jotaf Journal of Tekirdag Agricultural Faculty
Tekirdağ Ziraat Fakültesi Dergisi

Mart/March 2026, 23(2)
Başvuru/Received: 23/05/25
Kabul/Accepted: 18/02/26
DOI: 10.33462/jotaf.1704855

http://dergipark.gov.tr/jotaf
http://jotaf.nku.edu.tr/

ARAŞTIRMA MAKALESİ **RESEARCH ARTICLE**

Selection of Innovative Technologies in Greenhouse Systems: A Multi-Criteria Approach Based on Patent Data*

Sera Sistemlerinde Yenilikçi Teknoloji Seçimi: Patent Verilerine Dayalı Çok Kriterli Bir Yaklaşım

Ali KILIÇ[1], Hakan EREN[2*], Uğur GÖÇEN[3]

Abstract

The ever-increasing pace of technological development has led to shorter lifecycles for technological projects, making it crucial for researchers and technology companies to carefully consider when, where, and how much to invest. Moreover, the current trend of technological research and innovation transitioning from within a single sector to a more interdisciplinary and collaborative framework involving multiple companies and groups heightens the necessity for businesses to gather information to identify collaboration areas and potential partners. In this context, for an organization working in a certain field and aiming to innovate, the question of which technologies it should invest in to take its work forward is among the most important problems of innovation today. This study proposes an innovative and objective method for selecting technologies by leveraging patent data, data mining algorithms, and Multi-Criteria Decision-Making (MCDM) techniques. The research focuses on greenhouse technologies, a vital area within agricultural innovation, and uses patent data to identify promising technological directions. Patent records from the European Patent Office (EPO) were collected using a custom software tool that queries patents within the CPC classification Y02A40/25 (Greenhouse Technologies). After cleaning the data, the FP-Growth algorithm was applied to identify frequently co-occurring technology classifications. Five key criteria were used to evaluate these technology pairs: support value (frequency), average patent age, average forward citations, average backward citations, and the average patent strength of leading applicant companies. Using the Entropy Weight method, objective weights were assigned to each criterion. The TOPSIS method was then applied to rank the identified technology pairs in terms of their overall suitability for investment and innovation. The results indicated that greenhouse cultivation (A01G9), hydroponic farming (A01G31), plant processing (A01G7), agriculture-related technologies (Y02P60), and business-specific software systems (G06Q50) are the most strategic areas for innovation. Notably, the inclusion of G06Q50 underscores the growing importance of software and digital infrastructure in greenhouse innovation. This suggests that companies aiming to advance in this field should not only enhance their core greenhouse technologies but also invest in complementary software and algorithmic solutions. In conclusion, the study presents a novel framework for technology selection that can guide R&D investment decisions using patent data, especially in sectors where innovation plays a critical role.

Keywords: Technology selection, Greenhouse technologies, Patent analysis, Multi-Criteria Decision Making, FP-Growth, Entropy weight

[1]Ali Kılıç, Konya Food and Agriculture University, Faculty of Agriculture and Natural Sciences, Konya, Türkiye. E-mail: ali.kilic@gidatarim.edu.tr OrcID: 0000-0003-2777-0876
[2]***Sorumlu Yazar/Corresponding Author:** Hakan Eren, OSTIM Technical University, Department of Management Information Systems, Ankara, Türkiye. E-mail: hakan.eren@ostimteknik.edu.tr OrcID: 0000-0002-5222-2938
[3]Uğur Göçen, DB Masters Proje ve Tasarım Yönetimi A.Ş. Yıldız Teknopark, İstanbul, Türkiye. E-mail: ugocen@gmail.com OrcID: 0000-0002-1681-9841
Atıf: Kılıç, A., Eren, H., Göçen, U. (2026). Sera sistemlerinde yenilikçi teknoloji seçimi: patent verilerine dayalı çok kriterli bir yaklaşım. *Tekirdağ Ziraat Fakültesi Dergisi,* 23(2): 755-773.
Citation: Kılıç, A., Eren, H., Göçen, U. (2026). Selection of innovative technologies in greenhouse systems: a multi-criteria approach based on patent data. *Journal of Tekirdag Agricultural Faculty,* 23(2): 755-773.

*This study was summarized from Uğur Göçen's MSc/PhD thesis.

Öz

Teknolojik gelişmelerin artan hızı, teknolojik projelerin yaşam döngülerini kısaltmış ve bu durum, araştırmacılar ile teknoloji şirketlerinin ne zaman, nerede ve hangi ölçüde yatırım yapacaklarını titizlikle planlamalarını zorunlu kılmıştır. Teknolojik araştırma ve inovasyonun disiplinler arası ve çok paydaşlı bir yapıya yönelmesi, işletmelerin iş birliği alanlarını ve potansiyel ortaklarını belirlemek amacıyla kapsamlı veri toplama ihtiyacını artırmaktadır. Bu bağlamda, belirli bir alanda çalışan ve yenilik yapmayı hedefleyen bir organizasyon için, çalışmalarını ileriye taşımada, hangi teknolojilere yatırım yapması gerektiği sorusu, günümüz inovasyonunun en önemli sorunları arasındadır. Bu çalışma, patent verileri, veri madenciliği algoritmaları ve çok kriterli karar verme (ÇKKV) tekniklerinden yararlanarak teknolojilerin seçilmesi için yenilikçi ve nesnel bir yöntem önermektedir. Araştırma, tarımsal inovasyon içinde hayati bir alan olan sera teknolojilerine odaklanmakta ve gelecek vaat eden teknolojileri belirlemek için patent verilerini kullanmaktadır. Patent verileri, özel bir yazılım aracı kullanılarak Avrupa Patent Ofisi'nden (EPO) alınmıştır. CPC sınıflandırması Y02A40/25 (Sera Teknolojileri) içindeki patentler sorgulanmıştır. Veriler temizlendikten sonra, sıklıkla birlikte görülen teknoloji sınıflandırmalarını belirlemek için FP-Büyüme algoritması uygulanmıştır. Bu teknoloji çiftlerini değerlendirmek için beş temel ölçüt: destek değeri (frekans), ortalama patent yaşı, ortalama ileri atıflar, ortalama geri atıflar ve önde gelen başvuru şirketlerinin ortalama patent gücü kullanılmıştır. Entropi Ağırlığı yöntemi kullanılarak, her ölçüte nesnel ağırlıklar atanmıştır. Belirlenen teknoloji çiftleri, yatırım ve inovasyon için genel uygunlukları açısından TOPSIS yöntemi ile sıralanmıştır. Sonuçlar, sera yetiştiriciliğinin (A01G9), hidroponik çiftçiliğin (A01G31), bitki işlemenin (A01G7), tarımla ilgili teknolojilerin (Y02P60) ve işletmeye özgü yazılım sistemlerinin (G06Q50) inovasyon için en stratejik alanlar olduğunu göstermiştir. Özellikle, G06Q50'nin dahil edilmesi, sera inovasyonunda yazılım ve dijital altyapının artan önemini vurgulamaktadır. Bu, alanda ilerlemeyi hedefleyen şirketlerin yalnızca temel sera teknolojilerini geliştirmekle kalmayıp aynı zamanda tamamlayıcı yazılım ve algoritmik çözümlere de yatırım yapmaları gerektiğini göstermektedir. Sonuç olarak, çalışma, özellikle inovasyonun kritik bir rol oynadığı sektörlerde, patent verilerini kullanarak Ar-Ge yatırım kararlarına rehberlik edebilecek teknoloji seçimi için yeni bir çerçeve sunmaktadır.

Anahtar Kelimeler: Teknoloji seçimi, Sera teknolojileri, Patent analizi, Çok kriterli karar verme, FP-Büyüme algoritması, Entropi ağırlığı

1. Introduction

The rapid advancement of technology has led to shorter lifecycles for technological projects, making it crucial for researchers and technology companies to carefully consider when, where, and how much to invest. The change in technical R&D has shifted from a singular industry focus to a collaborative effort between different companies and groups. Moreover, the shift towards interdisciplinary and inter-company/group collaborations has increased the need for businesses to gather information about potential partners from various disciplines and organizations. To meet this demand, it is essential to explore various tracking methodologies for technical advancements and establish a decision-making system that evaluates progress in specific areas of work, potential investment opportunities, and viable collaboration. Each variable provides valuable insights into the potential success of an R&D project. However, relying solely on market data for research may not be sufficient, as the market is continuously evolving. To address this, decision makers require additional tools to predict technological advancements beyond market forecasts.

