

This study is unique because it represents a comprehensive data mining classification investigation to predict independent audit opinion types. By analyzing 12 different data mining methods using WEKA software and comparing their predictive performance, this study is expected to contribute to the literature on independent auditing. Additionally, it is anticipated that this study will provide a more precise definition of data mining, which will assist in solving current business problems and development. This study is expected to shed light on the future study in this area.

Future study could focus on alternative data mining or deep learning methods to predict audit opinion type. Hybrid models can be developed by integrating multiple models into a single framework. Second, the financial and non-financial variables can be expanded, and the results can be compared with those of this study. Finally, the results were evaluated using different data mining software (such as Python, R, Orange, RapidMiner, and Knime).

Acknowledgments

This study developed in the context of the first author's Doctoral Thesis.

Authors' Contributions

Zafer KARDEŞ: %75, Tuğrul KANDEMİR: %25.

Competing Interests

The authors declare that they have no conflict of interest.

References

[1] "Independent auditing standard 315." Accessed: Dec. 18, 2024. [Online]. Available: https://www.kgk.gov.tr/Portalv2Uploads/files/Duyurular/v2/BDS/bdsyeni25.12.2017/BDS%20315-Site.pdf

[2] H. Valipour, F. Salehi, and M. Bahrami, "Predicting audit reports using meta-heuristic algorithms," *J. Distrib. Sci.*, vol. 11, no. 6, pp. 13–19, 2013, doi: 10.13106/jds.2013.vol11.no6.13.

[3] A. K. Nawaiseh, M. F. Abbod, and T. Itagaki, "Financial statement audit using support vector machines, artificial neural networks and k-nearest neighbor: empirical study of UK and Ireland," *Int. J. Simul. Syst. Sci. Technol.*, Mar. 2020, doi: 10.5013/IJSSST.a.21.02.07.

[4] F. A. Amani and A. M. Fadlalla, "Data mining applications in accounting: A review of the literature and organizing framework," *Int. J. Account. Inf. Syst.*, vol. 24, pp. 32–58, 2017, doi: https://doi.org/10.1016/j.accinf.2016.12.004.

[5] O. Pourheydari, H. Nezamabadi-pour, and Z. Aazami, "Identifying qualified audit opinions by artificial neural networks," *Afr. J. Bus. Manag.*, vol. 6, no. 44, pp. 11077–11087, Nov. 2012, doi: 10.5897/AJBM12.855.

[6] S. M. Saif, M. Sarikhani, and F. Ebrahimi, "Finding rules for audit opinions prediction through data mining methods," *SSRN Electron. J.*, 2012, doi: 10.2139/ssrn.2185919.

[7] M. A. Fernández-Gámez, F. García-Lagos, and J. R. Sánchez-Serrano, "Integrating corporate governance and financial variables for the identification of qualified audit opinions with neural networks," *Neural Comput. Appl.*, vol. 27, no. 5, pp. 1427–1444, Jul. 2016, doi: 10.1007/s00521-015-1944-6.

[8] E. Kirkos, C. Spathis, A. Nanopoulos, and Y. Manolopoulos, "Identifying qualified auditors' opinions: a data mining approach," *J. Emerg. Technol. Account.*, vol. 4, no. 1, pp. 183–197, Jan. 2007, doi: 10.2308/jeta.2007.4.1.183.

[9] N. Dopuch, R. W. Holthausen, and R. W. Leftwich, "Predicting audit qualifications with financial and market variables," *Account. Rev.*, pp. 431–454, 1987.

[10] K. Keasey, R. Watson, and P. Wynarczyk, "The small company audit qualification: a preliminary investigation," *Account. Bus. Res.*, vol. 18, no. 72, pp. 323–334, Sep. 1988, doi: 10.1080/00014788.1988.9729379.

[11] D. Zdolšek, T. Jagrič, and M. Odar, "Identification of auditor's report qualifications: an empirical analysis for Slovenia," *Econ. Res.-Ekon. Istraživanja*, vol. 28, no. 1, pp. 994–1005, Jan. 2015, doi: 10.1080/1331677X.2015.1101960.

[12] E. Kirkos, C. Spathis, and Y. Manolopoulos, "Data mining techniques for the detection of fraudulent financial statements," *Expert Syst. Appl.*, vol. 32, no. 4, pp. 995–1003, 2007, doi: 10.1016/j.eswa.2006.02.016.

