THEORETICAL AND APPLIED APPROACHES IN ENGINEERING

Editor:
Dr. Mehmet Uzun

DUVAR

THEORETICAL AND APPLIED APPROACHES IN ENGINEERING

Editor

Dr. Mehmet Uzun

THEORETICAL AND APPLIED APPROACHES IN ENGINEERING

Editor: Dr. Mehmet Uzun

Editor in chief: Berkan Balpetek

Cover and Page Design: Duvar Design

Printing : May -2025

Publisher Certificate No: 49837

ISBN: 978-625-5885-45-6

853 Sokak No:13 P.10 Kemeraltı-Konak/İzmir

Tel: 0 232 484 88 68

www.duvaryayinlari.com

duvarkitabevi@gmail.com

TABLE OF CONTENTS

Chapter 1

Electrıc Scooter Project Homework

Ali Osman GENÇ[1], Ömer Faruk SEVİM[2], Halil İbrahim ERKAN[3], Mevlüt KAÇMAZ[4], Abdulrezzak BAKIŞ[5]

1. WHAT IS AN ELECTRIC SCOOTER?

Electric scooters are electrically powered vehicles designed especially to meet short-distance transportation needs. These vehicles have become very popular in urban transportation in recent years thanks to their environmentally friendly features, ease of use and fast transportation. In addition, in many cities, electric scooters are also widely used through rental systems according to the needs of users, thus being preferred as a practical and efficient means of transportation as an alternative to public transportation systems.

1.1 Working Principle of Electric Scooters

Electric scooters basically consist of a few main components:

The motor: This is the main component that makes the electric scooter move and is usually integrated into the wheels. It runs on electric power.

Battery: The batteries, which are the energy source of the scooter, are usually produced with lithium ion technology and determine the range of the vehicle.

Control System: This electronic system manages the speed of the scooter, monitors the battery level and controls the functioning of the engine.

Brake System: Electric scooters can be equipped with both mechanical and electric braking systems. Electric brakes reduce speed by operating the motor in the opposite direction.

Wheels and Suspension: These components affect the ride comfort of the scooter, generally providing better handling and comfort with larger wheels.

[1]
[2]
[3]
[4]
[5] Assoc. Prof. Dr.

Steering and Pedals: It is used to control the direction of the scooter. Some models may also have pedals, but most models are controlled only by the steering wheel. (Binns, P., & Richards, C. 2020)

1.2 Types of Electric Scooters

Personal Electric Scooters: These types of scooters are lightweight vehicles that are generally designed for individual users and are very practical in terms of use. Ideal for activities such as daily commuting, short-distance travel and shopping, personal electric scooters are frequently preferred in busy city life thanks to their ease of transportation and compact structure.

Shared Electric Scooters: These are scooters that are commonly found in urban transportation and that users can rent from a certain point and drop off at another point. These scooters can usually be rented through smartphone apps and can be easily found in certain areas. These types of scooters are usually equipped with larger batteries and are designed for more intensive use, thus offering a practical solution as an alternative to urban public transportation systems.

Sport and High Performance Electric Scooters: These scooters have more powerful motors and larger batteries and are designed to cover long distances quickly. Sports and high-performance electric scooters, which can generally reach faster speeds and are suitable for long-term use, are built to withstand more challenging road conditions and are an ideal choice for users looking for performance. (Sun, Y., & Li, J. 2020).

1.3 Advantages of Electric Scooters

Environmentally Friendly: Electric scooters offer environmentally friendly transportation options by completely eliminating carbon emissions. These vehicles, which do not use fossil fuels, not only reduce air pollution, but also contribute significantly to noise pollution, thus enabling a cleaner and quieter city life.

Low Cost: The operating costs of electric scooters are considerably lower when compared to conventional vehicles. These electric-powered scooters save significantly by using electricity instead of fuel oil. In addition, maintenance and repair costs are also more economical as mechanical parts wear less and have a simpler structure.

Fast and Efficient: Electric scooters offer a fast and efficient transportation solution, especially in cities with heavy traffic. Compared to public transport or cars, they can move much faster over short distances, thus minimizing time loss and speeding up transportation processes.

Portability: Many electric scooters are highly portable thanks to their lightweight construction and foldable features. These features provide a great advantage for users who want to take public transportation or carry the scooter to reach another point. Thus, scooters make transportation within the city more flexible.

Ease of Use: Electric scooters are very simple to use. In most models, only the throttle is used to accelerate and steering is done with the steering wheel. These simple control mechanisms allow scooter users to use their vehicles without any difficulty. (URL-1, URL-2)

1.4 Disadvantages of Electric Scooters

Limited Range and Speed: The range of electric scooters varies according to the capacity of the battery used and is usually limited. Most models can cover a distance of 20-50 kilometers on a single charge. In addition, scooters' speeds usually range between 25-30 km/h, which can be a limiting factor for users who prefer to travel at higher speeds.

Battery Life and Charging Time: Batteries in electric scooters can lose capacity over time, depending on usage time and charging cycles. Also, batteries often take a long time to fully charge, which can be an obstacle for users who plan to travel long distances.

Road Safety and Infrastructure Deficiencies: Appropriate infrastructure is needed for the safe use of electric scooters. However, especially in cities with heavy traffic, existing road infrastructure is not sufficient for scooters, which can put the safety of users at risk. Scooters interacting with other vehicles in traffic can lead to accidents.

Theft and Vandalism: Shared electric scooter systems can often face issues such as theft and vandalism. As scooters are widely used, especially through rental systems, vulnerabilities can trigger such problems, which can negatively impact the safety and maintenance costs of the scooters.

Physical Challenges: Electric scooters can pose physical challenges for some users. They may not be particularly suitable for people with disabilities, and long-term use may result in discomfort such as back or knee pain. This is a limiting factor for scooters to be used comfortably by a wider range of users (URL-1, URL-2).

1.5 The Future of Electric Scooters

Electric scooters are rapidly becoming widespread and are becoming an important alternative, especially in urban transportation. With the increasing popularity of smart city projects and the growing interest in environmentally friendly transportation solutions, the use of scooters will become even more widespread in the future. Emerging battery technologies have great potential to extend the range of electric scooters and shorten charging times. Moreover, with the integration of autonomous driving technologies, scooters can become safer, more efficient and user-friendly.

In order to increase the use of electric scooters, city governments are introducing regulations and accelerating infrastructure investments for this mode of transportation. At the same time, scooter sharing services are expected to be more accessible and user-friendly. These developments will contribute to a more sustainable and efficient urban transportation model by integrating with public transportation systems. (URL-3)

1.6 Electric Scooter Usage Tips

Wear a Helmet and Protective Equipment: When riding an electric scooter, there is a risk of falling and crashing, especially when you reach high speeds. To reduce these risks, using protective equipment such as helmets, knee pads and elbow pads greatly increases personal safety.

Slow and Controlled Driving: Carefully controlling the speed of scooters is important for safe riding, especially in heavy traffic and crowded pedestrian areas. While it is tempting to go fast, controlled driving protects both the rider and bystanders.

Follow the Rules of the Road: When using an electric scooter, it is extremely important to follow local traffic rules and special regulations for scooters. This ensures safety and avoids legal problems.

Maintenance and Charging: Regular maintenance of your electric scooter and proper charging of the battery improves its performance and prolongs its life. It is also important to detect malfunctions early for the scooter to work properly.

In conclusion, electric scooters, besides being an environmentally friendly and practical means of transportation, require users to be careful for safety and trouble-free use. Acquiring safe riding habits ensures effective and long-lasting use of scooters. (URL-4)

2. INFORMATION ABOUT BICYCLE PATHS

Bicycle Network Planning is a process that aims to increase community mobility by providing interconnected routes and appropriate facilities, taking into account the needs of bicycle users. It aims to establish a seamless transport network for cyclists and to organize the most appropriate infrastructure improvements (Bach and Diepens, 2000; LTSA 2004).

Bikeways are specially designed and marked paths that allow bicycle users to travel more safely and comfortably. They are separated from motor vehicles and pedestrians and allow cyclists to travel safely and without interruption. It is of great importance to promote cycling as a means of transportation in cities, to reduce traffic congestion, to provide environmentally friendly transportation alternatives and to promote healthy living. The development of this infrastructure can make urban transportation more efficient, environmentally friendly and healthy (Götschi, T., de Nazelle, A., & Brand, C. 2016).

Bicycle lanes can be shaped according to different standards and designs. Depending on factors such as the size of cities, transportation policies and the priorities of local governments, there may be various differences in the configuration of bikeway systems. However, despite all these differences, the main purpose of bicycle lanes is to provide cyclists with a safe riding space, minimize their interactions with traffic and enable more efficient use of the bicycle as a means of transportation. To this end, the design of bicycle lanes in each city should be optimized to maximize the safety of users. (MoEU)

2.1 Types of Bicycle Lanes

1. Cycle Lane (Bicycle Track):

- This is a specially designed path that can only be used by cyclists. It is usually separated from sidewalks or motor vehicles and dedicated entirely to bicycles. A cycle track offers a space reserved exclusively for cyclists and not shared with pedestrians. These paths allow cyclists to travel safely without interacting with traffic.

2. Bicycle Lane:

- A bicycle lane is a dedicated lane, usually on the edge of a main roadway, separated from motor vehicles and other road users. This lane is usually delimited by lines or physical barriers so that cyclists can travel at a certain distance from motor vehicles. However, this lane is located on the same road, offering a space reserved exclusively for cyclists.

3. Shared Road:

- Shared roads are areas that are shared by both bicycles and motor vehicles. On such roads, cyclists, motor vehicles and pedestrians usually share the same roadway. Some shared paths may have a dedicated lane or pattern for cyclists, but the road is still a shared space for multiple users.

4. Separate Bikeway (with Physical Separators):

- This type of bicycle lane is separated from motor vehicle paths by physical barriers. By using concrete barriers, high sidewalks or other types of physical barriers, cyclists are completely isolated from motor vehicles. This design improves the safety of cyclists and provides a more comfortable riding experience.

5. Complex Roads (Multimodal):

- Multimodal roads in large cities serve both cyclists and public transport (buses, trams, etc.). These roads allow different modes of transportation to be used together, thus making transportation within the city more efficient. Multimodal roads allow for the coexistence of various modes of transportation, creating a more integrated and efficient system.

6. Interregional cycle path (Veloroute):

- Regional cycle paths are long-distance road networks, usually designed to facilitate transportation between cities. These roads are particularly appealing to cyclists who want to travel between cities. Veloroute offers comfortable and sustainable transportation to ensure safe cycling over longer distances (Pucher, J., & Buehler, R. 2008).

2.2 Bike Lane Design

The design of bicycle lanes plays a critical role in both safety and user experience. The key elements of an effective bikeway design are the following:

Road Width:

The width of the bike lane is critical to ensure that cyclists can travel safely and comfortably. One-way cycle tracks are usually 1.2 to 1.5 meters wide, while two-way cycle tracks can be 2.5 to 3 meters wide. This width ensures that cyclists traveling at different speeds can safely pass and navigate each other.

Surface Quality:

A smooth and non-slip surface is essential for the safety of cyclists. Bike lanes should be made of durable materials such as asphalt or concrete and should be free of cracks, holes or deterioration anywhere on the road. Otherwise, it can lead to cycling accidents and threaten the safety of cyclists.

Physical Separation:

The physical separation of the bicycle lane from carriageways is essential for the safety of cyclists. Separating the bicycle lane from vehicles allows cyclists to ride in a safer environment, especially when traveling at high speeds. There are various ways to achieve this separation, such as with barriers, fences or wide sidewalks.

Turns and Intersections:

Intersections, turns and crossing points on bikeways should be carefully designed. Intersections are areas where cyclists interact with motorized vehicles and are often more at risk. At this point, a design that visually warns cyclists and uses markings and special traffic signals will increase safety. For example, special traffic lights for cyclists can be used.

Lighting:

Adequate lighting should be provided for night use of bicycle lanes. Good lighting allows the driver to see ahead and other road users to recognize cyclists.

Lighting is an important factor that increases safety, especially during the hours of darkness.

Markings and Signals:

Bike lanes should have proper directional signs, speed limits and warning signs. In addition, special traffic signals for cyclists (e.g. green lights for bicycles) and on-street signage are also important. These markings ensure that both cyclists and other road users can share the road safely.