Companies and countries have implemented techniques to predict future technological trends based on patent analysis (Lee et al., 2009a). Patents serve the dual purpose of providing legal safeguards for intellectual property rights (IPRs) and outlining specific information about the technology being developed (Park et al., 2014). The use of patent research to forecast potential technologies is crucial to create effective management strategies. It helps minimize unnecessary R&D expenses, prevent unexpected technology infringement costs from competitors, and develop R&D programs that protect core patents (Kim et al., 2008; Kim et al., 2015). R&D decision-makers can greatly benefit from using patent analysis-based technology forecasting methods. These methods provide valuable insights that can enhance a company's future competitiveness and serve as essential tools for making technology management decisions (Ernst, 2003).

When dealing with a large dataset such as patent data, several methods can be used for analysis. While numerous studies delve into patent analysis, most focus on comparing the number of patents related to the searched terms. The accuracy and efficiency of research results greatly depend on the methods used to analyze technically based and diverse patent data that contain information from various industrial sectors.

Greenhouse technology is an important field that showcases advancements in agriculture and its future potential. This allows controlled environments that enable agriculture to thrive. To ensure progress in this field and the agricultural sector, it is crucial to identify the key technological areas that are essential for innovation. Therefore, the main objective of this study is to determine the direction of technological development in greenhouse technology.

The choice of greenhouse technologies as the focus of this research was driven by the differences in methods used for analyzing patent data. A word-based analysis might direct researchers to innovative studies on greenhouse gases, which are not necessarily related to the primary research topic of greenhouse technologies. Therefore, the patent class defined for greenhouse technology is used to identify the relevant patent dataset for this study. This approach has become common in recent years, as patents are classified according to their technological fields, simplifying the identification of relevant data.

This study aims to tackle the challenge of identifying the most suitable technologies for investment by leveraging data mining algorithms on patents categorized under Patent Class for Greenhouse Technologies. By employing the Entropy-weighted TOPSIS method, the research seeks to offer an objective and innovative solution to the multi-criteria decision-making problem within the realm of innovation for organizations operating in specific industries. The study aimed to address its objectives by answering two key research questions. Firstly, it aimed to identify which technologies companies involved in innovation within a specific field should prioritize for development and enhancement alongside their primary technology domains to advance in the field. Secondly, it explored the feasibility of using patent data and analysis to overcome the challenge of selecting technologies in studies aiming for innovative outcomes.

2. Literature Review

2.1. Technology Selection

Managers often encounter the challenge of selecting the most suitable technology for R&D from a range of available choices. This is a difficult process because of the fast-paced advancement, intricate nature, and multiplicity of the technologies. Present methodologies for technology selection typically involve assessing the financial viability of available alternatives or relying on traditional investment justifications. Technology selection is a crucial aspect that businesses must consider. With the constant evolution of technology and the abundance of available options, selecting the appropriate technology can be challenging. However, embracing new technology can open new possibilities for diversifying products and generating new business ideas.

Extensive studies have focused on selecting the appropriate technology for R&D, along with numerous frameworks and methods proposed to justify investments. After a new technology passes all R&D stages and is demonstrated at bench, pilot plant, and semi-commercial scales, it can provide significant cost benefits. To understand patent protection, a thorough examination of intellectual property is crucial. Given the escalating complexity and abundance of technology options, businesses are seeking technologies that are straightforward to understand and implement. To fulfill technological needs, businesses employ diverse technology management methodologies and enact technological strategies and planning.

As technology advances, navigating its complexities becomes increasingly challenging. This challenge is compounded by factors such as the convergence of multiple technologies, numerous potential applications, rising development costs, and accelerated dissemination. Chan et al. (2000) analyzed different evaluation techniques and provided a sample method that employs fuzzy and non-fuzzy decision support systems for technology selection. In their study, Torkkeli and Tuominen (2002) emphasized the importance of a company's technological core competencies and the need to keep them updated to maintain the company's sustainability, competitiveness, and R&D capabilities. Their recommendation suggests that decision-makers should focus on developing and reinforcing their company's core competencies when selecting technology. This is because a company's survival relies on the added value it generates through its core competencies and the competitive advantages it offers in the market. Therefore, it is essential to use the right tools and information when selecting technologies that can enhance a company's core competencies.

Salvadó et al. (2022) devised a decision support system for selecting technologies appropriate for a particular domain, as evidenced by their case study on near-zero-energy building technology. Instead of adopting a specific Multi-Criteria Decision Making (MCDM) approach, they developed a novel system to rank the gathered data according to predefined criteria. Oztaysi (2014) highlighted the challenge of selecting the appropriate technology for creating a content management system. He used the AHP-TOPSIS-Gray method to rank the technologies under consideration.

According to Wang (2012), an enterprise's choice of technology can impact its investments in R&D, and making a wise selection of technology can enable a company to optimize its resources and expedite the market launch of its products. To make informed decisions about technology, patent analysis is a reliable method that provides decision-makers with valuable information on forecasts, and development prospects. Wang (2012) suggested that by analyzing patent searches and the number of patents obtained in specific fields during a selected timeline, one can infer the technology life cycle. This trend analysis can serve as a crucial indicator of technological choice.

Shen et al. (2010) proposed a hybrid approach to technology selection, which first involved identifying patents that were interlinked or shared a common source within a specific technological domain. They then applied the Analytic Hierarchy Process (AHP) to assign weights to predefined criteria, thereby categorizing and prioritizing technologies in the field of Organic Light-Emitting Diodes (OLEDs).

2.2 Utilizing Patent Data to Explain and Predict Technological Development

Patents function both as legal safeguards for intellectual property and as critical sources of technological information. They provide early signals of innovation that precede market data, making them essential for technology forecasting and strategic R&D planning (Griliches, 1998; Abbas et al., 2014). Beyond protecting intellectual property rights, patent information offers insights into the novelty, maturity, and diffusion of technologies, supporting firms in minimizing unnecessary R&D costs, avoiding infringement risks, and identifying new market opportunities (Ernst, 1997b; Kim et al., 2015). Filing patents in multiple jurisdictions, despite its high cost, often signals significant commercial expectations and strengthens a company’s technological positioning in international markets.

The analytical use of patent data has evolved significantly. Early studies relied on simple patent counts as proxies for technological progress and innovation, often linking patent numbers with national economic performance (Crosby, 2000; Paci et al., 1997). This approach, however, overlooked qualitative differences in patent value. Later, citation analysis was introduced to assess the technological influence of patents by examining forward and backward citations. Content-based methods, including keyword analysis and text mining of abstracts and claims, further expanded analytical possibilities by identifying emerging topics and innovation gaps (Lee et al., 2009b; Yoon et al., 2008). Nevertheless, these approaches face challenges of ambiguity and limited ability to capture functional and structural aspects of inventions (Cascini and Zini, 2008; An et al., 2018).

More recently, classification-based methods have gained prominence. The Cooperative Patent Classification (CPC), jointly developed by the EPO and USPTO, provides a detailed hierarchical taxonomy of technologies, consisting of sections, classes, subclasses, groups, and subgroups. Unlike keyword-based searches, CPC classification enables more precise retrieval of relevant patents and allows for systematic analysis of technological domains. The granularity of CPC coding enables researchers to detect co-occurrence patterns between technological domains, uncovering complementarities that conventional approaches often overlook (Kim and Bae, 2017; Altuntaş and Sezer, 2021).

The integration of big data analytics and machine learning has further advanced patent intelligence. Techniques such as association rule mining, clustering, and network analysis are increasingly applied to identify frequently co-occurring classifications, map innovation networks, and forecast technological trajectories (Jun, 2011; Park et al., 2015; Choi et al., 2021). More recent approaches integrate topic mining with patent evaluation to systematically uncover technology opportunities in smart agriculture, demonstrating the value of combining text-based and evaluative methods for agricultural innovation studies (Song and Ran, 2023). Dimensionality-reduction methods such as principal component analysis and self-organizing maps have been used to visualize technology vacuums and highlight underexplored opportunities (Lee and Park, 2005). Collectively, these developments demonstrate that CPC-driven and data-mining-based patent analysis provides a robust framework for forecasting technological development, supporting evidence-based decision-making, and guiding R&D investment strategies.