[13] P. Ravisankar, V. Ravi, G. Raghava Rao, and I. Bose, "Detection of financial statement fraud and feature selection using data mining techniques," *Decis. Support Syst.*, vol. 50, no. 2, pp. 491–500, Jan. 2011, doi: 10.1016/j.dss.2010.11.006.

[14] G. L. Gray and R. S. Debreceny, "A taxonomy to guide research on the application of data mining to fraud detection in financial statement audits," *Int. J. Account. Inf. Syst.*, vol. 15, no. 4, pp. 357–380, Dec. 2014, doi: 10.1016/j.accinf.2014.05.006.

[15] M. Cecchini, H. Aytug, G. J. Koehler, and P. Pathak, “Detecting management fraud in public companies,” *Manag. Sci.*, vol. 56, no. 7, pp. 1146–1160, Jul. 2010, doi: 10.1287/mnsc.1100.1174.

[16] K. Takahashi, Y. Kurokawa, and K. Watase, “Corporate bankruptcy prediction in Japan,” *J. Bank. Finance*, vol. 8, no. 2, pp. 229–247, Jun. 1984, doi: 10.1016/0378-4266(84)90005-0.

[17] W. Kwak, Y. Shi, and C. F. Lee, “Data mining applications in accounting and finance context,” in *Handbook of Financial Econometrics, Mathematics, Statistics, and Machine Learning*, World Scientific, 2020, pp. 823–857. doi: 10.1142/9789811202391_0021.

[18] I. Dutta, S. Dutta, and B. Raahemi, “Detecting financial restatements using data mining techniques,” *Expert Syst. Appl.*, vol. 90, pp. 374–393, Dec. 2017, doi: 10.1016/j.eswa.2017.08.030.

[19] E. Kirkos, C. Spathis, and Y. Manolopoulos, “Applying data mining methodologies for auditor selection,” in *Proceedings 11th Pan-Hellenic conference in informatics (PCI)*, 2007, pp. 165–178. [Online]. Available: https://datalab-old.csd.auth.gr/wp-content/uploads/publications/PCI07ksm.pdf

[20] T. G. Calderon and J. J. Cheh, “A roadmap for future neural networks research in auditing and risk assessment,” *Int. J. Account. Inf. Syst.*, vol. 3, no. 4, pp. 203–236, Dec. 2002, doi: 10.1016/S1467-0895(02)00068-4.

[21] E. Koskivaara, “Artificial neural networks in analytical review procedures,” *Manag. Audit. J.*, vol. 19, no. 2, pp. 191–223, Feb. 2004, doi: 10.1108/02686900410517821.

[22] E. Koskivaara and B. Back, “Artificial neural network assistant (anna) for continuous auditing and monitoring of financial data,” *J. Emerg. Technol. Account.*, vol. 4, no. 1, pp. 29–45, Jan. 2007, doi: 10.2308/jeta.2007.4.1.29.

[23] A. Yaşar, E. Yakut, and M. M. Gutnu, “Predicting qualified audit opinions using financial ratios: Evidence from the Istanbul Stock Exchange,” *Int. J. Bus. Soc. Sci.*, vol. 6, no. 8, pp. 57–67, 2015.

[24] A. Yaşar, “Olumlu görüş dışındaki denetim görüşlerinin veri madenciliği yöntemleriyle tahminine ilişkin karar ve birliktelik kuralları,” *Financ. Anal. Cozum Derg.*, vol. 26, no. 133, 2016, Accessed: Jan. 07, 2025. [Online]. Available: https://search.ebscohost.com

[25] T. Büyüktanir and T. Toraman, “Auditor’s opinions prediction with machine learning algorithms,” in *2020 28th Signal Processing and Communications Applications Conference (SIU)*, IEEE, 2020, pp. 1–4. Accessed: Dec. 18, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9302441/

[26] A. Özcan, “Determining factors affecting audit opinion: evidence from Turkey,” *Int. J. Account. Financ. Report.*, vol. 6, no. 2, Art. no. 2, Aug. 2016, doi: 10.5296/ijafr.v6i2.9775.

[27] B. Adiloglu and B. Vuran, “Identification of key performance ındicators of auditor’s reports: evidence from Borsa Istanbul (BIST),” *Press. Procedia*, vol. 3, no. 1, pp. 854–859, 2017, doi: https://doi.org/10.17261/Pressacademia.2017.665.