Barriers and Obstacles:

Bike lanes should be free of physical obstacles. Obstacles such as sidewalks or overhangs can make it difficult for cyclists to pass and can cause accidents. A clear and unobstructed space should be provided at each point of the cycle path that does not obstruct the ride. This ensures that users can travel more comfortably and safely (Pucher, J., & Buehler, R. 2012).

2.3 Benefits of Bike Lanes

1. Safety:

Bike lanes allow cyclists to travel more safely without interacting with motorized vehicles in traffic. They also reduce accidents and injuries.

2. Environmentally Friendly Transportation:

Cycling is a zero-emission means of transportation. The widespread use of bicycle lanes reduces carbon emissions from motor vehicle trips and promotes an environmentally friendly transportation alternative.

3. Health:

Cycling increases physical activity. They promote a healthy lifestyle, improve heart health and prevent overweight. In addition, cycle paths encourage people to spend time outdoors.(WHO)

4. Traffic Reduction:

Increasing the use of bicycles in heavy traffic in cities can alleviate motor vehicle traffic. Thanks to bicycle lanes, bicycles are preferred as an option for urban transportation.(TÜİK)

5. Economic Benefits:

The construction of bicycle lanes can provide economic benefits for cities in the long run. People traveling by bicycle reduce fuel costs, especially when they choose not to drive over short distances. (Aydın, M.)

2.4 Bikeway Infrastructure Deployment and Challenges

Many cities are investing in expanding their bicycle road networks. However, this process can face some challenges, especially in larger cities:

- Lack of Infrastructure: Especially in older cities, there may not be enough space for bicycle lanes. In this case, existing roads may need to be reconstructed or adapted to the existing infrastructure.

- Interaction with Vehicle Drivers: Careful design and supervision is required in areas where the bikeway interacts with motorized vehicles. Arrangements should be made to ensure that vehicle drivers do not violate bicycle lanes.

- Social and Cultural Factors: In places where bicycle culture is widespread, people use these roads more, while in societies that are not accustomed to bicycle use, this transition may take longer.(Aktaş, T)

As a result, bicycle lanes offer a safe, environmentally friendly and healthy alternative to urban transportation. However, an effective bikeway infrastructure requires careful planning, proper design and sustainable policy.

3. BICYCLE PATHS IN BATMAN

Table 3.1 Table showing bicycle lanes in Batman

STREET - STREET	**BICYCLE ROAD**		**EXPLANATION**
	TO BE	**NOT TO BE**	
BETWEEN BATMAN AIRPORT AND GÜLTEPE INTERSECTION		+	A BICYCLE PATH SHOULD BE BUILT FOR TRAFFIC SAFETY.
GÜLTEPE INTERSECTION - BATMAN PARK		+	A BICYCLE PATH SHOULD BE BUILT FOR TRAFFIC SAFETY.
BETWEEN BATMAN PARK AND DÖRTYOL	+		BECAUSE THERE IS NO OBSTACLE BETWEEN THE BICYCLE PATH AND THE SIDE STREET, THE BICYCLE PATH IS USED BY OTHER VEHICLES MANY TIMES. THIS ENDANGERS THE SAFETY OF THE BICYCLE PATH. THE BICYCLE PATH AND THE PASSAGE ROUTES OF OTHER VEHICLES SHOULD BE SEPARATED AS SOON AS POSSIBLE. PREVENTIVE MEASURES SHOULD BE TAKEN.

BETWEEN DÖRTYOL – BASIN JUNCTION		+	A BICYCLE PATH MUST BE BUILT FOR TRAFFIC SAFETY.
BETWEEN BASIN JUNCTION – BATMAN MUNICIPALITY		+	A BICYCLE PATH MUST BE BUILT FOR TRAFFIC SAFETY.
BETWEEN DORTYOL – HASANKEYF JUNCTION		+	A BICYCLE PATH MUST BE BUILT FOR TRAFFIC SAFETY.
SİİRT RING ROAD		+	A BICYCLE PATH MUST BE BUILT FOR TRAFFIC SAFETY.
BETWEEN SİİRT RING ROAD JUNCTION – HASANKEYF JUNCTION		+	A BICYCLE PATH MUST BE BUILT FOR TRAFFIC SAFETY.
UNIVERSITY RING ROAD		+	A BICYCLE PATH MUST BE BUILT FOR TRAFFIC SAFETY.
TOKİ 1 – TOKİ 2 BETWEEN	+		DUE TO THE BUSY ROAD, THE BICYCLE PATH MUST CONTINUE ALL THE WAY.
TOKİ 2 – UNIVERSITY BETWEEN		+	A BICYCLE PATH MUST BE BUILT FOR TRAFFIC SAFETY.
UNIVERSITY BATIRAMAN CAMPUS	+		THERE IS A BICYCLE PATH, BUT IT IS USED IN SHARED SHARES WITH OTHER VEHICLES AND SINCE THE ROAD IS EXCESSIVELY BUMPY, THE ROAD NEEDS TO BE REPARED AND THE BORDERS OF THE BICYCLE PATH NEED TO BE DETERMINED.

4. IMAGES OF BICYCLE PATHS IN BATMAN

Figure 4.1 Turgut Özal Boulevard

Figure 4.2 Turgut Özal Boulevard

Figure 4.3 Turgut Özal Boulevard

Figure 4.4 Turgut Özal Boulevard

Figure 4.5 Turgut Özal Boulevard

Figure 4.6 Kuyubasi Toki

Figure 4.7 Airport city center bird's eye view bicycle route

Figure 4.8 Turgut Özal Boulevard city center bird's eye view bicycle path

Figure 4.9 Batman University Batıraman Campus bird's eye view bicycle path


Figure 4.10 Electric scooters operating in Batman

5. NUMBER OF ELECTRIC SCOOTERS AND BICYCLES PASSING PER HOUR IN BATMAN

Table 5.1 5. Number of Electric Scooters and Bicycles Per Hour in Batman

STREET-BROAD	NUMBER OF BICYCLES/HOUR	NUMBER OF SCOOTERS/HOUR
TURGUT ÖZAL BOULEVARD	35	7
BATMAN UNIVERSITY BATIRAMAN CAMPUS	4	0
KUYUBAŞI TOKİ	13	0
BATMAN AIRPORT-CITY CENTER	6	4

6. CONCLUSION

In this study, the infrastructure of electric scooters and bicycle lanes in Batman province was examined and the current situation was analyzed. The main findings of the study can be summarized as follows:

1. Electric Scooter Usage: Electric scooters have potential as an environmentally friendly, economical and fast transportation alternative in Batman. However, the usage rate is low due to limited infrastructure and lack of awareness.

2. Bike Lane Infrastructure: The number of existing bicycle lanes in Batman is insufficient, some of the existing lanes are shared with other vehicles and there are insufficient safety measures. It is recommended to separate bicycle lanes with physical barriers and improve the infrastructure to increase traffic safety on the routes specified in the study.

3. Traffic Safety and Usage Data: Especially in busy areas such as Turgut Özal Boulevard, an average of 35 bicycles and 7 scooters per hour were observed. On the other hand, in less dense areas such as around Batman University, utilization is quite low. This data shows that usage can increase if the infrastructure is improved.

4. Infrastructure Needs: The study emphasized that in order to improve traffic safety in Batman, bicycle lanes should be expanded, existing roads should be improved, and these areas should be protected from the use of other vehicles. It is also recommended to increase regulatory signage and organize awareness campaigns for bicycle and scooter users.

7. REFERENCES

URL-1 www.hiscoot.com Accessed on 23.11.2024

URL-2 www.motoblog.com.tr Accessed on 23.11.2024

URL-3 www.motobox.com.tr Accessed on 23.11.2024

URL-4 www.binbin.tech Accessed 23.11.2024

REPUBLIC OF TURKEY MINISTRY OF ENVIRONMENT, URBANISM AND CLIMATE CHANGE, 2024, Bicycle Road Transportation Guide, Environmentally Friendly Transportation Alternatives

World Health Organization (WHO) 2023, Active Transport and Health

European Cyclists' Federation (ECF), 2022 Cycling Infrastructure and Urban Planning

Aydın, M. (2021). “Urban Bikeway Planning and Design”. Journal of Urbanism and Transportation

Pucher, J., & Buehler, R. (2020). Cycling for Sustainable Cities. MIT Press.

Turkish Statistical Institute (TurkStat). (2024). Transportation Statistics and Bicycle Use.

Aktaş, T. (2019). “Bicycle Lane Policies and Practices in Turkey”. Journal of Planning

Binns, P., & Richards, C. (2020). Electric Scooters: The Mechanics and Future of Micro-Mobility. Transportation Research Part C: Emerging Technologies

Sun, Y., & Li, J. (2020). Shared electric scooters: The future of urban micro-mobility. Transport Policy, 85

Götschi, T., de Nazelle, A., & Brand, C. (2016). Environmental health benefits of cycling: A review of evidence. European Journal of Public Health

Pucher, J., & Buehler, R. (2008). Bisiklet Sürmeyi Dayanılmaz Kılmak: Hollanda, Danimarka ve Almanya'dan Dersler. Ulaşım İncelemeleri

Pucher, J., & Buehler, R. (2012). Ulaşım için Bisiklet Kullanımının Teşvik Edilmesi: Araştırma İhtiyaçları ve Temel Sorular. Ulaşım İncelemeleri

Bach, R., & Diepens, J. (2000). Bisiklet ağlarının kentsel ulaşım sistemlerindeki rolü. Ulaşım Planlaması Dergisi

Chapter 2

AI Model Validation with Zero-Knowledge Proof: Trust and Transparency Without Data Access

Özge TAŞ[1]

Summary

Zero-Knowledge Proof Machine Learning (ZKML) is a new approach that combines zero-knowledge proofs (ZKP) with machine learning (ML) to develop privacy-focused and secure artificial intelligence systems. ZKP are cryptographic techniques that enable one party to prove the validity of certain information without disclosing any additional data. This mechanism is particularly important in fields that require high levels of privacy, such as finance, healthcare, and identity verification.

ZKML enables the cryptographic verification of the accuracy of model inferences or training processes without disclosing model parameters or user data. In the context of federated learning, the accuracy of each participant's contribution to the model training process can be verified using proof systems such as zk-SNARKs, thereby enabling a secure collaboration environment without the risk of data leakage. Similarly, during the inference phase, it can be verified whether the model produced a specific output, which builds trust in fields such as medicine and finance where sensitive decisions are made.

Currently, the production of ZK proofs requires high computing power. However, thanks to advances in hardware, distributed systems, and cryptography, proof production is now more feasible even for larger and more complex models. Startups like Modulus Labs and tools like the ezkl library enable the production of ZK proofs on models in ONNX format, offering practical solutions to

[1] Öğretim Görevlisi, Kapadokya Üniversitesi, Kapadokya Meslek Yüksek Okulu, Bilgisayar Teknolojileri Bölümü, Bilgisayar Programcılığı Pr, 50400, Nevşehir, Türkiye. ozge.tas@kapadokya.edu.tr
ORCID: https://orcid.org/0000-0001-7220-5054

developers. Systems like Plonky2 have reduced proof production for models with millions of parameters to just minutes.

ZKML has a bunch of use cases, including on-chain ML verification (e.g., in DeFi protocols), transparency of ML services (MLaaS), fraud detection, and private inference. For example, in decentralized Kaggle-like systems, the accuracy of a model can be proven without revealing its details. In healthcare, patients can access diagnostic results without disclosing their data.

In conclusion, ZKML combines privacy protection with the security of verification processes, enabling the development of more ethical and reliable artificial intelligence systems. This approach, which lies at the intersection of cryptography and machine learning disciplines, has the potential to increase the transparency and security of AI systems at both technical and societal levels.

Keywords: Machine Learning, Zero-Knowledge, Artificial Intelligence.

1.Introduction

The investigation of the security of AI language models (LLMs) using zero-knowledge proofs (ZKP) represents a critical intersection of artificial intelligence and cryptography, addressing growing concerns about data privacy and model integrity in an increasingly digital world. As machine learning models are integrated into various applications, they often handle sensitive information and raise important issues related to trust, data breaches, and the potential misuse of algorithms. Significant incidents such as the FlexBooker data breach, which affected millions of people, highlight the urgency of advanced security measures in AI systems and underscore the need for frameworks that ensure both performance transparency and privacy protection (Peng Z et al. 2025; Scherlis, B. 2024).