2.3 Greenhouse Technologies

Greenhouse technologies have become a central pillar of modern agricultural innovation, providing controlled environments that increase yield, reduce vulnerability to weather extremes, and enable year-round production (Çinkılıç et al., 2014). Beyond their traditional role of protecting crops, contemporary greenhouses integrate advanced cultivation techniques-such as soilless farming (hydroponics, aeroponics, and aquaponics)-that reduce dependence on arable land and allow agricultural production in urban and constrained environments. These soilless approaches improve resource-use efficiency and lessen the risk of soil degradation while supporting high-density, high-value cropping systems (Resh, 2022; Al-Chalabi, 2015). Greenhouses have emerged as hubs for the development and testing of numerous agricultural innovations due to their controlled production environments. This allows for experimentation of techniques before implementation in the field or other settings (Sivri and Çanakçı, 2024).

2.3.1 Digitalization, Sensing and Automation

Recent literature emphasizes the accelerating adoption of digital technologies in greenhouse systems. IoT sensors, remote monitoring, and automated climate control enable precise regulation of temperature, humidity, CO_2 and light, improving both productivity and energy efficiency (Mohammadian et al., 2020). Machine-learning algorithms and computer-vision methods are increasingly applied for early disease detection, phenotyping, and predictive control of microclimates, while robotics and automation support labor-intensive tasks such as transplanting and harvesting (Van Straten et al., 2019). These digital tools transform greenhouses into cyber-physical systems where data streams drive adaptive management and continuous optimization.

2.3.2 Sustainability and Resource Efficiency

Sustainability is a recurring theme in greenhouse research. Studies report progress on energy-saving measures (e.g., improved insulation, LED lighting strategies, heat recovery), integrated water management (closed-loop irrigation and nutrient recycling), and methods to reduce greenhouse operations' carbon footprints (Pérez-Alonso et al., 2020). The interest in energy-efficient greenhouse designs is reflected in patent activity and R&D projects focused on both hardware (structural materials, glazing, thermal screens) and software (energy management and scheduling algorithms) solutions, pointing to a convergence between engineering and information systems.

2.3.3 Novel Production Systems and Urban Applications

Greenhouses are increasingly being combined with vertical farming and urban agriculture paradigms to address local food security and reduce supply-chain distances. Vertical and multi-tier greenhouse designs, often coupled with aeroponic or hydroponic subsystems, demonstrate strong potential for high-yield production in urban settings, while enabling precise nutrient and water control (Al-Chalabi, 2015; Resh, 2022). These systems also serve as testbeds for circular-economy practices (e.g., waste-to-nutrient streams, energy integration) that help close resource loops at the facility level.

2.3.4 Patents, Technology Intelligence, and Methodological Advances

Patent analytics have become a preferred method to identify emerging technological directions in greenhouse systems. Bibliometric and patent-class based studies reveal growth in technologies related to irrigation, microclimate control, and energy systems within greenhouse classifications (Aznar-Sánchez et al., 2020). Combining association mining (e.g., FP-Growth) with MCDM tools enables more nuanced identification of promising, cross-disciplinary technology pairs - for instance, the coupling of cultivation techniques with business-adapted information systems - which conventional keyword searches often miss (Altuntaş and Sezer, 2021; Park et al., 2014).

While technical sophistication is increasing, several gaps remain. Integration of digital systems across hardware and software layers (interoperability), lifecycle assessments of novel greenhouse configurations, and socio-economic analyses of automation adoption (labor displacement, skills needs) require further study. Moreover, semantic analyses and natural language processing of patent claims and full texts can improve technology mapping beyond CPC labels, especially for recent patents that may lack complete classification entries. These methodological refinements will support more actionable technology intelligence for R&D managers and policymakers.

3. Research Methodology

The research primarily analyzes patent data through statistical techniques based on descriptive and comparative approaches. MCDA methods are employed to develop the decision support system. These methods structure statistical data in line with expert judgments and aim to uncover the underlying rationale behind the data. Accordingly, both qualitative and quantitative research approaches were applied in this study.

To systematically identify promising technological domains and potential investment targets, a multi-stage methodological framework was developed, as illustrated in *Figure 1*. The process begins with the collection and cleaning of raw patent data, followed by the extraction of frequent co-occurring CPC code pairs using the FP-Growth algorithm. Decision criteria—including frequency, patent age, citations, and firm strength—are then defined to construct a decision table. Objective weighting of these criteria is performed through the entropy method, and the resulting weighted matrix is analyzed using the TOPSIS technique to rank CPC pairs against an ideal solution. Finally, strategic areas are identified based on ranked alternatives, guiding the selection of high-potential innovation and investment opportunities.

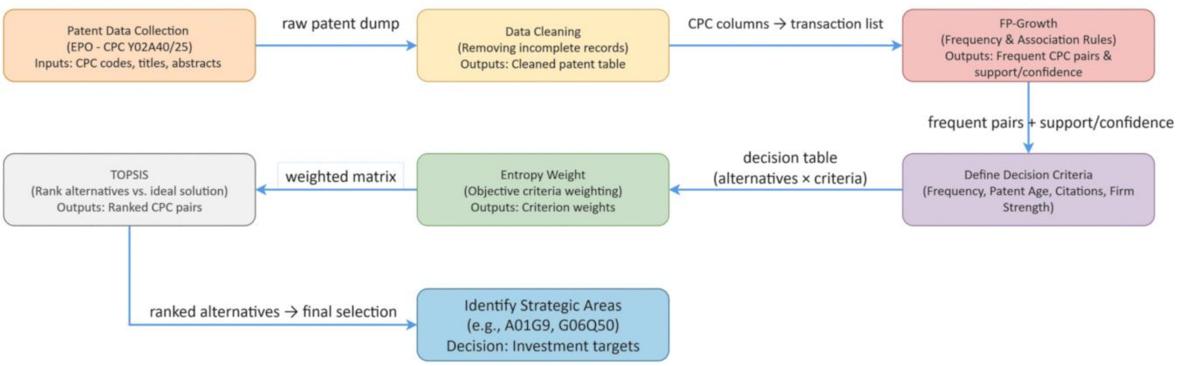

Figure 1. Research Methodology

3.1 Patent Data Collection

The primary data source for this study is the patent records of the European Patent Office (EPO). The EPO permits researchers with suitable authorization to access both processed and unprocessed patent data. This analysis relies on assessing patent data, information, and regional R&D indicators generated by organizations such as the European Union, UNDP, and the Turkish Statistical Institute to enrich the geographical aspect of the evaluation. Data analysis, natural language processing, and machine learning techniques were also employed in the study.

Previous studies have typically focused solely on analyzing patent registration methods and the information contained within, often limited to examining class (IPC - CPC) or title names, or a relatively small subset of patents compared to the total. This method neglects the original information embedded in patents, thereby restricting the potential insights that could be obtained from analysis. Therefore, this study plans to perform semantic analyses on patent claims and summary information, as well as conduct analytical classifications based on factors like class, inventor, and patent family. Additionally, more precise estimation and identification analysis will be carried out by providing access to a much wider patent pool with the program to be developed within the research using the Open Patent Services (OPS) tool of the European Patent Office.

Purely patent-based analyses often fall short of accurately portraying the current landscape because each patent carries varying significance and impact. It provides only basic vector data regarding potential technological directions. However, achieving more precise analyses is feasible by supplementing patent data with information from sources such as the Innovation Scorecard, project acceptance data, and export-import data. By acquiring clean and relevant data, it will be feasible to develop artificial intelligence systems capable of uncovering fundamental insights and information nodes within the data through machine learning techniques.

3.2 Software Development for Patent Retrieval

This study used The International Patent Documentation File (INPADOC) database and OPS Internet services. INPADOC is updated each week by the European Patent Office (European Patent Office, 2022) and a worldwide repository of bibliographic and legal status data. The INPADOC database gathers patent information from patent authority gazettes, registries, and web services, enabling it to upload data from over 100 patent authorities (Lingua, 2005). The ESPACENET database is similar to the INPADOC database, which is accessible via the European Patent Office. ESPACENET offers free open access to data on inventions and scientific discoveries from 1782 to the present. This database, which contains details of more than 110 million patent papers worldwide, is updated daily. The study used the European Patent Office's OPS online service to gather data via a software application.