[28] M. Mat and S. Onal, “Bağımsız denetim raporlarında denetim görüşünü etkileyen faktörlerin belirlenmesi: Borsa İstanbul’da bir uygulama,” *Muhasebe Bilim Dünya. Derg.*, vol. 21, no. 3, pp. 733–760, 2019, doi: 10.31460/mbdd.533127.

[29] İ. Sakin, “Bağımsız denetim görüşünü etkileyen faktörlerin tespiti: Borsa İstanbula kayıtlı işletmeler üzerine bir uygulama,” Master’s Thesis, Sosyal Bilimler Enstitüsü, 2019. Accessed: Jan. 06, 2025. [Online]. Available: https://acikbilim.yok.gov.tr/handle/20.500.12812/124427

[30] H. A. Ata and I. H. Seyrek, “The use of data mining techniques in detecting fraudulent financial statements: an application on manufacturing firms,” *Suleyman Demirel Univ. J. Fac. Econ. Adm. Sci.*, vol. 14, no. 2, 2009, Accessed: Dec. 18, 2024. [Online]. Available: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=27ff227919838c7c0991b061751f0e246557788d

[31] S. Terzi, “Hile ve usulsüzlüklerin tespitinde veri madenciliğinin kullanımı,” *Muhasebe Ve Finans. Derg.*, no. 54, Art. no. 54, Apr. 2012.

[32] S. Terzi and İ. Şen, “Finansal tablo hilelerinin veri madenciliği yardımıyla tespit edilmesi: üretim sektöründe bir araştırma,” *J. Account. Tax. Stud.*, vol. 5, no. 2, Art. no. 2, Jul. 2012.

[33] M. C. Sorkun and T. Toraman, “Fraud detection on financial statements using data mining techniques,” *Int. J. Intell. Syst. Appl. Eng.*, vol. 3, no. 5, pp. 132–134, Sep. 2017, doi: 10.18201/ijisae.2017531428.

[34] K. Kirda and M. Katkat Özcelik, “Finansal tablo hilesi riski taşıyan şirketlerin veri madenciliği ile belirlenmesi,” *Muhasebe Ve Vergi Uygulamaları Derg.*, vol. 14, no. 2, pp. 609–639, Jul. 2021, doi: 10.29067/muvu.802703.

[35] B. Tatar and H. Kiymik, “Finansal tablolarda hile riskinin tespit edilmesinde veri madenciliği yöntemlerinin kullanılmasına yönelik bir araştırma,” *J. Yaşar Univ.*, vol. 16, no. 64, pp. 1700–1719, Oct. 2021, doi: 10.19168/jyasar.956623.

[36] I. Kilic and S. Onal, “Finansal hilelerin tespit edilmesinde kullanılan veri madenciliği yöntemleri ve Borsa Istanbul’da bir uygulama,” *Muhasebe Ve Denetime Bakış*, vol. 22, no. 67, pp. 181–208, Sep. 2022, doi: 10.55322/mdbakis.1068503.

[37] G. Ozdagoglu, A. Ozdagoglu, Y. Gumus, and G. Kurt Gumus, “The application of data mining techniques in manipulated financial statement classification: the case of Turkey,” *J. AI Data Min.*, no. Online First, 2017, doi: 10.22044/jadm.2016.664.

[38] I. F. Ceyhan, “Bağımsız denetim kalitesini artırıcı bir yöntem olarak veri madenciliği: Borsa Istanbul uygulaması,” PhD Thesis, Sakarya Universitesi (Turkey), 2014. Accessed: Dec. 18, 2024. [Online]. Available: https://search.proquest.com/openview/692e0761f5f5ea36a5a8559ac464da35/1?pq-origsite=gscholar&cbl=2026366&diss=y&casa_token=zh2KnS5lw3sAAAAA:PdmSn3AqhdFbuV9omDXFsQ-lyOuy1_dnp9Q0HpkNiebBNVEMyW_TafJoFqO4F1QbkgcFl3Zozw

[39] N. Cömert and M. Kaymaz, “Araç sigortası hilelerinde veri madenciliğinin kullanımı,” *Marmara Üniversitesi İktisadi Ve İdari Bilim. Derg.*, vol. 41, no. 2, pp. 364–390, Jan. 2020, doi: 10.14780/muiibd.665058.