Zero-knowledge proofs, a cryptographic technique developed in the 1980s, enable one party to verify the truth of a statement without disclosing any additional information. This capability has profound implications for evaluating AI models and allows customers to verify the accuracy of outputs without disclosing sensitive data (Brundage, M. et al. 2018)(Lavin, R. et. 2024).

The adoption of ZKPs in AI is particularly important as the industry grapples with vulnerabilities such as data poisoning and instant injection that threaten the reliability and safety of AI systems. This research investigates how ZKPs can mitigate such risks by enhancing the security of LLMs while preserving the privacy of input data and ensuring the integrity of outputs.

Despite the promise of ZKPs, the investigation reveals persistent challenges such as implementation complexity, the balance between transparency and privacy, and the need for continuous adaptation to emerging threats.

For example, while ZKPs provide protection against data breaches, they may also introduce performance overhead that complicates real-time applications. Furthermore, the opaque nature of AI models can hide security vulnerabilities and require specialized strategies for detection and mitigation to ensure that robust security measures are in place (Yao, Yifan, et al 2024).

As AI continues to evolve, the integration of zero-knowledge proofs is positioned as a fundamental strategy to enhance the security and privacy of language models and promote trust in their deployment in sensitive applications such as healthcare, finance, and more. Ongoing research and development efforts will be important in improving these techniques, making them more accessible and effective for real-world applications, and emphasizing the importance of collaborative initiatives to establish comprehensive security frameworks across industries (Singh, S. 2024).

2. Conceptual Framework

The integration of machine learning (ML) models into various applications has raised significant concerns about the security and privacy of the data within these systems. Verifiable testing is a process designed to prove the true performance of an ML model and ensure that its claimed performance reflects its actual generalization ability rather than just its performance on the training data (Peng, Z. et al. 2025).

In this framework, an ML client sends a model and test data and specifies evaluation metrics such as accuracy or F1 score. The service provider runs the tests and creates a report using encryption techniques and third-party verification tools to maintain the integrity of the testing process and ensure the authenticity of the results (Peng, Z. Et.al. 2025).

Despite these security measures, the relationship between ML clients and service providers is fraught with trust issues. Customers are often reluctant to share sensitive data due to the risk of data breaches, as highlighted by incidents such as the 2022 FlexBooker breach, which exposed the personal information of approximately 3.2 million users, and a significant database breach at Alibaba Cloud that exposed the data of over one billion Chinese citizens (Peng, Z. et al. 2025). Additionally, there are concerns about the potential for service providers to use the least suitable models while presenting deceptively positive performance metrics (Peng, Z. et al., 2025).

The security landscape of modern AI becomes even more complex due to the integration of code and data within AI models. Unlike traditional software, which has a clear distinction between code and data, AI models can treat processed data as executable instructions, making them vulnerable to security flaws similar to code injection attacks in software security (Ahmet, M. 2025).

These weaknesses can lead to integrity risks, where an AI model's outputs can be manipulated or incorrect, privacy risks, where sensitive input data is exposed, and governance risks, which affect the wider application of AI technologies.

Recent developments in zero-knowledge proofs (ZKPs) offer a potential solution to these challenges. Introduced in the 1980s, ZKPs enable one party to prove the truth of a statement to another without disclosing any additional information. In the context of AI, ZKPs can enhance the transparency and verifiability of ML models by enabling customers to verify model execution without disclosing their data.

This capability is important for addressing privacy concerns associated with the sharing of sensitive data and can help mitigate risks associated with adversarial attacks targeting generative AI systems, which have become increasingly prevalent in recent years.

As the application of AI continues to expand, understanding these security dynamics and the role of ZKPs is critical to fostering trust and ensuring the secure deployment of AI technologies in sensitive environments.

3. Methodology

The investigation of the security of AI language models (LLMs) using zero-knowledge proofs (ZKP) involves a multifaceted approach that aims to address the inherent challenges of ensuring output integrity while preserving privacy. This methodology combines principles from cryptography, data verification, and AI system operational frameworks to develop robust mechanisms for protecting sensitive information. Methods are categorized according to these principles.

3.1.Yapay Zekada Sıfır Bilgi Kanıtları

ZKP serves as a fundamental element in verifying the accuracy of information without disclosing the underlying data. By applying ZKPs to outputs generated by AI, researchers can create zero-knowledge circuit demonstrations that enable the verification of outputs from models such as GPT-4 or text-to-image models without revealing their internal workings or the characteristics of the inputs used (Lavin, R. et al. 2024).

This feature is of critical importance in scenarios involving sensitive data, such as financial transactions or private voting systems, as it ensures both the integrity of the process and the preservation of privacy (Lavin, R. et al. 2024). Figure 1 illustrates the working principle of zero-knowledge proofs.

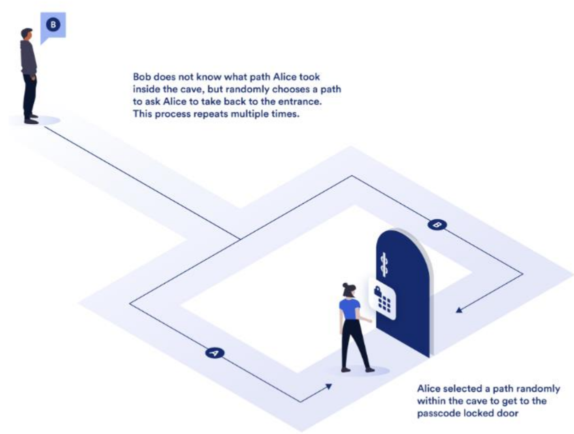


Figure 1. Conceptual example of how zero-knowledge proof works.

3.2. Collaborative Data Analysis

The methodology involves the use of collaborative zk-SNARKs for data analysis between institutions such as hospitals and research facilities, ensuring compliance with privacy regulations. This allows for the examination of aggregate data such as treatment effectiveness or disease trends without revealing individual patient details (Lavin, R. et al. 2024). Such applications highlight the importance of privacy protection in collaborative research environments where data security is paramount.

3.3. Addressing Model Security Vulnerabilities

To mitigate risks associated with LLMs, the methodology includes strategies to address various security threats. Key challenges include unlimited consumption attacks, where malicious users exploit the resource-intensive nature of LLMs, and sensitive information disclosure, where confidential data may inadvertently appear in model outputs (Scherlis, B., 2024).

Providing developers with comprehensive training on these security vulnerabilities and establishing robust data verification channels are important steps in this process.

3.4. Security Framework

A proposed security framework outlines the integration of ZKPs into AI system operational workflows. It emphasizes the necessity of cross-verification processes and human oversight to enhance the security of critical information. Furthermore, the framework advocates for automated verification mechanisms tailored for high-risk environments (Peng, Z. et al., 2025).

This layered approach enables AI systems to effectively manage sensitive data and provide verifiable outputs.

3.5. Evaluation of Effectiveness

The evaluation of the proposed methodologies involves assessing their effectiveness in addressing the unique challenges posed by AI language models. Metrics will focus on the scalability and flexibility of the framework in handling sparse and quantitative models, ensuring that ZKP-based approaches are not only secure but also efficient in real-world applications (Peng, Z. et al. 2025).

Experimental results will be used to demonstrate the effectiveness of the methodologies, particularly in optimizing the speed of evidence generation for different model architectures (Peng, Z. et al. 2025).

By systematically applying these methodologies, the research aims to enhance the security and privacy of artificial intelligence language models while leveraging the innovative capabilities of zero-knowledge proofs. Figure highlights security risks in large language models

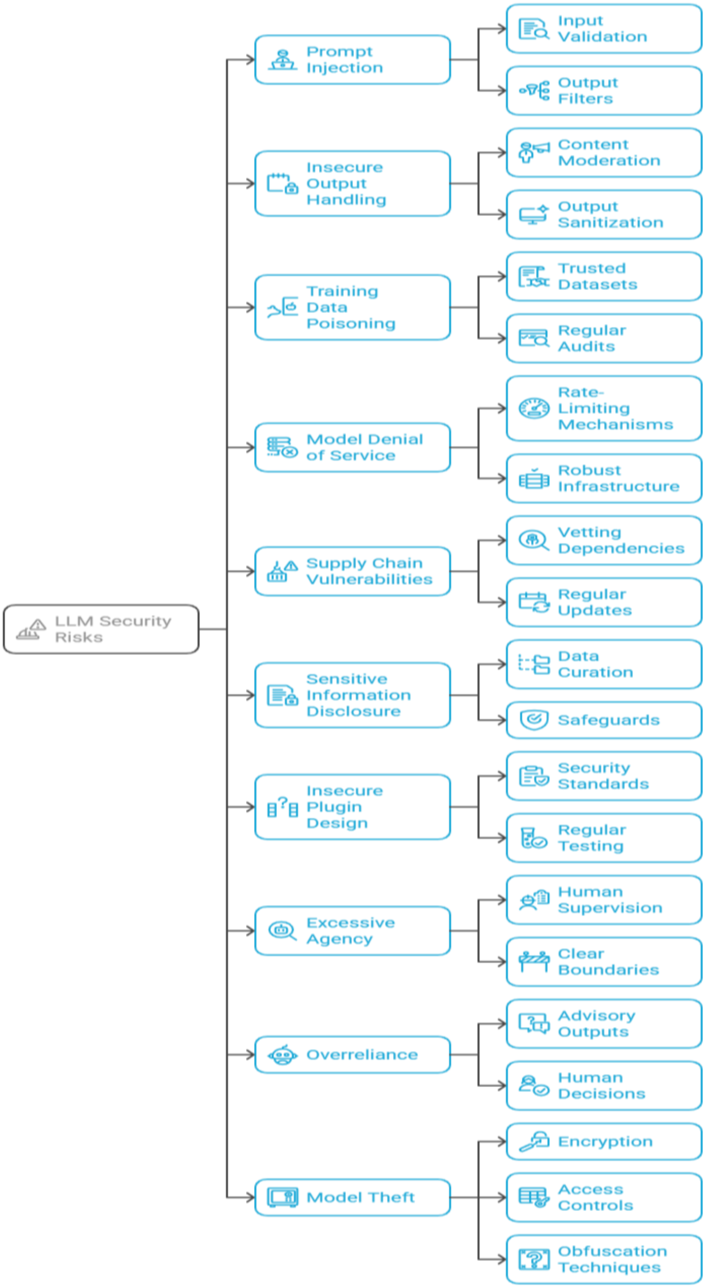


Figure 2. LLM Security Risks and Mitigation Strategies

4. Results

Research into the security of artificial intelligence language models (LLMs) has revealed a number of security vulnerabilities that pose significant risks. Key vulnerabilities include data poisoning, prompt injection, and sensitive information disclosure, all of which have been increasingly identified in recent evaluations of open-source AI and machine learning (ML) models (Ajayi, A. et al. 2024).

The evolving nature of these vulnerabilities necessitates the continuous evaluation and adaptation of security measures.

Recent evaluations show that innovative approaches such as Apollo and Artemis significantly outperform traditional methods in addressing these security vulnerabilities. These improve state-of-the-art security for ML models by an order of magnitude. Notably, Artemis achieves comparable performance to Apollo even when used without a trusted setup, demonstrating its robust design (Peng, Z. Et.al. 2025).

This performance metric highlights the effectiveness of new security frameworks in protecting LLMs against emerging threats.

Zero-knowledge proofs (ZKP) have been identified as a critical technology for enhancing the security of AI systems. By allowing data to be verified without revealing sensitive information, ZKP can effectively reduce the risks associated with data breaches. This cryptographic approach enables organizations to protect their data assets by ensuring that even if a system is compromised, the disclosed information remains incomprehensible to hackers

Implementing ZKPs as part of a zero trust framework further strengthens access controls and reduces potential attack surfaces.

4.1. Challenges and Recommendations

Despite these advancements, research highlights ongoing challenges in securing LLMs. The complexity and opacity of statistical models can hide security vulnerabilities and make it difficult to design effective evaluation schemes. Each identified security vulnerability, such as data leakage and model reversal attacks, requires specific strategies for detection and mitigation.