The research involved using an API designed to interface with the OPS system at various stages. A Python-based software program was developed to obtain the necessary data. A key feature of the software application is its ability to connect to the OPS system through the appropriate API and execute search queries using the Contextual Query Language (CQL). CQL is a formal language that can be used to express searches for information retrieval systems, including online directories, bibliographic catalogs, and museum collections. The aim of query design is to preserve the expressive capacity of complex languages while ensuring queries remain understandable and writable by humans. This involves maintaining a natural flow of thought and ensuring human readability. A CQL query, utilizing Boolean operators, can comprise one or multiple search clauses. It may begin with a sort specification, followed by the keyword "sortBy." Additionally, prefix assignments, which assign shorter names to context set identifiers, may be included. The following is an example of a CQL search that would be used to look for the terms "precision agriculture" and "smart agriculture" in all patent text fields for patents filed between 2016 and 2021.

(ftxt = "precision agriculture" OR ftxt = "smart agriculture") AND pd within "2016,2021"

The following limitations of the EPO-OPS system affect software effectiveness, even though it can be used to analyze the INPADOC patent database. The OPS system:

1- Restricts the retrieval of patent data to a maximum of 100 records per query.

2- Imposes a maximum data retrieval size limit of 1 MB per minute.

3- Allows researchers to use up to 4GB of data transfer for free each week.

As a result of these limitations, the software's operation is designed to break down queries into small batches of no more than 100 patents across multiple cycles, ensuring compliance with the data transfer rate specified for each minute.

3.3 The process of acquiring, organizing, and refining the dataset

The first step in the research process entails collecting patent data, which forms the basis of the study. This involves identifying the most appropriate query and employing specialized computer software designed for patent analysis, as detailed in the methodology section. By implementing text-based query methods, it was discovered that there are patents related to greenhouse gases and technologies, in addition to greenhouse technologies. In patent research studies, utilizing patent classification is crucial rather than relying solely on text-based searches. This is because patent databases classify patents based on their content and technological fields during the application and examination stages. Therefore, patent classification enables access to patents in targeted technology fields with high accuracy. The CPC system is commonly used owing to its detailed and modern approach. The research study commenced by identifying the CPC code for Greenhouse Technologies to retrieve pertinent patents. A patent search was conducted utilizing the Y02A40/25 class in the EPO OPS platform using our patent search software. The CPC categorizes technologies based on their functions and further classifies them into lower and upper classes.

The Y02A40/25 class is referred to as "Greenhouse Technologies" in this particular context. This classification is based on function and includes neighboring classes, such as water and soil control and conservation, as well as seed and other agricultural technologies. Patents related to greenhouse technologies falling under the "Y02A40/25" class were extracted from the OPS system. For patent analysis, the date range of January 1, 2001, to December 31, 2021, was selected to ensure the accuracy and relevance of the data. An automatic query cycle was employed to extract the patent data on a monthly basis, with each query requesting fewer than 100 patents. For months containing more than 100 patents, the data extraction process involved partitioning within each month.

A raw dataset comprising 67,506 patents was successfully downloaded. This was achieved through the execution of 872 queries on OPS using our software. The downloaded data were initially listed in the software database and sorted chronologically before undergoing data-cleaning procedures. To clean up the raw patent data, we need to segregate the files lacking crucial information that cannot be utilized for analysis. This process entails identifying patents essential for our targeted analyses. The INPADOC database system, which stores all the legal data of patents, may have some missing information in its patent files, particularly for recently filed patents or variations in data from certain Far Eastern countries. To ensure precise analysis, we defined specific criteria for identifying patent documents. An example of a patent bibliography is illustrated in *Figure 2*. These criteria:

1. The patent application requires a title and abstract, both of which must be submitted into the system. This signifies that the application has been submitted and is being processed, meeting the necessary criteria.

STRAWBERRY PLANTING FRAME WITH AUTOMATIC IRRIGATION FUNCTION

Page bookmark	WO2022222338 (A1) - STRAWBERRY PLANTING FRAME WITH AUTOMATIC IRRIGATION FUNCTION
Inventor(s):	WU CHONG [CN]; JIANG LILI [CN] +
Applicant(s):	SHANDONG INST OF POMOLOGY [CN] +
Classification:	- international: ***A01G13/02; A01G27/00; A01G9/02***
	- cooperative: **A01G13/02 (CN); A01G27/008 (CN); A01G9/023 (CN); Y02A40/25** (EP)
Application number:	WO2021CN114876 20210827 Global Dossier
Priority number(s):	CN202110427780 20210421
Also published as:	CN113508703 (A) CN113508703 (B)

Figure 2. An Example of a Patent Bibliography

2. The patent has a designated applicant. As the details of the applicant's company are also employed in patent classification and ranking, this requirement was established to exclude patents that lacked this crucial information from the analysis.

The total number of patent listings decreased to 59,257 as patents lacking abstracts, titles, and applicant data were removed from the list.

In the second step of the data cleaning process, the focus was on analyzing CPC classes. The primary goal of this research was to identify technologies that could be used by individuals and organizations for future studies on Greenhouse Technology. The main aim was to identify technologies mentioned in greenhouse patents, as well as those related to it. The research question pertaining to this topic focuses on identifying technology companies conducting innovative studies in a specific field that should invest in, apart from their primary technology, to make progress and generate innovative outcomes.

In this scenario, patents classified solely under the Y02A40/25 CPC class do not offer any novel technological opportunities. Therefore, we excluded patents exclusively falling within the CPC class Y02A40/25 from our list without further examination. After completing this process, the final number of patents in the list has been reduced to 31,040. The patent dataset has been finalized for data analysis following the conclusion of the final data selection process.

3.4 Multi-Criteria Decision Analysis (MCDA) Framework

3.4.1 FP Growth Algorithm

In association rule mining, the primary objective is to uncover meaningful relationships among data points within a dataset. As the importance of association rules has increased, a variety of algorithms have been developed and refined to meet evolving industry needs. Among these, the FP-Growth algorithm has become one of the most widely adopted approaches due to its efficiency and low computational cost. FP-Growth is particularly effective in identifying frequent patterns within large datasets, making it a cornerstone technique in modern association rule mining. This is largely owing to the frequent pattern tree (FP-Tree), which keeps the entire database in a small and dense data structure. The entire database is only examined twice in FP-Growth, in contrast to Apriori-based algorithms. The second is used to build the tree structure, and the first is used to determine each item's support value. For large datasets, FP-Growth is advantageous because it eliminates the need to continually create new candidates and search databases. The time required to construct association rules using Apriori grows exponentially with database size, while the scanning time for FP growth increases linearly. Therefore, the FP growth algorithm is significantly faster than Apriori (Han et al., 2004).

The first stage of the FP-Growth algorithm involves scanning the database and determining the support value for each item. Items with support values equal to or greater than the assigned threshold are then ranked in descending order. Additionally, the elements within each transaction in the database are also ranked in descending order based on their support value. To construct the frequent pattern tree, the first step involves establishing a new node known as the root. Subsequently, each transaction is incorporated into the tree, maintaining the sequence of its items. If an item is absent in the tree, a new node is created with a support value of 1. Additionally, support values for individual items are monitored. If an item already exists, only the support value of the node is incremented by 1. Moreover, a header table is maintained to establish connections between nodes, indicating each node's origin. This table facilitates simultaneous connection of identical nodes in the tree through pointers. Once the tree-building process is completed, the FP-Growth algorithm is applied to this structure. It begins with the least frequent item, identifying the paths where the item occurs. The support value of the item is assigned to each path as the path's support value, forming the basis of the conditional pattern for that item. Each conditional pattern base gives rise to a conditional pattern tree. Subsequently, a recursive application of the procedure is conducted on this conditional pattern tree. This process is repeated for every item in the table, resulting in a list of frequently occurring items. The FP-Growth Algorithm employs a divide-and-conquer strategy to partition the main workload (Han et al., 2004).

3.4.2 Entropy Weight Method

Rudolph Clausius initially introduced the idea of entropy to quantify the level of disorder and unpredictability in a system. Claudie E. Shannon later applied this concept to information theory, where entropy is now used to measure the level of uncertainty in information based on probability theory. The core principle of the entropy approach is that information is derived from the degree of clustering within the data.