[40] E. K. Laitinen and T. Laitinen, “Qualified audit reports in Finland: evidence from large companies,” *Eur. Account. Rev.*, vol. 7, no. 4, pp. 639–653, Dec. 1998, doi: 10.1080/096381898336231.

[41] C. T. Spathis, “Audit qualification, firm litigation, and financial ınformation: an empirical analysis in greece,” *Int. J. Audit.*, vol. 7, no. 1, pp. 71–85, Mar. 2003, doi: 10.1111/1099-1123.00006.

[42] C. Spathis, M. Doumpos, and C. Zopounidis, “Using client performance measures to identify pre-engagement factors associated with qualified audit reports in Greece,” *Int. J. Account.*, vol. 38, no. 3, pp. 267–284, Sep. 2003, doi: 10.1016/S0020-7063(03)00047-5.

[43] M. Doumpos, C. Gaganis, and F. Pasiouras, “Explaining qualifications in audit reports using a support vector machine methodology,” *Intell. Syst. Account. Finance Manag.*, vol. 13, no. 4, pp. 197–215, Dec. 2005, doi: 10.1002/isaf.268.

[44] C. Caramanis and C. Spathis, “Auditee and audit firm characteristics as determinants of audit qualifications: Evidence from the Athens stock exchange,” *Manag. Audit. J.*, vol. 21, no. 9, pp. 905–920, Dec. 2006, doi: 10.1108/02686900610705000.

[45] C. Gaganis, F. Pasiouras, and M. Doumpos, “Probabilistic neural networks for the identification of qualified audit opinions,” *Expert Syst. Appl.*, vol. 32, no. 1, pp. 114–124, Jan. 2007, doi: 10.1016/j.eswa.2005.11.003.

[46] C. Gaganis, F. Pasiouras, C. Spathis, and C. Zopounidis, “A comparison of nearest neighbours, discriminant and logit models for auditing decisions,” *Intell. Syst. Account. Finance Manag.*, vol. 15, no. 1–2, pp. 23–40, Jan. 2007, doi: 10.1002/isaf.283.

[47] N. Stanišić, T. Radojević, and N. Stanić, “Predicting the type of auditor opinion: Statistics, machine learning, or a combination of the two?,” *Eur. J. Appl. Econ.*, vol. 16, no. 2, pp. 1–58, 2019, doi: 10.5937/EJAE16-21832.

[48] A. K. Nawaiseh and M. F. Abbod, “Financial statement audit utilising naive bayes networks, decision trees, linear discriminant analysis and logistic regression,” in *The ımportance of new technologies and entrepreneurship in business development: in the context of economic diversity in developing countries*, B. Alareeni, A. Hamdan, and I. Elgedawy, Eds., Cham: Springer International Publishing, 2021, pp. 1305–1320. doi: 10.1007/978-3-030-69221-6_97.

[49] H. Zarei, H. Yazdifar, M. Dahmarde Ghaleno, and R. Azhmaneh, “Predicting auditors’ opinions using financial ratios and non-financial metrics: evidence from Iran,” *J. Account. Emerg. Econ.*, vol. 10, no. 3, pp. 425–446, Jun. 2020, doi: 10.1108/JAEE-03-2018-0027.

[50] M. I. H.-P. Wu and L. Li, “The BP neural network with adam optimizer for predicting audit opinions of listed companies,” 2021. Accessed: Jan. 07, 2025. [Online]. Available: https://www.semanticscholar.org/paper/The-BP-Neural-Network-with-Adam-Optimizer-for-Audit-Wu-Li/a08f165c1ad8b530c0a1a356d9233639e23bb52f

[51] A. Saeedi, “Audit opinion prediction: a comparison of data mining techniques,” *J. Emerg. Technol. Account.*, vol. 18, no. 2, pp. 125–147, Sep. 2021, doi: 10.2308/JETA-19-10-02-40.

[52] N. Japkowicz and M. Shah, *Evaluating learning algorithms: a classification perspective*. Cambridge University Press, 2011.

[53] J. R. Landis and G. G. Koch, “The measurement of observer agreement for categorical data.,” *Biometrics*, vol. 33, no. 1, p. 159, Mar. 1977, doi: 10.2307/2529310.

[54] G. S. Monroe and S. T. Teh, “Predicting uncertainty audit qualifications in Australia using publicly available information,” *Account. Finance*, vol. 33, no. 2, pp. 79–106, Nov. 1993, doi: 10.1111/j.1467-629X.1993.tb00200.x.