Organizations must take proactive measures, such as continuous security testing and human oversight, to address these vulnerabilities comprehensively. The OWASP Top 10 Risks for LLMs have been updated to reflect the latest insights on security vulnerabilities and emphasize the need for collaborative efforts to develop and implement robust security protocols across industries (Bertaccini, M. 2022).

Ongoing education and research on the evolving landscape of AI security will be critical to maintaining the integrity and security of AI language models.

Healthcare Collaboration

In a notable application of zero-knowledge proofs, two hospitals can collaborate on medical research while ensuring patient privacy. Using collaborative zk-SNARKs (zero-knowledge succinct non-interactive arguments), hospitals can analyze aggregate patient data such as treatment effectiveness or disease trends without disclosing individual patient details (Lavin, R. et al. 2024).

This is particularly important in the healthcare sector, where data privacy regulations are stringent and meaningful research must be enabled without compromising patient privacy.

Financial Transparency

Zero-knowledge proofs have emerged as a critical tool for enhancing transparency in financial transactions, particularly within cryptocurrency exchanges and wallets. Through mechanisms such as Zero-Knowledge Proof of Reserves (ZK-PoR), companies can demonstrate their ability to make payments without disclosing sensitive information about individual transactions (Lavin, R. et al. 2024). This trustless mechanism helps reduce fraud and mismanagement risks and increases trust among users of decentralized financial systems.

Artificial Intelligence and Employment Applications

Another case study highlights the application of zero-knowledge proofs in AI systems that screen job applications. By using zero-knowledge identity verification solutions, organizations can simplify the hiring process while protecting applicants' personal information. This mechanism ensures that sensitive data remains confidential even when hiring decisions are made, thereby supporting a zero-trust framework in human resources applications.

Approach to Cyber Security Threats

The integration of zero-knowledge proofs into cyber security frameworks has also been explored. For example, organizations can use these proofs to allow employees to access networks securely without disclosing their personal identity information. This approach helps reduce risks associated with data breaches, thereby strengthening the integrity of company networks against external threats.

Through these case studies, it is evident that zero-knowledge proofs provide innovative solutions across various industries, address privacy and security concerns, and facilitate collaboration in the healthcare, finance, and employment sectors.

Integrity Risks

One of the fundamental challenges of using AI models, especially in critical applications, is integrity risks.

These risks arise when the results produced by an AI model are incorrect due to unintended errors or deliberate manipulation by malicious actors (Scherlis, B, 2024). This concern is heightened in mission-critical areas where effectiveness, safety, security, and resilience are paramount, as outlined in the NIST AI Risk Management Framework (RMF).

Unfortunately, developers currently lack rigorous methods to patch or reliably identify these vulnerabilities, making it imperative to measure the effectiveness of security measures that require best efforts at both the system and operational levels (Scherlis, B, 2024).

Scalability and Performance

Zero-Knowledge Proof (ZKP) technology faces significant scalability and performance challenges. ZKP protocols can be computationally intensive, requiring significant processing power and time, which may render them unsuitable for real-time applications or systems with high transaction volumes. To address this, ongoing research is focused on discovering efficient implementation techniques and optimizing the fundamental cryptographic algorithms used in ZKP protocols. Advances in hardware acceleration, such as specialized chips and hardware modules, could also improve the performance of these protocols.

Complexity and User Experience

The complexity of implementing ZKP systems can negatively impact the user experience. Users may face difficulties understanding and interacting with complex ZKP frameworks, which could hinder broader adoption (Barbera, 2025)

Developing user-friendly interfaces and documentation to facilitate a better understanding of ZKP applications and increase their accessibility to a wider audience is of great importance.

Trade-offs Between Privacy and Transparency

In modern digital systems, there is a critical balance between transparency and privacy. While blockchain technology emphasizes transparency to promote trust and prevent fraud, this openness can compromise user privacy. Advanced analytics can link on-chain and off-chain data, revealing users' anonymity, exposing transaction histories, and leading to privacy breaches (Lavin, R. et al., 2024). ZKP offers a promising solution to this dilemma by providing proof of information accuracy without revealing underlying data, thereby ensuring the integrity of processes such as confidential financial transactions and private voting systems while preserving privacy.

Engineering Challenges

The application of artificial intelligence engineering presents unique challenges when compared to traditional fault-tolerant design methodologies. Traditional approaches typically rely on statistical independence assumptions between errors and use redundancy to minimize risks (Scherlis, B, 2024). In contrast, AI systems are inherently composed of less reliable components, which requires new engineering patterns and workflows to effectively manage and mitigate these security vulnerabilities.

Advances in Zero-Knowledge Proofs for AI Security

The integration of zero-knowledge proofs (ZKPs) into AI security offers a promising avenue for verifying AI language models and enhancing their privacy. Research by Lavin and others highlights the need for effective ZKP implementations in machine learning environments, focusing on balancing security guarantees with practical overhead to ensure these proofs remain concise and manageable for verifiers in real-world applications (Lavin, R. et al. 2024) (Peng, Z. et al. 2025). This approach not only addresses computational complexity but also opens the door to scalable solutions that can support more complex AI models without compromising efficiency.

Addressing Security and Safety Risks

As AI language models become more prevalent in critical systems, identifying and mitigating security risks is essential. Future research should focus on enhancing the robustness of AI models against integrity, privacy, and governance risks that can arise from both unintended errors and malicious manipulation (Scherlis, B, 2024).

By collaborating with developers, policymakers, and users, the AI community can establish guidelines and best practices that promote the ethical and secure use of these technologies.

The introduction of comprehensive security training and robust infrastructure designs will be critical in creating a safer AI environment.

Commercial Applications and Industry Adoption

The commercial potential of zero-knowledge proofs in AI security continues to expand with applications in sectors such as blockchain, secure communication, and data privacy.

Organizations are encouraged to adopt zero-knowledge protocols to enhance the privacy-preserving capabilities of AI models while ensuring compliance with data protection principles.

Future directions should include the discovery of practical frameworks that facilitate the widespread adoption of these technologies within businesses, thereby minimizing security risks while encouraging innovation (Singh, Shridhar 2024).

Research and Development Initiatives

Researchers are encouraged to investigate the scalability of ZKP applications across various machine learning frameworks to further advance the field of zero-knowledge proofs in artificial intelligence. This includes addressing existing challenges related to computational demands and creating proofs for complex tasks

The ongoing evolution of distributed systems and cryptographic techniques provides a fertile ground for innovation, potentially leading to new applications and improved performance of ZKPs in securing artificial intelligence language models.

5.Sonuç

In conclusion, the role of zero-knowledge proofs in large language model chains is critical for ensuring data privacy in today's digital environment. By incorporating these cryptographic protocols, individuals and organizations can protect sensitive information while leveraging the capabilities of advanced language models. As data privacy remains a top priority, the use of zero-knowledge proofs offers a promising solution for safeguarding privacy and trust in data-driven applications.

References

Ajayi, A. A., Igba, E., Soyele, A. D., & Enyejo, J. O. (2024). Enhancing digital identity and financial security in decentralized finance (DeFi) through zero-knowledge proofs (ZKPs) and blockchain solutions for regulatory compliance and privacy.

Ahmed, M. (2025). *AI as emerging technology: Opportunities, vulnerabilities and future directions*.

Barbera, I. (2025). *AI privacy risks & mitigations in large language models (LLMs)*.

Bertaccini, M. (2022). *Cryptography algorithms: A guide to algorithms in blockchain, quantum cryptography, zero-knowledge protocols, and homomorphic encryption*. Packt Publishing Ltd.

Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., ... & Amodei, D. (2018). *The malicious use of artificial intelligence: Forecasting, prevention, and mitigation*. arXiv. https://arxiv.org/abs/1802.07228

Lavin, R., Liu, X., Mohanty, H., Norman, L., Zaarour, G., & Krishnamachari, B. (2024). *A survey on the applications of zero-knowledge proofs*. arXiv. https://arxiv.org/abs/2408.00243

Peng, Z., Wang, T., Zhao, C., Liao, G., Lin, Z., Liu, Y., ... & Zhang, S. (2025). *A survey of zero-knowledge proof based verifiable machine learning*. arXiv. https://arxiv.org/abs/2502.18535

Scherlis, B. (2024, July 29). *Weaknesses and vulnerabilities in modern AI: Why security and safety are so challenging*. Carnegie Mellon University, Software Engineering Institute. https://doi.org/10.58012/h970-7f87

Singh, S. (2024, March). Enhancing privacy and security in large-language models: A zero-knowledge proof approach. In *International Conference on Cyber Warfare and Security* (pp. 574–582). Academic Conferences International Limited.

Yao, Y., Duan, J., Xu, K., Cai, Y., Sun, Z., & Zhang, Y. (2024). A survey on large language model (LLM) security and privacy: The good, the bad, and the ugly. *High-Confidence Computing*, *100211*.

Chapter 3

Analysis Of Energy Efficiency For An Indoor Swimming Pool

Selmin ENER RÜŞEN[1], Aydın RÜŞEN[2]

ABSTRACT

Increasing expenses and environmental concerns have made it necessary for buildings today to consume less energy. Especially, indoor swimming pools typically have different energy systems and use more water and energy than other types of construction. The economical weight of energy prices on the budget will be lessened by the effective use of energy in indoor swimming pools, which is also a crucial step toward ecological sustainability. In this study, the energy consumption in the selected indoor swimming pool (ISP) was examined and energy efficiency potential as well as carbon footprint was revealed as a case study. Initially, the energy consumption values were determined in the selected ISP, then the regular measurements were made with appropriate equipment at energy consumption points, and energy efficiency enhancing projects were suggested for these initiatives. Also, average greenhouse gas emissions of the ISP for the last three years were calculated as 208.28 tCO_2. According to the results, energy usage in the swimming pool might decrease up to 7.5% with energy-saving solutions with a relatively quick payback period (less than 3 years).

Key words: Energy consumption, Indoor swimming pools, Energy Saving, Carbon Footprint.

[1] Karamanoğlu Mehmetbey University, Faculty of Engineering, Department of Energy Systems Engineering, Karaman, Türkiye
ORCID: 0000-0003-3389-5739, selminerusen@kmu.edu.tr
[2] Karamanoğlu Mehmetbey University, Faculty of Engineering, Department of Metallurgical and Materials Engineering, Karaman, Türkiye
0000-0001-5592-1411, aydinrusen@kmu.edu.tr

1. INTRODUCTION

The building sector accounts for around one third of all final energy use globally. It is known that the primary uses of energy in buildings are to heat and/or cool the interior spaces. While doing this, determination of energy efficient method is the most critical part. Buildings should be energy efficient initially due to global warming and the depletion of fossil fuels, followed by renewable energy applications. Unquestionably, the construction industry could dramatically reduce its greenhouse gas emissions by improving its energy efficiency [1,2]

Since they require heating, dehumidification, and ventilation, indoor swimming pools (ISP) have high energy-consumption structures and they are also significant social and sporting venues [3]. There are many ways that swimming pools lose energy, but evaporation is by far the biggest reason. It is known that the amount of heat required to change the temperature of one gram of water by 1 °C is 1 cal (4.184 J). Water evaporation also consumes a great deal of energy. Evaporation accounts for a startling 70% of the energy lost from indoor swimming pools. It is anticipated that indoor swimming pools use three times more energy than offices of the same size since these facilities require forced ventilation [4,5]. In swimming pools, the heating, ventilation, and air conditioning (HVAC) system (27–50%) and pool heating (28–50%) use the majority of the energy. A complete comfort system that can be utilized to heat and cool your building as well as improve indoor air quality is referred to as an HVAC system.

In average standard swimming pools, 10-40% of the total energy provided is used by domestic hot water, 4-12% by lighting, 3-30% by building insulation and 10-30% by pump systems [6,7]. Different percentages are due to varying facility sizes, the number of swimming pools in a facility, the work schedule, the insulation properties of the building, and the effectiveness of energy-consuming machines. The impact of pool size, area, construction year, and building isolation on energy use should also be examined [7-10]. The architecture, together with its design and upkeep, is the only factor that influences how much energy is used by the equipment for energy consumption. In addition, energy consumption is greatly influenced by thermal and lighting systems [11-13].