The Entropy Weight method is beneficial for determining the weights of criteria because it provides an objective approach to decision-making problems, distinct from relying on subjective information. The Entropy Weight method was used to measure the useful information provided by the available data. It computes the weights of criteria by evaluating the contrast density, which represents the disparity in values among alternatives for each criterion. A higher contrast indicates that the relevant criterion provides more information. Conversely, if all alternatives produce similar outputs, the criterion is considered to have a minimal impact on the decision-making process. In cases where all outputs are identical, the criterion can be entirely disregarded. The Entropy weight method is a suitable option for decision-making with multiple criteria, owing to its ability to objectively calculate the importance weights of the criteria. This eliminates the need for expert opinion and personal judgment. The entropy method steps are outlined as follows.

1. The decision matrix is normalized to exclude any inconsistencies that may arise from variations in the evaluation entities. The initial decision matrix serves as the basis (Eq.1) (Zhu et al., 2020).

$$P_{ij} = \frac{X_{ij}}{\sum_{i=1}^{m} X_{ij}} \quad \text{(Eq. 1)}$$

2. The equations provided below are used to determine the entropy value for each criterion (Eq.2-3).

$$e_j = -k \sum_{j=1}^{n} p_{ij} \ln(p_{ij}) \quad \text{(Eq. 2)}$$

$$k = (\ln(m))^{-1} \quad \text{(Eq. 3)}$$

3. The equation given below is used to calculate the deviation values for each criterion (Eq.4).

$$d_j = 1 - e_j \quad \text{(Eq. 4)}$$

4. The equation for calculating the weight value used to weigh the deviation values (Eq.5).

$$w_j = \frac{1 - e_j}{\sum_{p=1}^{n} (1 - e_p)} \quad \text{(Eq. 5)}$$

3.4.3 TOPSIS Method

The TOPSIS method, introduced by Hwang and Yoon (1981), addresses MCDM problems by employing a multidimensional weighting technique alongside ideal points to establish the preference ranking for alternatives. It stands out for its comprehensiveness compared to similar approaches, as it considers proximity to both positive and negative ideal solutions. TOPSIS recognizes that each criterion can either enhance or diminish its utility, making it useful for identifying both the most favorable and the least desirable solutions. The positive ideal solution seeks to maximize utility criteria while minimizing cost criteria. In contrast, the negative ideal solution aims to maximize cost criteria while minimizing utility criteria (Cocis et al., 2021).

4. Results

4.1 Patent Data Analysis with FP-Growth

The FP-Growth algorithm is employed to analyze the frequency and interconnectedness of 280,809 CPC codes across all patents in the raw dataset. Our approach involved implementing a Python code snippet utilizing FP-Growth and Panda libraries within our data analysis software. After completing the data cleaning, the column containing CPC codes in the dataset table undergoes analysis using the FP-Growth algorithm. This analysis aims to assess the frequency of occurrence and confidence frequencies associated with the CPC codes present in the dataset.

The "class symbol" is completed by a two-digit number, followed by the "subclass" defined in the next letter. The "group" number is then presented with a range of one to three digits. Each phrase in CPC codes is denoted by a symbol, such as "Y02A40/25", which stands for "GREENHOUSE TECHNOLOGIES.” The first letter is called a "section sign.” The next step involves assigning a numeric code to identify the "major group" or "subgroup". The code consists of at least two digits and is preceded by a slash. The patent examiner uses this code to classify the patent application, or any other relevant document based on its content.

4.2 CPC Code Frequency and Co-Occurrence Patterns

CPC codes serve as valuable identifiers for patents due to their detailed hierarchical structure. However, employing complete CPC codes can complicate the establishment of relevant relationships in relationality studies. To mitigate this challenge, the CPC codes gathered in this analysis were redefined at the group level, disregarding the subgroups. *Figure 3* illustrates an example of how a CPC code can be partitioned in this manner.

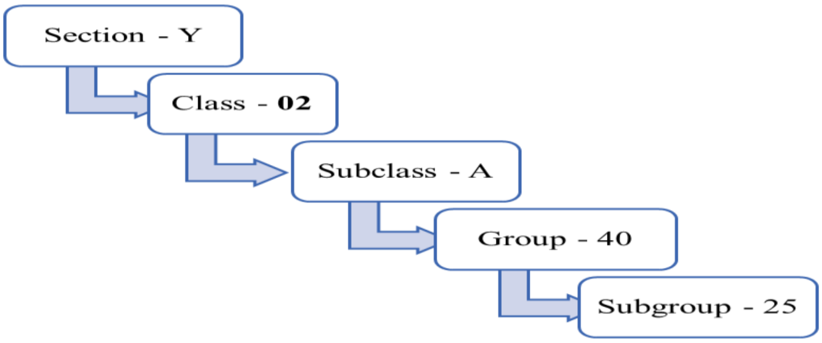

Figure 3. CPC Segmentation of a Sample CPC Code (Y02A40/25 for Greenhouse Technologies)

After refining the CPC granularity, the FP-Growth algorithm was applied. *Table 1* presents the CPC group codes with a support level greater than 0.02, as identified through the FP-Growth analysis.

Table 1. Frequency of Occurrence of CPC Codes

Rank	CPC Codes	Occurrence Frequency	Rank	CPC Codes	Occurrence Frequency
1	A01G9	0,751353	10	G06Q50	0,029285
2	Y02P60	0,397487	11	A01G22	0,028608
3	A01G7	0,128673	12	B32B27	0,028093
4	A01G31	0,087758	13	A01G27	0,027738
5	Y02E10	0,058344	14	A01G25	0,026095
6	A01G13	0,052126	15	C08J5	0,023518
7	Y02A20	0,050387	16	E04D13	0,023293
8	Y02B10	0,032152	17	E04B1	0,020651
9	A01C23	0,029671			

Table 1 illustrates that two distinct CPC codes, namely "Cultivation in receptacles, forcing-frames, or greenhouses" and "Technologies relating to agriculture, livestock, or agroalimentary industries," exhibit a higher frequency of utilization in the patents within the dataset compared to other codes.

When it comes to choosing the right technology, it may seem that "A01G9", "Y02P60", and "A01G7" are the most popular options based on their appearance in at least 10% of the patents. Relying solely on the frequency of CPC codes may not offer a comprehensive understanding of the ideal technology for selection. It's crucial to consider the combined impact of multiple technological fields rather than examining them in isolation. Researchers can achieve this by utilizing the FP-Growth Algorithm, which aids in evaluating the association of values in a dataset across various groups. This algorithm also helps in identifying the technologies used together in a patent dataset.

The information presented in *Table 2* illustrates the frequency of CPC code pairings along with the confidence level of their co-occurrence in patents. Additionally, it signifies the strength of the relationship between the codes by indicating their rates of co-occurrence and separate occurrence.

4.3 Multi-Criteria Decision Making

4.3.1 Defining Evaluation Criteria

The significance of a technology cannot be evaluated solely based on its frequency of use, despite the search for associations and relationships in the patent data. This kind of predictive information may simply confirm what is already known. Therefore, it is crucial to classify and identify valuable information in patent data according to specific criteria before selecting.