This study investigates an energy consumption analysis as the basis for a case study of an ISP building in Karaman. In terms of the implementation costs, some energy efficiency techniques are proposed and assessed. As is customary in energy audits of buildings, the measures were graded according to two criteria: payback time and carbon dioxide emission [14-16]. It is crucial to stress that the ranking of the measures according to importance varies for each indoor swimming pool and is determined by the opinion of specialists. This research suggests conducting an energy audit of facilities with ISP, which would involve the use of isolation during the implementation of energy-saving measures. The findings of every energy consumption of undiscovered situations are particularly important because each swimming pool facility represents a distinct energy system [17-18]. As a result, the study also expands the knowledge pertaining to this type of buildings' energy efficiency.

2. METHODOLOGY

Energy consumption by device type and usage, energy cost and tariff data, and energy management techniques are among the details and papers needed during the thorough consumption and analysis. Energy efficiency aims to reduce energy bills while keeping existing conditions the same. While utilizing conventional loads, the energy consumption for swimming pools is calculated for each scenario and contrasted with the energy consumption while using energy-efficient loads. This comparison was made in terms of how much money was spent and how much electricity and natural gas were used. To determine the extent to which a certain load has an impact on the overall consumption of electricity, a load distribution has also been deduced from this information.

2.1. Energy Data Collection

Semi-Olympic Indoor Swimming Pool Building with an area of 6,456 m^2 built in 2015, is located in the city of Karaman. The swimming pool was heated with a central heating system before 2018, and then it was separated from the central system by installing a cascade system. The main energy consumptions of the building are electricity and natural gas.

Between 2016 and 2017, the swimming pool was heated by underground galleries from a heat center approximately 700 meters away. Rusen et al. reported that the flow rate was measured weekly in the direction of the hot water inlet of the swimming pool building during these periods (between November 2016 and April 2017), but the flow rate and temperature values of the hot water provided from the heat center showed instantaneous changes depending on time [19]. Although it is not possible to reveal exact values in the systems without continuous monitoring, an idea can be obtained about the heat load of the pool during that period by taking the average of all measurements. Considering the average flow rate of 18 m3/hour and the temperature difference of approximately 10 ^{o}C between the flow and return lines, the heating capacity provided to the ISP building can be calculated with the help of the basic heat transfer rate equation (Q = m.Cp.ΔT). In the light of these data, the amount of heat transferred to the pool during the specified period is 180,000 Kcal/hour. Considering the daily 24-hour usage (14 hours active + 10 hours passive), 4,320,000 Kcal/day and approximately 5200 m^3 of natural gas (low heating value: 8250 Kcal/m^3) is required to provide this energy. This corresponds to an annual natural gas usage of 93,600 m^3 in a system that operates continuously throughout the heating season (6 months). As stated above, since the flow rate and temperature of the water fed from the central system can change instantly, this calculation gives us an idea about the annual average natural gas consumption. However, it is not possible to reach the actual consumption values for this period. In this study, the same values were accepted throughout the 2-year period in which central heating was provided, ignoring seasonal conditions.

After separation from the central heating system, 4 wall-mounted condensing boilers (cascade) were used to meet the building and pool heating needs. The amount of natural gas consumed in the boilers was monitored monthly from the invoices. Since the average daily user number is stated as 200 on weekdays and 300 on weekends, the daily user number is accepted as 250. The exterior view of the building is presented in Figure 1.

Figure 1. Semi-Olympic Indoor Swimming Pool Building in Karaman.

2.2. Energy Consumption Analysis

The energy consumption of this building was examined for two different periods: i) Central heating system (2016-2017), ii) Cascade system (2021-2023) on a monthly basis. Data only belonging to electricity and natural gas are collected in the selected swimming pool building and are presented in the Table

1 with cost analysis. In addition, Figure 2 shows percentage distribution charts of energy consumption and costs for the selected years.

Table 1. Swimming Pool Building energy consumption values for the years 2016-2023

		Natural Gas				Electricity			
	Years	Cost ($)/$Sm^3$	Amount (Sm^3)	TOE	Cost ($)/TOE	Cost ($)/kWh	Amount (kWh)	TOE	Cost ($)/TOE
I.Period	2016	0.25	93600	77.12	303.42	0.13	269207	23.15	1511.74
	2017	0.22	93600	77.12	267.01	0.11	209030	17.97	1279.53
II. Period	2021	0.28	83031	68.42	339.79	0.13	196181	16.87	1511.76
	2022	0.60	83145	68.51	728.17	0.17	187026	16.08	1977.26
	2023	0.42	73217	60.33	509.71	0.18	180916	15.56	2092.85

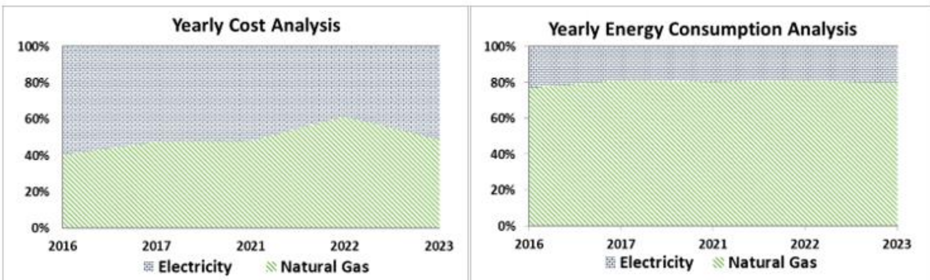


Figure 2. Percentage Distribution Charts of Yearly Energy Consumptions and Costs

When the energy consumptions for the periods are examined, it is observed that the share of electricity consumption in total energy consumption corresponds to 23% at 2016, while its share in the total cost is 60%. The reason for this difference in consumption and cost percentage distributions is the changes in energy unit costs. The unit cost of natural gas is approximately 5 times more expensive than the unit cost of electricity in I.Period. However, in the II.Period, there is a change in the unit costs of natural gas and electricity between 3 and 5 times. This fluctuation can be explained especially by the negative impact of the Covid-19 pandemic on the energy sector.

a. Average Monthly Energy Consumption Analysis for Three-Year

Since the electrical measurements for the first period include a single tariff, monthly electricity consumption according to the electricity tariff was examined only for the second period. The results are given in Figure 3, including monthly natural gas consumption for the same period.

When monthly electricity consumption is examined, the months with high electricity consumption vary from year to year. In August, when the air temperature is high, the use of air conditioning systems in particular increases electricity consumption, while a similar situation shows that electricity consumption for heating and lighting purposes increases in the winter months.

As for natural gas consumption, it can be said that it varies in the same months depending on the year. This situation can be attributed to the number of heating degree days (HDD) and cooling degree days (CDD) during the year.

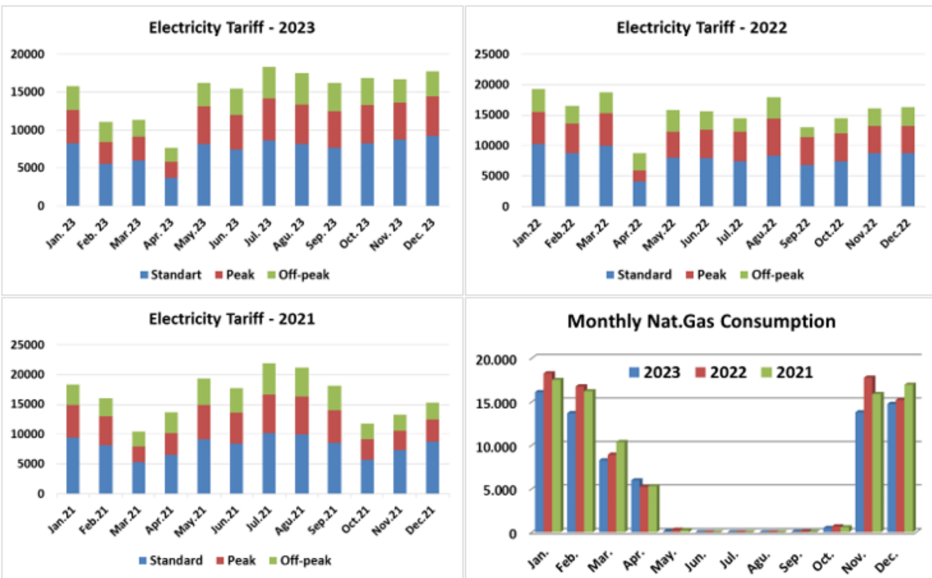


Figure 3. Monthly electricity consumption according to the electricity tariff for II.Period as well as natural gas consumptions on monthly base.

b. Average Monthly Energy Consumption Analysis with Respect to the HDD and CDD Values for Three-Year

The number of heating degree days (HDD) and the number of cooling degree days (CDD) in a region play an important role in determining the cost and capacity of the heating and cooling system. For this reason, monitoring energy consumption together with HDD and CDD is important in terms of determining energy efficiency in buildings. Figure 4 shows the Heating and Cooling Day Degrees for selected region (MGM 2024[20]). In addition, the comparison of the electricity and natural gas energy consumption of the ISP according to HDD and CDD was given in Figure 5 for last three years.

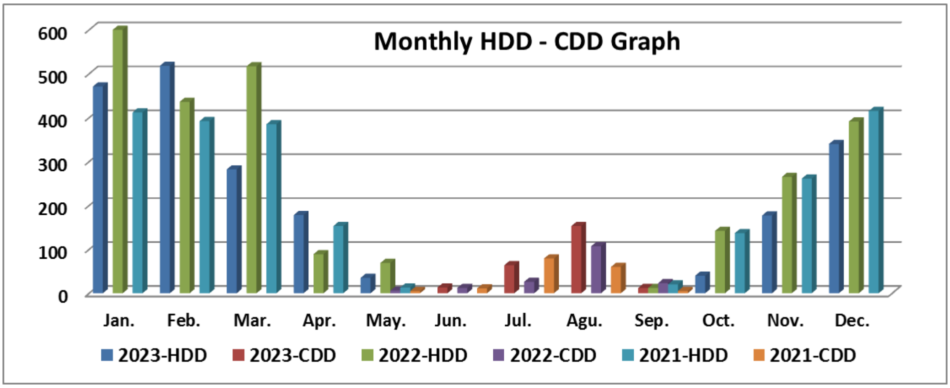


Figure 4. Number of degree days for three-year

When the monthly consumption graphs of natural gas and electricity energy are examined together with HDD and CDD values (Fig.5), it is understood that the increase in the energy consumption of the building changes according to seasonal conditions for all years. While there is intense natural gas consumption due to high HDD in the winter months, natural gas consumption is zero in the summer months because there is no need for energy. In general, CDD values for natural gas and electricity consumption show the same trend.

It is observed from Figure 5 that natural gas consumption decreases in proportion to the decrease in HDD value by starting from April. Also, it is seen that electricity consumption starts to increase after this month. Furthermore, it is observed that electricity consumption peaks in August, when the highest CDD values are reached, due to the increase in cooling needs.