Table 2. Results of FP-Growth Algorithm on CPC Code Pairs

Antecedents	Consequents	Antecedent support	Consequent support	Support	Confidence
A01G9	Y02P60	0.7514	0.3975	0.2031	0.2703
Y02P60	A01G9	0.3975	0.7514	0.2031	0.5109
A01G7	A01G9	0.1287	0.7514	0.1244	0.9664
A01G9	A01G7	0.7514	0.1287	0.1244	0.1655
A01G31	A01G9	0.0878	0.7514	0.0834	0.9501
A01G9	A01G31	0.7514	0.0878	0.0834	0.1110
A01G31	Y02P60	0.0878	0.3975	0.0797	0.9082
Y02P60	A01G31	0.3975	0.0878	0.0797	0.2005
A01G7	Y02P60	0.1287	0.3975	0.0659	0.5123
Y02P60	A01G7	0.3975	0.1287	0.0659	0.1658
Y02E10	Y02P60	0.0583	0.3975	0.0502	0.8597
Y02P60	Y02E10	0.3975	0.0583	0.0502	0.1262
A01G13	A01G9	0.0521	0.7514	0.0500	0.9592
A01G9	A01G13	0.7514	0.0521	0.0500	0.0665
A01G31	A01G7	0.0878	0.1287	0.0355	0.4042
A01G7	A01G31	0.1287	0.0878	0.0355	0.2757
A01G9	Y02E10	0.7514	0.0583	0.0338	0.0449
Y02E10	A01G9	0.0583	0.7514	0.0338	0.5787
A01C23	A01G9	0.0297	0.7514	0.0292	0.9826
A01G9	A01C23	0.7514	0.0297	0.0292	0.0388
A01G9	G06Q50	0.7514	0.0293	0.0286	0.0381
G06Q50	A01G9	0.0293	0.7514	0.0286	0.9769
A01G9	B32B27	0.7514	0.0281	0.0276	0.0368
B32B27	A01G9	0.0281	0.7514	0.0276	0.9839
Y02B10	Y02P60	0.0322	0.3975	0.0273	0.8477
Y02P60	Y02B10	0.3975	0.0322	0.0273	0.0686
A01G22	A01G9	0.0286	0.7514	0.0270	0.9426
A01G27	A01G9	0.0277	0.7514	0.0270	0.9721
A01G9	A01G22	0.7514	0.0286	0.0270	0.0359
A01G9	A01G27	0.7514	0.0277	0.0270	0.0359
Y02B10	Y02E10	0.0322	0.0583	0.0256	0.7976
Y02E10	Y02B10	0.0583	0.0322	0.0256	0.4395
A01G25	A01G9	0.0261	0.7514	0.0240	0.9185
A01G9	A01G25	0.7514	0.0261	0.0240	0.0319
A01G9	E04D13	0.7514	0.0233	0.0229	0.0305
E04D13	A01G9	0.0233	0.7514	0.0229	0.9848
A01G9	C08J5	0.7514	0.0235	0.0226	0.0300
C08J5	A01G9	0.0235	0.7514	0.0226	0.9589

In the context of this research, five criteria were established by analyzing valuable data within the patent information to create the decision table:

I. Support value: The frequency of CPC pairs in the total data set: These criteria will be determined based on the values obtained using the FP-Growth algorithm, and the data presented in *Table 2.*

II. Average patent age: The importance of innovation in technology studies should be taken into consideration. In this context, the average age values of patents associated with a CPC code pair will also offer important insights into the novelty and currency of those patent applications. To calculate the values for this criterion, the ages of the patents with the relevant CPC code pair will be summed, divided by the total number of occurrences, and then averaged.

III and IV Average back citation value and average forward citation value: As in all technical studies, citations highlight the value of a patent and the significance of its technology. They are considered one of the important factors when assessing the importance of a patent. To calculate the values within the scope of these criteria, the forward and backward citation values of patents with the relevant CPC code pair will be summed separately, divided by the total number of occurrences, and then averaged.

Table 3. List of Applicants with The Highest Number of Patents in The Data (Top 10)

No	Institution Name[1]	Number of Patents[2]
1	GROW SOLUTIONS TECH LLC	455
2	EASTMAN CHEM CO	179
3	OCADO INNOVATION LTD	99
4	HAYGROVE LTD	76
5	UNIV CHINA AGRICULTURAL	75
6	REPUBLIC KOREA MAN RURAL DEV	69
7	ASAHI GLASS CO LTD	67
8	UNIV SHENYANG AGRICULTURAL	64
9	SVENSSON LUDVIG AB	62
10	LG ELECTRONICS INC	55

1. The organizations' names are transcribed precisely as they appear in the patent documents.
2. The number of patents value represents the number of patents with the CPC code "Y02A40/25" (which represents "GREENHOUSE TECHNOLOGIES") of the listed companies

Table 4. Raw Decision Table

CPC Code Pairings	Frequency of Occurrence	Average Patent Age	Average Power of Company	Average Number of Patents Cited	Average Number of Patents Cited by
A01G9:Y02P60	0.2031	4.8652	38.7498	4.8087	0.3544
A01G9:A01G7	0.1244	3.9699	37.1974	4.7632	0.3896
A01G9:A01G31	0.0834	4.1878	87.9494	5.0193	0.4097
Y02P60:A01G31	0.0797	4.1722	87.7726	4.9903	0.4155
Y02P60:A01G7	0.0659	4.1452	65.1972	5.0582	0.4481
Y02P60:Y02E10	0.0502	5.3513	6.0057	3.5055	0.3047
A01G9:A01G13	0.0500	5.4884	13.4070	4.0670	0.2218
A01G7:A01G31	0.0355	3.7866	114.4184	4.6948	0.4360
A01G9:Y02E10	0.0338	5.6918	6.7526	5.3950	0.3542
A01G9:A01C23	0.0292	2.9249	15.6075	3.6773	0.3396
A01G9:G06Q50	0.0286	2.8953	135.0447	2.5383	0.2908
A01G9:B32B27	0.0276	8.0303	45.5698	14.1131	0.2520
Y02P60:Y02B10	0.0273	5.9243	5.8680	3.5626	0.3365
A01G9:A01G22	0.0270	3.8029	12.2338	4.1768	0.3429
Y02E10:Y02B10	0.0256	5.9749	5.8727	3.7374	0.3345
A01G9:A01G25	0.0240	4.5121	59.9765	4.3011	0.3202
A01G9:E04D13	0.0229	8.1433	48.4002	16.6854	0.3246
A01G9:C08J5	0.0226	8.2057	51.1258	16.4129	0.2531

V. The average number of patents held by leading firms in a given field, referred to as the average firm strength: This value, regarded as an innovative criterion, is based on the principle that the competence and strength of the applicant company are crucial in determining the value of a patent. To calculate values under this criterion, we initially used a software code to determine the total number of patents held by each applicant in the field of Greenhouse Technology within a specified range. These patent numbers served as a numerical indicator for the respective applicants. Then, the total number of patent applications in the relevant field for the applicants with the corresponding CPC code pair, as shown in *Table 3*, will be summed up, divided by the total number of occurrences, and averaged. With this method, the aim is to consider the strength of the applicant company in the field within the analysis.

The values were computed according to the criteria definitions and organized into a decision table, presented in *Table 4*.

4.3.2 Entropy Weighting and Criterion Significance

The original decision matrix was normalized to eliminate irregularities due to different evaluation entities and ratios. Next, entropy, deviation, and weight values were calculated for each criterion and recorded in *Table 5.*

Table 5. Entropy Weight Values

	Frequency of Occurrence	Average Patent Age	Average Power of Company	Average Number of Patents Cited	Average Number of Patents Cited by
Entropy	0.90906	0.98449	0.87835	0.93697	0.99428
Deviation value	0.09094	0.01551	0.12165	0.06303	0.00572
Weight Value	0.30636	0.05223	0.40982	0.21232	0.01926

Table 5 provides the entropy, deviation, and weight values for five evaluation criteria. The entropy values reflect the degree of disorder or uniformity in each criterion's data distribution. Criteria with lower entropy (e.g., Average Power of Company) offer more discriminative information, resulting in higher deviation values. Consequently, the most influential criterion in the decision-making process is Average Power of Company, with a weight of 0.40982, followed by Frequency of Occurrence and Average Number of Patents Cited. In contrast, Average Patent Age and Average Number of Patents Cited by contribute less due to their high entropy and low deviation values. These weight values were later applied in the TOPSIS analysis to ensure an objective, data-driven ranking of technology alternatives.

4.3.3 TOPSIS-Based Ranking of Technology Alternatives

Following the determination of weight values, TOPSIS was applied. In the TOPSIS analysis, initially, the normalized decision matrix was formulated. Next, the weighted normalized decision matrix was derived, as shown in *Table 6*, using the weighting values obtained from the previous study, which are detailed in *Table 5*.

Table 6. Entropy Weighted Normalized Matrix

CPC Code Pairings	Frequency of Occurrence	Average Patent Age	Average Power of Company	Average Number of Patents Cited	Average Number of Patents Cited by
A01G9:Y02P60	0.2084	0.0108	0.0620	0.0314	0.0045
A01G9:A01G7	0.1276	0.0088	0.0595	0.0311	0.0049
A01G9:A01G31	0.0856	0.0093	0.1406	0.0327	0.0052
Y02P60:A01G31	0.0818	0.0093	0.1404	0.0325	0.0052
Y02P60:A01G7	0.0676	0.0092	0.1043	0.0330	0.0057
Y02P60:Y02E10	0.0515	0.0119	0.0096	0.0229	0.0038
A01G9:A01G13	0.0513	0.0122	0.0214	0.0265	0.0028
A01G7:A01G31	0.0364	0.0084	0.1830	0.0306	0.0055
A01G9:Y02E10	0.0346	0.0127	0.0108	0.0352	0.0045
A01G9:A01C23	0.0299	0.0065	0.0250	0.0240	0.0043
A01G9:G06Q50	0.0294	0.0065	0.2160	0.0165	0.0037
A01G9:B32B27	0.0284	0.0179	0.0729	0.0920	0.0032
Y02P60:Y02B10	0.0280	0.0132	0.0094	0.0232	0.0042
A01G9:A01G22	0.0277	0.0085	0.0196	0.0272	0.0043
A01G22:A01G9	0.0277	0.0085	0.0196	0.0272	0.0043
Y02E10:Y02B10	0.0263	0.0133	0.0094	0.0244	0.0042
A01G9:A01G25	0.0246	0.0101	0.0959	0.0280	0.0040
A01G9:E04D13	0.0235	0.0181	0.0774	0.1088	0.0041
A01G9:C08J5	0.0231	0.0183	0.0818	0.1070	0.0032
A01G9:A01G2009	0.2084	0.0108	0.0620	0.0314	0.0045

Using the data from the normalized weighted decision matrix, the optimal and least favorable options for each criterion were identified and recorded in *Table 7*.

Table 7. Positive and Negative Ideal Solutions

Positive and Negative Ideal Solutions	Frequency of Occurrence	Average Patent Age	Average Power of Company	Average Number of Patents Cited	Average Number of Patents Cited by
V+	0.2084	0.0065	0.2160	0.1088	0.0057
V-	0.0221	0.0183	0.0094	0.0165	0.0028

During the final phase of TOPSIS, the ideal solution values were computed, and based on these values, the preference options were ranked, as depicted in Table 8.

Table 8. Ranking of Alternatives

Rank	CPC Group Codes	Ideal Solution Value
1	A01G9:Y02P60	0.52976
2	A01G9:G06Q50	0.50688
3	A01G7:A01G31	0.47715
4	A01G9:A01G31	0.47418
5	Y02P60:A01G31	0.46635
6	A01G9:A01G7	0.38011
7	Y02P60:A01G7	0.35406
8	A01G9:C08J5	0.33587
9	A01G9:E04D13	0.33132
10	A01G9:B32B27	0.29973

Upon examining the ranked *Table 8* after conducting the decision analysis using the TOPSIS method, a significant disparity is evident between the rankings obtained from FP-Growth. The CPC code pair "A01G9- Cultivation in receptacles, forcing-frames or greenhouses" and "G06Q50-Systems or methods specially adapted for specific business sectors", which was previously ranked 11th, stands out as the most ideal solution as revealed by TOPSIS. However, the other 3 solutions listed in *Table* with an ideal solution value above 0.45 are also close to the ideal solution.

Table 9. Ideal Solutions with Value Bigger Than 0.45

No	CPC Group Codes	CPC Definitions	Ideal Solution Value
1	Y02P60	Technologies relating to agriculture, livestock or agroalimentary industries	0.5298
	A01G9	Cultivation in receptacles, forcing-frames or greenhouses	
2	A01G9	Cultivation in receptacles, forcing-frames or greenhouses	0.5069
	G06Q50	Systems or methods specially adapted for specific business sectors	
3	A01G7	Botany, treatment of plants	0.4772
	A01G31	Soilless cultivation, e.g. Hydroponics	
4	A01G9	Cultivation in receptacles, forcing-frames or greenhouses	0.4742
	A01G31	Soilless cultivation, e.g. Hydroponics	
5	Y02P60	Technologies relating to agriculture, livestock or agroalimentary industries	0.4664
	A01G31	Soilless cultivation, e.g. Hydroponics	

Upon analyzing these results, it becomes evident that Greenhouse technology has developed alongside the technology areas of Cultivation (A01G9), Botany and plant processing (A01G7), Hydroponics (A01G31), Agriculture-related technologies (Y02P60), and notably, systems or methods adapted for specific business sectors (G06Q50). The term systems or methods are used to describe information systems and related algorithms according to the patent methodology. Agriculture-related technologies also refer to developments in electronic and information system infrastructures in agriculture.

Moreover, focusing on the code Y02P60, it is notable that the most prevalent sub-group within this CPC code group in the dataset is Y02P60/14. Efforts to improve energy saving in greenhouses emphasize that both completed and ongoing innovation projects in this sector are centered around enhancing energy efficiency. In essence, the CPC code G06Q50, positioned 10th in the ranking of unique CPC occurrence frequency and appearing in less than 3% of existing patents, suggests that prominent companies in Greenhouse technology predominantly prioritize software applications within Greenhouse technologies. This inference is supported by the substantial weight assigned to the company strength criterion in the outcomes of the entropy weighting method.

5. Conclusions

This study developed a novel framework for technology selection by integrating patent analytics with MCDM. Unlike market data, which often reflect changes only after commercialization, patents provide the earliest and most comprehensive signals of technological progress. By leveraging FP-Growth to detect co-occurrence patterns among CPC codes and applying Entropy-weighted TOPSIS to evaluate their relative significance, this study demonstrated how patent data can be transformed into a structured decision-support framework.

The findings revealed several CPC codes as particularly influential in shaping the trajectory of greenhouse technologies. A01G9 (greenhouse cultivation) represents the foundation of protected agriculture, encompassing structural design, environmental control, and cultivation in controlled environments. Its consistent presence underscores the enduring importance of core greenhouse systems as the backbone of agricultural modernization. Alongside this, A01G31 (soilless cultivation) points to the increasing adoption of hydroponics, aeroponics, and aquaponics, which enable resource-efficient production and support the expansion of urban agriculture. A01G7 (plant treatment and processing) highlights the biological dimension of innovation, including plant health, growth optimization, and biotechnology applications within controlled-environment systems.

Beyond traditional cultivation, two additional CPC domains reveal the broader transformation of greenhouse technologies. Y02P60 (climate change mitigation in agriculture, forestry, and livestock) illustrates the alignment of agricultural innovation with global sustainability goals, particularly the European Green Deal and climate-smart agriculture initiatives. Patents classified under this category emphasize energy-efficient designs, carbon reduction, and adaptive systems for climate resilience. At the same time, G06Q50 (business-sector-specific information systems) demonstrates the digitalization of agriculture, capturing patents related to farm management software, logistics optimization, and data-driven decision-support systems. The prominence of this class confirms that future competitiveness will not only rely on physical cultivation infrastructure but also on the integration of artificial intelligence, IoT, and big data into agricultural management.

Taken together, these findings indicate that companies seeking to strengthen their position in the greenhouse technology sector should not only advance core cultivation techniques but also invest in sustainability-oriented technologies and digital infrastructures. This conclusion is consistent with broader international analyses, such as Cano et al. (2025), who identified similar growth trajectories in U.S. agricultural patents, particularly in digital farming and sustainable production domains. The convergence of biological, structural, and digital innovations suggests that future competitiveness will depend on integrated solutions that combine resource-efficient cultivation with climate resilience and data-driven management.

From a scholarly perspective, this study contributes to the literature by showing that CPC-based patent analytics, when combined with advanced data-mining and MCDM techniques, provide a robust and scalable method for technology forecasting. The framework presented here can assist R&D managers, investors, and policymakers in identifying technological opportunities, minimizing risks, and developing evidence-based roadmaps.

Future research could expand the scope of this methodology by incorporating semantic analysis of patent claims, integrating patent family data to capture international strategies, and conducting longitudinal studies to analyze how technological trajectories evolve over time. In doing so, the approach may evolve into a comprehensive decision-support system capable of guiding innovation management not only in agriculture but also across other technology-intensive sectors.

Ethical Statement

There is no need to obtain permission from the ethics committee for this study.

Conflicts of Interest

We declare that there is no conflict of interest between us as the article authors.

Authorship Contribution Statement

Concept: Kılıç, A.; Design: Eren, H.; Data Collection or Processing: Göçen, U.; Statistical Analyses: Göçen, U.; Literature Search: Kılıç, A., Eren, H., Göçen, U.; Writing, Review and Editing: Kılıç, A., Eren, H., Göçen, U.