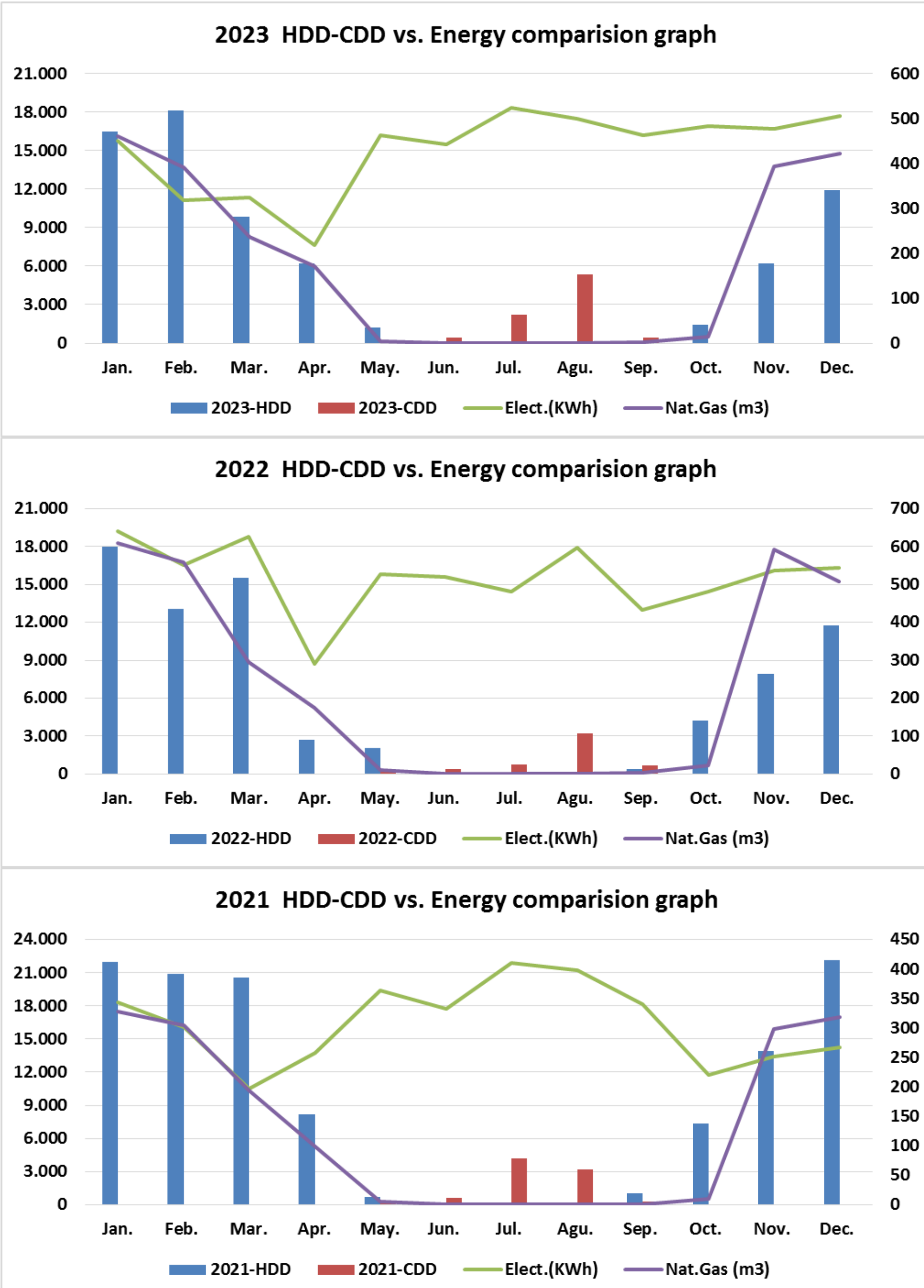


Figure 5. The comparison of the electricity and natural gas energy consumption of the ISP according to HDD and CDD for last three years.

3. ENERGY AUDIT RESULTS

Within the scope of this study, an energy audit of a swimming pool was conducted and various measures were determined to increase energy efficiency. With the improvement suggestions that can be applied to the ISP building, the energy consumption of the building will be reduced, financial savings will be provided, and environmental conditions will be improved. Within the scope of the energy study, basic components such as insulation, heating system, lighting system were examined and various efficiency-enhancing suggestions were made.

3.1. Insulation

ISP building exterior was measured by a thermal camera (Testo 875 – 2i). In addition, the image taken by thermal camera was evaluated by IR – Soft Pc thermal camera software and the average surface temperature was used in the calculations. When the total energy consumption of the swimming pool is examined, it is determined that the natural gas consumption is higher than expected. Therefore, some studies have been focused on reducing natural gas consumption. As a result of the research, it has been determined that the existing building exterior insulation is partially done. There are thermal bridges and energy leaks in the insulation. A sample thermal camera measurement is given in Figure 6.

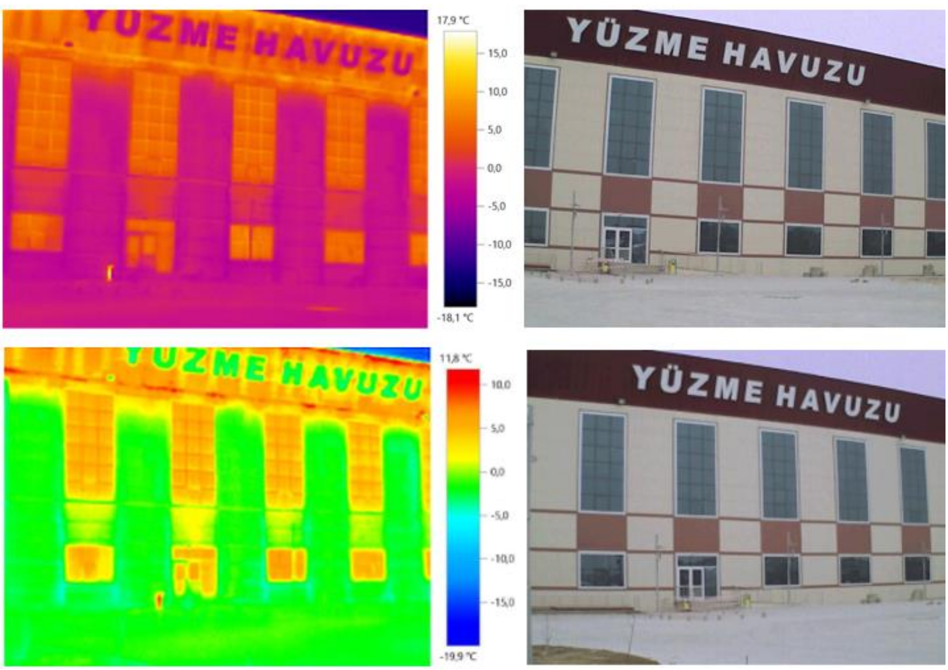


Figure 6. A sample thermal camera measurement of ISP

The building (ISP) has glass and reinforced concrete on the front surface, and only reinforced concrete on the back. There are aluminum framed windows on the side surface. It was read from the project that thermal insulated double glass with 12 mm air gap was used, and it was determined on site. Various thermal bridges are detected from the image. As can be seen in Figure 6, thermal bridges are seen especially at the junction points for the ISP building. It is foreseen that energy saving can be achieved by eliminating insulation problems with the improvement works to be carried out at these points. Thereby, electricity from the cooling systems in the summer and natural gas from the heating system in the winter will be saved.

Another project proposal to increase energy efficiency is to eliminate insulation deficiencies in the heating installation system. The selected ISP building heating installation was examined, and it was seen that there was partial insulation in the hot water transport lines. It was calculated that energy savings would be achieved by properly insulating the uninsulated lines and connection elements.

Energy losses before and after the insulation can be calculated with using the data from the measurements by following heat transfer formulas [21];

- **Heat losses before insulation (Eq. 1)**

$$Q_1 = (U_c + U_r)*A*(T_s - T_a)$$

Heat transfer coefficient by radiation

$$U_r = [(E * 5{,}67)/(T_s - T_a)]*[(T_s/100)^4 - (T_a/100)^4]$$

Heat transfer coefficient by convection

$$U_c = B*(T_s - T_a)^{0.25}$$

- **Heat losses after insulation(Eq. 2)**

$$Q_2 = A \times (T_s - T_a) / [(L/\lambda) + (1/\alpha_d)]$$

- **Heat savings(Eq. 3)**

$$Q_S = Q_1 - Q_2$$

where Ts and Ta are the surface and ambient temperatures, respectively. A is the surface area of the building needed insulation (~300m^2), B is the multiplying factor which equals to 1.75 when the building is located to vertical to the ground. E is the emissivity coefficient of surface material of the building (0.94), L is the insulation thickness (0.05 m), λ is the thermal conductivity coefficient of the insulation material (0.04 W/m.K) and α_d is the thermal conductivity coefficient by heat convection (20 $W/m^2.K$).

After calculation by using the above equations and data, the annual heat loss of the swimming pool building due to inadequate insulation was calculated as

40,230,167 kcal. This value corresponds to 4,876 Sm³/year natural gas consumption and 4.02 TOE. When compared to the total annual TEP amount (~76 TEP) consumed for 2023, it will provide an improvement (energy saving) of around 5%.

In the study, it was determined that the building's thermal insulation was not good. The adequacy of the insulation can also be revealed by measuring the U-value in the building. Thus, its compliance with TS-825 is also determined. With appropriate insulation processes to be carried out in buildings with insufficient insulation, significant energy savings can be achieved by making the heat transmission coefficients of the building shells compatible with the TS-825 standard. In addition, significant energy savings can be achieved by improving the building's glass, providing a heat transmission coefficient (2.4 W/m^2K) compatible with the TS-825 standard.

3.2. Heating System

As mentioned before, after the building's heating system is separated from the central system, it is heated with 4 wall-mounted condensing boilers (cascade system). Building heating by pressing the hot water produced in the boilers to the radiators is provided. There are thermostatic valves in radiators. The general view of the wall-mounted condensing boilers of the building heating system and the label image are presented in Figure 7.

Figure 7. The general view of the wall-mounted condensing boilers of the ISP building heating system.

The flue gas measurements of the heating system were made with the TESTO 435 Flue Gas Measurement Device. With the help of the flue gas analyzer, parameters such as O_2 percentage in the flue gas, CO, SO_2, NOx emission amount, CO_2 percentage, gas and combustion air temperatures, efficiency and excess air ratio can be monitored instantly. Thus, by improving the combustion efficiency of the boiler, fuel savings and environmental pollution emission rates can be reduced.

According to the measurements, the flue gas temperature is 68 ^{o}C, the O_2 ratio in the flue gas is 3.1% and the CO ratio in the flue gas is 0 ppm. These results show that there is no problem in the combustion system of the boiler and that it works efficiently. It has been determined that there is no need for burner adjustment, economizer or recuperator applications to benefit from waste heat, which are generally recommended in energy efficiency studies in boilers. It is recommended that periodic maintenance be carried out on time in the boiler system and that the fault be fixed by authorized technical personnel in case of a fault.

In addition to the energy coming from the heating system of the ISP building, it is suggested to be equipped with solar energy-supported water heating systems and solar energy supported electric resistances to be used when needed.

3.3. Lighting System

In this study, the efficiency-increasing project proposed for the improvement of lighting systems was carried out without reducing the quality of lighting and by fulfilling the requirements of good lighting. It was determined that there were 528 lamps of different powers (18 - 400 W) and types (incandescent filament, Fluorescent, Metal Halide, etc.) in the ISP building. They have 46,774 KWh total power and their average annual usages are 2000 hours. It is known that LED lamps, which can be up to 80% more efficient, are preferred instead of existing lighting systems for quality and economical lighting [22,23]. Therefore, LED conversion was suggested as an energy efficiency project.

According to the calculations made by considering the difference in energy consumed in case LEDs are used instead of fluorescent or other types of lamps in ISP building, which will provide the same light comfort with the annual electricity consumed; it has been determined that approximately 28,100 KWh of energy can be saved in one year when efficient lighting systems are preferred. By this way, approximately 15% energy saving can be achieved in annual electric consumption, which corresponds to ~2.5% of the total annual energy consumption for 2023. It has been determined that the investment made will amortize itself less than 3 years according to the calculation made with the simple payback method.

The most effective action in energy saving studies in the field of lighting is the use of natural light. For this, it is recommended that the windows be positioned correctly and, if necessary, daylight lamps be used. In addition, energy-saving lighting fixtures, usage of motion sensors, appropriate color use, main switch and time control applications can be recommended for energy efficiency in lighting.

4. CARBON FOOTPRINT CALCULATION RESULTS

In this part of the study, greenhouse gas emissions (also carbon footprint) were calculated for the ISP building for the selected years (Figure 8). Tier 1 and Tier 2 approaches recommended by IPCC were used in the emission calculations. Accordingly, the IPS building's natural gas fuel consumption for heating purposes was evaluated as a direct emission (Scope-1), while the emissions released into the atmosphere as a result of electricity consumption purchased from outside by the institution were evaluated as indirect emission (under Scope-2). Within Scope 1, the natural gas fuel lower calorific value and the appropriate emission factor for the fuel were taken into account in the IPCC 2006 Guide. On the other hand, for electricity consumption within Scope 2, the total carbon dioxide (CO_2) amount was reached by using the special emission factor defined for the relevant year in the national inventory.