References

Abbas, A., Zhang, L. and Khan, S.U. (2014). A literature review on the state-of-the-art in patent analysis, *World Patent Information*, 37: 3-13.

Al-Chalabi, M. (2015). Overview of hydroponic and aeroponic systems for urban agriculture. *Journal of Agricultural Science*, 7(5): 1–10.

Altuntaş, S. and Sezer, M. (2021). A novel technology intelligence tool based on utility mining. *IEEE Transactions on Engineering Management*, 70(7): 2480-2492.

An, J., Kim, K., Mortara, L. and Lee, S. (2018). Deriving technology intelligence from patents: Preposition-based semantic analysis. *Journal of Informetrics*, 12(1): 217-236.

Aznar-Sánchez, J. A., Velasco-Muñoz, J. F., López-Felices, B. and Román-Sánchez, I. M. (2020). An analysis of global research trends on greenhouse technology: Towards a sustainable agriculture. *International Journal of Environmental Research and Public Health*, 17(2): 4.

Cano, P. B., Carcedo, A. J. P., Hernández, C. M. and García, C. M. (2025). Trends in agricultural technology: A review of US patents. *Precision Agriculture,* 26(59): 1-17.

Cascini, G. and Zini, M. (2008). Measuring patent similarity by comparing inventions functional trees. *IFIP International Federation for Information Processing*, 277: 31–42.

Chan, F. T. S., Chan, M. H. and Tang, N. K. H. (2000). Evaluation methodologies for technology selection. *Journal of Materials Processing Technology*, 107(1–3): 330–337.

Choi, Y., Park, S. and Lee, S. (2021). Identifying emerging technologies to envision a future innovation ecosystem: A machine learning approach to patent data. *Scientometrics,* 126(7):5431-5476.

Cocis, A.-D., Batrancea, L. and Tulai, H. (2021). The link between corporate reputation and financial performance and equilibrium within the airline industry. *Mathematics*, 9(17): 2150.

Crosby, M. (2000). Patents, innovation and growth. *Economic Record,* 76(234): 255-262.

Çinkılıç, L., Varış, S. and Kubaş, A. (2014). Greenhouse vegetable growing and its problems in Thrace Regio. *Journal of Tekirdag Agricultural Faculty,* 11(2): 1-10. (In Turkish)

Ernst, H. (1997a). The use of patent data for technological forecasting: the diffusion of CNC-technology in the machine tool industry. *Small Business Economics*, 9(4): 361-381.

Ernst, H. (1997b). The Patent Portfolio for Strategic R&D Planning. Innovation in Technology Management - The Key to Global Leadership, *PICMET 1997: Portland International Conference on Management and Technology*, 31 July, P. 491-496, Portland, OR, USA.

Ernst, H. (2003). Patent information for strategic technology management. *World Patent Information*, 25(3): 233-242.

European Patent Office. Legal event data. https://www.epo.org/en/searching-for-patents/helpful-resources/first-time-here/legal-event-data (Accessed Date: 13.05.2025).

European Patent Office. (2022). EPO worldwide legal event data (INPADOC). Bulk Data Sets. https://www.epo.org/en/searching-for-patents/data/bulk-data-sets/inpadoc (Accessed Date: 13.05.2025).

Griliches, Z. (1998). Patent Statistics as Economic Indicators: A Survey. In: R&D and Productivity: The Econometric Evidence Z., Ed(s): Griliches, University of Chicago Press.

Han, J., Pei, J. and Yin, Y. (2004). Mining frequent patterns without candidate generation: A frequent-pattern tree approach. *Data Mining and Knowledge Discovery,* 8(1): 53-87.

Hwang, C.L. and Yoon, K. (1981). Methods for Multiple Attribute Decision Making–Methods and Applications: A State-of-the-art Survey. Springer, Berlin/Heidelberg, Germany.

Jun, S. (2011). IPC Code Analysis of Patent Documents Using Association Rules and Maps – Patent Analysis of Database Technology. *International Conference on Bio-Science and Bio-Technology (ICBB).* 11-12 October, P. 21-30, Yogyakarta, Indonesia.

Kim, G. and Bae, J. (2017). A novel approach to forecast promising technology through patent analysis. *Technological Forecasting and Social Change*, 117: 228-237.

Kim, Y. G., Suh, J. H. and Park, S. C. (2008). Visualization of patent analysis for emerging technology. *Expert Systems with Applications*, 34(3): 1804-1812.

Kim, G. J., Park, S. S. and Jang, D. S. (2015). Technology forecasting using topic-based patent analysis. *Journal of Scientific and Industrial Research*, 74(5): 265-270.

Lee, S. and Park, Y. (2005). Customization of technology roadmaps according to roadmapping purposes: Overall process and detailed modules. *Technological Forecasting and Social Change*, 72(5): 567-583.

Lee, S., Yoon, B., Lee, C. and Park, J. (2009a). Business planning based on technological capabilities: Patent analysis for technology-driven roadmapping. *Technological Forecasting and Social Change*, 76(6): 769-786.

Lee, S., Yoon, B. and Park, Y. (2009b). An approach to discovering new technology opportunities: Keyword-based patent map approach. *Technovation*, 29(6–7): 481-497.

Lingua, D. G. (2005). INPADOC: 30 years of endeavours yet unmapped territories remain. *World Patent Information*, 27(2): 105-111.

Mohammadian, A., Dahooie, J. H. and Qorbani, A. R. (2020). Prioritizing the applications of internet of things in the agriculture by using sustainable development indicators. *Iranian Journal of Agricultural Economics and Development Research*, 51(4):745-759.

Oztaysi, B. (2014). A decision model for information technology selection using AHP integrated TOPSIS-Grey: The case of content management systems. *Knowledge-Based Systems*, 70: 44-54.

Paci, R., Sassu, A. and Usai, S. (1997). International patenting and national technological specialization. *Technovation*, 17(1): 25-38.

Park, I., Jeong, Y., Yoon, B. and Mortara, L. (2014). Exploring potential R&D collaboration partners through patent analysis based on bibliographic coupling and latent semantic analysis. *Technology Analysis & Strategic Management,* 27(7): 759-781.

Park, S., Lee, S. J. and Jun, S. (2015). A network analysis model for selecting sustainable technology. *Sustainability,* 7(10): 13126-13141.

Pérez-Alonso, J., García-Martínez, A., and López, M. (2020). Energy efficiency and resource management in modern greenhouses. *Renewable Energy*, 148: 1116-1127.

Resh, H. M. (2022). Hydroponic Food Production: A Definitive Guidebook for the Advanced Home Gardener and the Commercial Hydroponic Grower. CRC Press, Boca Raton, USA.

Salvadó, L.L., Villeneuve, E., Masson, D., Abi Akle, A. and Bur, N. (2022). Decision Support System for technology selection based on multi-criteria ranking: Application to NZEB refurbishment. Building and Environment, 212.

Shen, Y.-C., Chang, S.-H., Lin, G. T. R., and Yu, H.-C. (2010). A hybrid selection model for emerging technology. *Technological Forecasting and Social Change*, 77(1): 151-166.

Sivri, M. and Çanakcı, M. (2024). Determination of spraying properties of nozzle plates in greenhouse sprayers in use. *Journal of Tekirdag Agricultural Faculty,* 21(3): 648-665. (In Turkish).

Song, K. and Ran, C. (2023). Research on technology opportunity identification based on topic mining and patent evaluation: A case study of smart agriculture. *Library and Information Service,* 67(3): 61-71.

Torkkeli, M. and Tuominen, M. (2002). The contribution of technology selection to core competencies. *International Journal of Production Economics*, 77(3): 271-284.

Van Straten, G., van Willigenburg, L. and van Henten, E. (2019). Robotics and automation in greenhouse crop production. *Biosystems Engineering*, 187: 1-14.

Wang, Y. L. (2012). Research on technology selection for enterprises with tools of patent analysis. *International Conference on Management Science and Engineering-Annual Conference Proceedings*, 1: 1651-1657.

Yoon, B., Phaal, R. and Probert, D. (2008). Morphology analysis for technology roadmapping: Application of text mining. *R and D Management*, 38(1): 51-68.

Zhu, Y., Tian, D., and Yan, F. (2020). Effectiveness of entropy weight method in decision-making. *Mathematical Problems in Engineering,* 2020: 1-5.