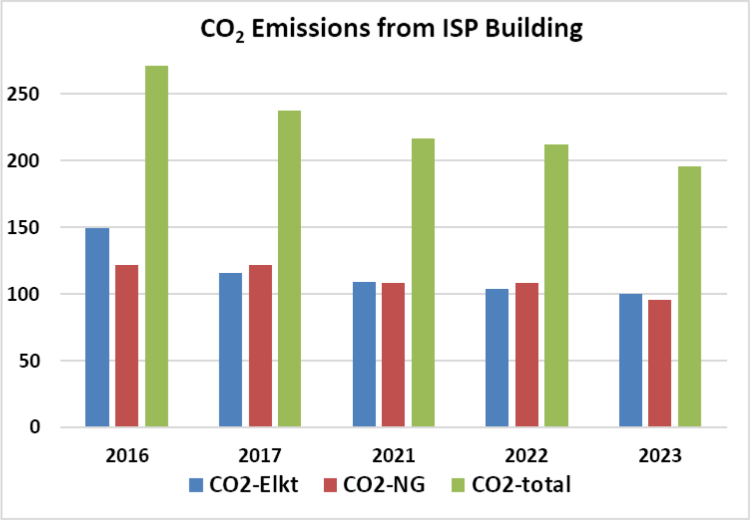


Figure 8. Greenhouse gas (CO_2) emissions from the ISP building for the selected years.

The amount of CO_2 emissions per person was calculated by considering the number of people (~250 per day) receiving service from the ISP. Also, the amount of CO_2 emissions per square meter area were calculated depending on the total usage areas for the ISP building (6454 m^2). The results of the calculation are given in Figure 9.

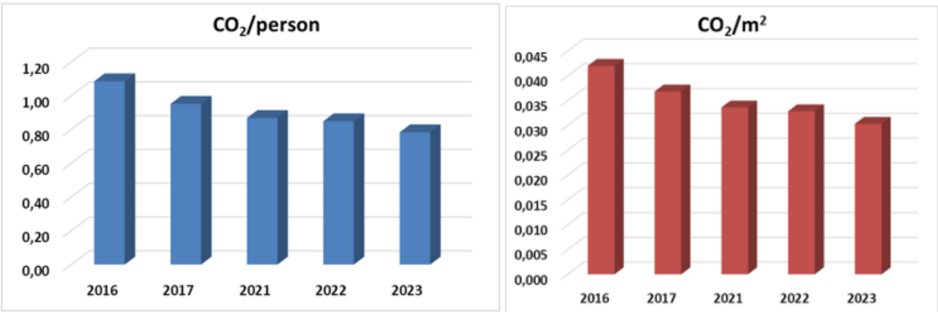


Figure 9. CO_2 emissions per person and per m^2 for the ISP building for the selected years.

As compared the Table 1 with Figure 8, it can be concluded that the change in the amount of CO2 emissions depending on the years is parallel to energy consumption. This situation is also reflected in the CO_2 emissions per person and per m^2.

In the medium and long-term emission reduction action planning of the institution, the base (reference) year should be determined for the reduction target. Since it best represents the current emission level and reflects the most up-to-date data, it is more appropriate to select the inventory period of 2023 as the "Base Year". The total amount of greenhouse gas emissions for the reference year was determined as 195.76 tons of CO_2.

According to the results of the study, it has been calculated that approximately 7.5% energy savings will be achieved by properly insulating and lighting, which corresponds to the fact that approximately 12.33 tons of CO_2 carbon emissions per year can be prevented.

5. CONCLUSIONS

This study is focused on an energy consumption analysis as the basis for a case study of an indoor swimming pool (ISP) building in Karaman. Energy consumption has been evaluated separately as electricity and natural gas. The values for the selected years with two different periods are calculated according to the amount of energy consumption. The first period consists of the years 2016 and 2017 when central heating was used, and the second period includes the last 3 years (2021, 2022 and 2023) when condensing boilers were used for ISP building.

Within the scope of this study, an energy audit of a swimming pool was conducted, and various measures were determined to increase energy efficiency. The improvement suggestions that can be applied to the ISP building aimed to reduce the energy consumption of the building, provide financial savings and improve environmental conditions. In this study, basic components such as insulation, heating system, lighting system of the ISP were examined, and various efficiency-enhancing suggestions were made.

It is stated by researchers that many indoor swimming pools have a financial gain potential between 5% and 20% at the end of energy efficiency projects. In particular, insulation, HVAC and lighting projects with short payback periods but different investment costs are at the forefront in these studies. A similar study was conducted here, and the results showed that approximately 7.5% energy savings could be achieved by properly insulating and lighting with simple payback period less than 3 years. After applying these projects, up to 12.33 tons of CO_2 carbon emissions per year could be prevented. In addition, instead of using energy obtained from fossil fuels, it has been suggested to reduce greenhouse gas (CO_2) emissions by supporting solar energy-based electrical resistances and water heating systems.

ACKNOWLEDGMENTS

The author acknowledges the management of Karamanoğlu Mehmetbey University for support this research study.

REFERENCES

[1] Rusen, S.E., Rusen, A. (2022). A Sample Study for Determining Energy Consumption Values in Public Buildings: Central Anatolia Region. In: Gökçekuş, H., Kassem, Y. (eds) Environmental Earth Sciences. Springer, Cham. https://doi.org/10.1007/978-3-031-04375-8_12.

[2] ISO (2014). Energy audits-requirements with guidance for use, ISO 5002:2014. International Organization for Standardization.

[3] Lebon, M., Fellouah, H., Galanis, N., Limane, A., & Guerfala, N. (2017). Numerical analysis and field measurements of the airflow patterns and thermal comfort in an indoor swimming pool: A case study. Energy Efficiency, 10(3), 527–548. https://doi.org/10.1007/s12053-016-9469-0.

[4] Seneviratne, M. (2006). A practical approach to water conservation for commercial and industrial facilities. Elsevier. https://doi.org/10.1016/B978-1-85617-489-3.X5001-2.

[5] Jelena Nikolic & Dusan Gordic & Nebojsa Jurisevic & Vladimir Vukasinovic & Dobrica Milovanović (2021); Energy auditing of indoor swimming facility with multi-criteria decision analysis for ranking the proposed energy savings measures", Energy Efficiency, 14: 36.

[6] Leitat & Sea. (2014). D2.2. Integrated report, Energy audits and Energy Performance Certification (p. 49). http://step2 sport.eu/sport/wp-content/uploads/2015/09 /WP2_20150612_2_STEP-2-SPORT_D2-2_Integrated_ report.pdf

[7] Kampel, W. (2015). Energy Efficiency in Swimming Facilities, PhD Thesis, Norwegian University of Science and Technology, Trondheim.

[8] Kampel, W., Aas, B., & Bruland, A. (2013). Energy-use in Norwegian swimming halls. Energy and Buildings, 59, 181–186. https://doi.org/10.1016/j.enbuild.2012.11.011.

[9] Kampel, W., Aas, B., & Bruland, A. (2014). Characteristics of energy-efficient swimming facilities – A case study. Energy, 75, 508–512. https://doi.org/10.1016/j.energy.2014.08.007.

[10] Kampel, W., Carlucci, S., Aas, B., & Bruland, A. (2016). A proposal of energy performance indicators for a reliable benchmark of swimming facilities. Energy and Buildings, 129, 186–198. https://doi.org/10.1016/j.enbuild.2016.07.033.

[11] Papadopoulos, A.M. "Forty years of regulations on the thermal performance of the building envelop in Europe: achievements, perspectives and challenges". Energy and Buildings, 127 (2016), pg. 942-52.

[12] Açıkkalp, E., Yucer, C. T., Hepbasli, A., Karakoc, T. H. "Advanced low exergy (ADLOWEX) modeling and analysis of a building from the primary energy transformation to the environment". Energy and Buildings, 81 (2014), pg. 281-286.

[13] Anagha Darshan, Naman Girdhar, Rohan Bhojwani, Kanish Rastogi, S. Angalaeswari , L. Natrayan and Prabhu Paramasivam; Hindawi Advances in Civil Engineering Volume (2022), Article ID 4400874, 10 pages, Energy Audit of a Residential Building to Reduce Energy Cost and Carbon Footprint for Sustainable Development with Renewable Energy Sources.

[14] Coşgun, A. E., 2021, The potential of Agrivoltaic systems in Turkey, Energy Reports, vol. 7, pp. 105–1.

[15] Ibrahim, M., Biwole, P. H., Achard, P., Wurtz, E., & Ansart, G. "Building envelope with a new aerogel-based insulating rendering: Experimental and numerical study, cost analysis, and thickness optimization". Applied Energy, 159 (2015), pg. 490-501.

[16] Sartor, K., Dewallef, P. Exergy analysis applied to performance of buildings in Europe. Energy and Buildings, 148 (2017), pg. 348-354.

[17] Saari, A., & Sekki, T. (2008). Energy consumption of a public swimming bath. The Open Construction and Building Technology Journal, 2(1), 202–206. https://doi.org/10.2174 /1874836800802010202.

[18] Trianti-Stourna, E., Spyropoulou, K., Theofylaktos, C., Droutsa, K., Balaras, C. A., Santamouris, M., Asimakopoulos, D. N., Lazaropoulou, G., & Papanikolaou, N. (1998). Energy conservation strategies for sports centers: Part B. Swimming pools. Energy and Buildings, 27, 123–135.

[19] Rusen S.E, Topçu, M.A., Celep, G.K., Çeltek, S.A., Rüşen A. (2018) Üniversite Kampüs Binaları için Enerji Etüdü: Örnek Çalışma, Çukurova Mühendislik Mimarlık Fakültesi Dergisi 33 (2), 83-92

[20] MGM 2024. Meteoroloji Genel Müdürlüğü. https://www.mgm.gov.tr/veridegerlendirme/gun-derece.aspx

[21] Topçu M.A. & Rusen A. (2016). Insulation of Boiler to Save Energy. 8th Ege Energy Symposium and Exhibition (IEESE), Afyon, 1-7.

[22] Singh, D., Basu, C., Meinhardt-Wollweber, M.,Roth, B., 2015. LEDs for Energy Efficient Greenhouse Lighting. Renewable and Sustainable Energy Reviews, 49, 139-147.

[23] Jones, B.A., 2018. Measuring Externalities of Energy Efficiency Investments using Subjective Well-Being Data: The Case of LED Streetlights. Resource and Energy Economics, 52, 18-32.

Chapter 4

The Impact of Big Data Preprocessing Techniques on Data Analytics: A Bibliometric Analysis (2000-2024)

Muhammet BERİGEL[1], Ayşegül ŞAHİN BAYSAN[2], Abdülaziz TÜRKYILMAZ[3]

Abstract

Big data is an area where intensive academic studies are conducted, especially in the field of technology. The development of big data processing techniques in Data Analytics has accelerated in recent years and has been changing in many ways. In this study, the academic literature on big data preprocessing techniques in the context of data analytics was scanned in the Scopus database between 2000-2024 with the bibliometric analysis performed, and 1089 articles were analyzed from different perspectives. The R software, together with the Biliometrix library and the biblioshiny tool, were used in the analysis of the data. The study results show that the number of publications and citations in this field has increased in recent years, China and the USA are the dominant countries, and research is concentrated in the areas of edge computing, real-time processing, machine learning, and cloud computing.

[1] Assoc.Prof.Dr, Karadeniz Technical University, Management Information System Department, ORCID: 0000-0001-5682-8956, berigel@ktu.edu.tr

[2] Researcher, Karadeniz Technical University, Department Of Biostatistics And Medical Informatics, ORCID: (0009-0000-7751-9044, sahinaysegul91@gmail.com

[3] Researcher, Karadeniz Technical University, Management Information System Department, ORCID: 0000-0001-7793-2859, abdulazizturkyilmaz@gmail.com

1. Introduction

In recent years, with the advancement of digital technologies and the acceleration of digitalization processes, the concept of big data has become increasingly important. The increasing prevalence of smart devices, interconnected systems and cloud technologies has significantly contributed to the exponential increase in the volume of data generated worldwide (Oussous et al., 2018). This increase includes structured, semi-structured and unstructured data generated through various sources such as texts, images, videos, sensor signals, social media and financial transactions (Gahi, Guennoun & El-Khatib 2016). Big data has become a transformative force in many areas, including healthcare, education, government, finance, marketing and social media (Mayer-Schönberger and Cukier, 2014,Garcia et al,2016).

The concept of big data is not only about size; it is often characterized by the "3Vs": volume, velocity and variety (Laney, 2001, Hu et al., 2014). Over time, scientists have expanded this definition to include concepts such as accuracy (data reliability) and value (usefulness of data) (Chen and Zhang, 2014; Yaqoob et al., 2016). These elements highlight the complexity of processing big data, where accurate and timely analysis of large and heterogeneous datasets requires specialized preprocessing techniques and advanced analytics.

Preprocessing techniques are fundamental to big data analytics. Raw data is often incomplete, inconsistent, noisy, and therefore not suitable for immediate analysis (García, Luengo & Herrera, 2015). Before data can be analyzed, it must undergo a series of processes, such as data cleaning, integration, transformation, feature selection, and normalization. These steps help increase the reliability and quality of the data, which in turn increases the validity of analytical outputs (Oussous et al., 2018, Golov and Rönnbäck 2017). In addition, data collection itself is prone to errors and inconsistencies. Imperfect data collection methods often result in noisy datasets that contain inaccurate or irrelevant information. Such deficiencies reduce the reliability of subsequent mining and analysis efforts and highlight the importance of robust data cleaning and validation processes (Prakash, A., Navya, N., & Natarajan, J., 2019).

Big data offers organizations unique opportunities to gain insights, optimize operations, and develop innovative services. The ability to uncover hidden patterns and trends from large data sets has transformed industries such as healthcare, finance, and energy (Chen and Zhang, 2014). However, these opportunities come with significant challenges. Another critical challenge is privacy. Since big data often involves the collection of personal and sensitive information, concerns about individual privacy rights and data misuse have increased (O'Leary, 2015). The ethical implications of large-scale data collection, the need for secure storage, and the risks of misinterpretation all require critical attention (Mayer-Schönberger and Cukier, 2014).

Finally, the role of preprocessing has been increasingly emphasized in the literature in various fields. Yaqoob et al. (2016) emphasize that preprocessing not only reduces data complexity but also provides meaningful analytical results, especially in systems that process real-time or heterogeneous data streams. Similarly, Chen and Zhang (2014) emphasize that proper data cleaning and dimensionality reduction are critical to improving efficiency and scalability in big data frameworks. In the energy sector, Li et al. (2017) emphasize preprocessing as a necessary step to ensure reliability in smart grid analytics. Furthermore, Oussous et al. (2018) argue that with the increasing use of distributed computing systems such as Hadoop and Spark, well-structured preprocessing pipelines are essential to effectively handle volume, speed, and variety. Taken together, these perspectives emphasize that preprocessing is not just a preparatory task, but also a strategic and indispensable phase within the broader big data analytics lifecycle.

In conclusion, while the potential of big data is enormous, realizing its full value requires a coordinated effort to address critical preprocessing challenges. Confidentiality, security and size of data processing, data collection point and ensuring data quality are included in this scope.Given this backdrop, the academic community has shown increased interest in exploring how data preprocessing impacts the effectiveness of data analytics. However, there is a lack of consolidated knowledge that captures the evolution, trends, and influential contributions within this field. This bibliometric study is designed to address this gap.This study aims to conduct a comprehensive bibliometric analysis of academic publications between 2000 and 2024 focused on big data preprocessing techniques and their application in data analytics.

2. Methodology

This study adopts an objectivist research philosophy, employing a quantitative bibliometric analysis to examine the scholarly landscape of big data preprocessing techniques and their applications in data analytics. Bibliometric analysis, including citation, co-citation, and keyword co-occurrence analysis, provides a robust methodological framework to uncover publication patterns, influential contributors, thematic clusters, and intellectual structures within a research field (Diodato & Gellatly, 2013,Berigel and Şılbır 2024).

2.1 Data Source and Search Strategy

The bibliometric data for this study were extracted from the **Scopus database**, which was selected due to its wide coverage of peer-reviewed academic literature across disciplines and authorized institutional access. The dataset was constructed using a comprehensive keyword-based search strategy that included terms relevant to big data preprocessing and analytics, such as:

"Big Data Preprocessing" OR "Data Preparation for Big Data Analytics" OR "Data Cleaning in Big Data" OR "Data Integration in Big Data" OR "Feature Selection in Big Data" OR "Data Transformation for Big Data" OR "Data Wrangling in Big Data" OR "Big Data ETL" OR "Big Data Pipeline" OR "Data Preprocessing for Machine Learning" OR "Data Quality in Big Data" OR "Scalability of Data Processing" OR "Distributed Data Processing" OR "Big Data Processing Frameworks" OR "Real-time Data Processing" OR "Data Curation in Big Data" OR "Automated Data Cleaning" OR "Anomaly Detection in Big Data" OR "Data Reduction Techniques for Big Data" OR "Handling Missing Data in Big Data" OR "Data Normalization in Big Data"

The search was limited to documents published between 2000 and 2024 and written in English. The inclusion of this time span reflects the modern evolution of big data technologies since the early 21st century. All document types indexed in Scopus were considered, including journal articles, conference proceedings, book chapters, reviews, and editorials. This inclusive approach allows for a comprehensive understanding of theoretical and practical advancements in the field.

2.2 Data Processing and Analysis Tools

A total of **1089 publications** were retrieved from Scopus based on the defined search criteria. The dataset was exported in BibTeX format and processed using the Bibliometrix R package and its graphical interface Biblioshiny (Aria & Cuccurullo, 2017). These tools enabled efficient data cleaning, transformation, and visualization to perform quantitative analyses.

2.3 Limitations

The analysis was limited to the **Scopus** database due to access availability. Future research may consider incorporating other databases such as Web of Science or Dimensions to ensure broader coverage. Furthermore, the use of bibliometric tools provides insights into publication trends and network structures but does not replace a systematic content analysis or qualitative assessment.

3. FINDINGS

3.1 Main Information About Data

Description	Results
Timespan	2000:2025
Sources (Journals, Books, etc)	609
Documents	1089
Annual Growth Rate %	-9,15
Document Average Age	6,23
Average citations per doc	21,6
References	0
DOCUMENT CONTENTS	
Keywords Plus (ID)	7763
Author's Keywords (DE)	3400
AUTHORS	
Authors	5524
Authors of single-authored docs	62
AUTHORS COLLABORATION	
Single-authored docs	65
Co-Authors per Doc	6,03
International co-authorships %	26,63
DOCUMENT TYPES	
article	1089

Figure 1. General bibliometric indicators of publications on big data preprocessing between 2000 and 2025

Figure 1 summarizes key bibliometric indicators of 1089 documents published between 2000 and 2025 in the field of big data. These publications span 609 sources, with an average of 21.6 citations per document. The high number of authors (5524) and keyword diversity reflect the conceptual richness and collaborative nature of the field. Most documents are co-authored, with 6.03 authors per paper and 26.63% involving international collaboration. The average document age of 6.23 years suggests the research is relatively recent.

3.2 Annual Scientific Production

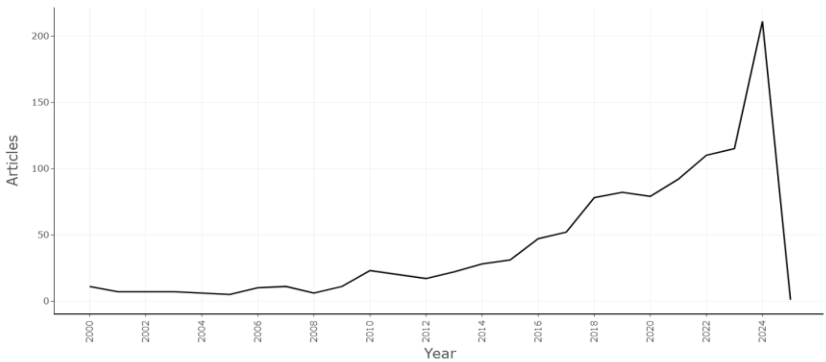


Figure 2. Annual Scientific Production Graph between 2000-2024

The ability to process and use big data together with three related concepts has played a very important role. This situation has also led to an increase in the number of academic studies on the subject. According to Figure 2, there has been a general increase in the number of publications from 2012 to the present. In particular, a significant increase can be observed from 2020 onwards, largely due to the increase in internet use caused by the COVID-19 pandemic. Another significant upward trend is seen in studies published from 2022 to the present. It can be argued that developments in the field of artificial intelligence, especially developments in language processing-based AI systems, have contributed to this increase.

3.3 Average Citations Per Year

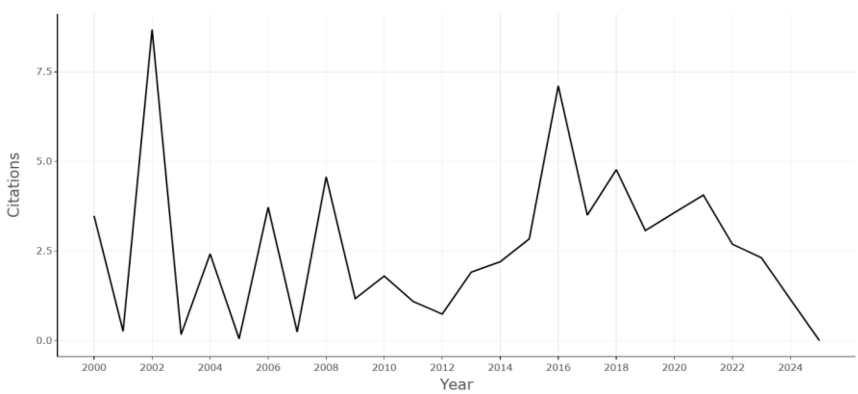


Figure 3. Average citations per year for publications on big data preprocessing (2000–2024)

Figure 3 presents the average number of citations per year for publications between 2000 and 2025. A distinct peak observed in 2002 indicates that the publications from this early phase had a substantial impact on the field. Following 2002, significant fluctuations in citation counts were noted, with marked declines during certain years. Particularly between 2010 and 2013, citation rates remained relatively low, suggesting a quieter period in big data research activity.

From 2014 onwards, a clear upward trend in citations can be seen, culminating in a second major peak in 2016. This pattern suggests a renewed academic interest and influence in the field of big data research during the mid-2010s. In the years following 2022, a noticeable decline in citation rates is evident. However, this decline is likely attributable to the recency of the publications, as newly published works typically require more time to accumulate citations.

Overall, **Figure 3** highlights the fluctuating influence of big data research over time, with certain periods standing out for their significant scholarly impact.

3.4 Most Relevant Sources

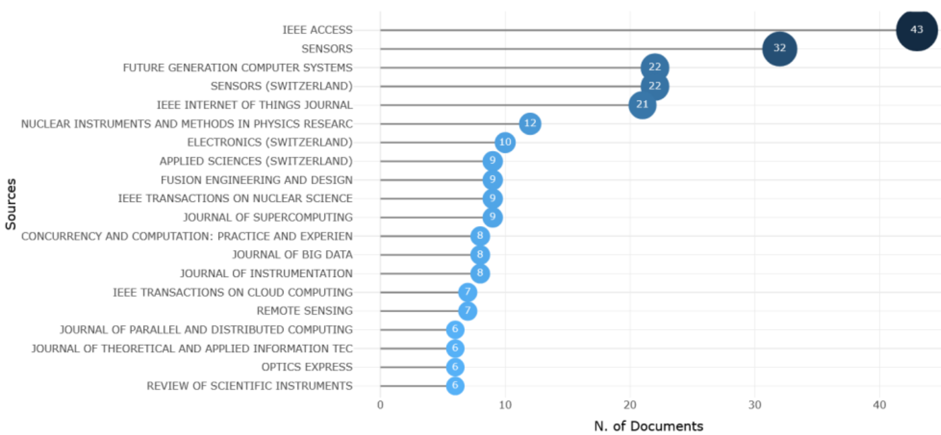


Figure 4. Most Relevant Sources

Figure 4 shows that engineering-oriented journals such as IEEE Access, Future Generation Computer Systems and Electronics stand out. This reflects the importance of processing big data and establishing infrastructure in terms of engineering. In addition, the fact that physics-oriented journals such as Nuclear Instruments in Physics Research are at the top shows that this field has a multidisciplinary structure.

3.5 Most Relevant Authors

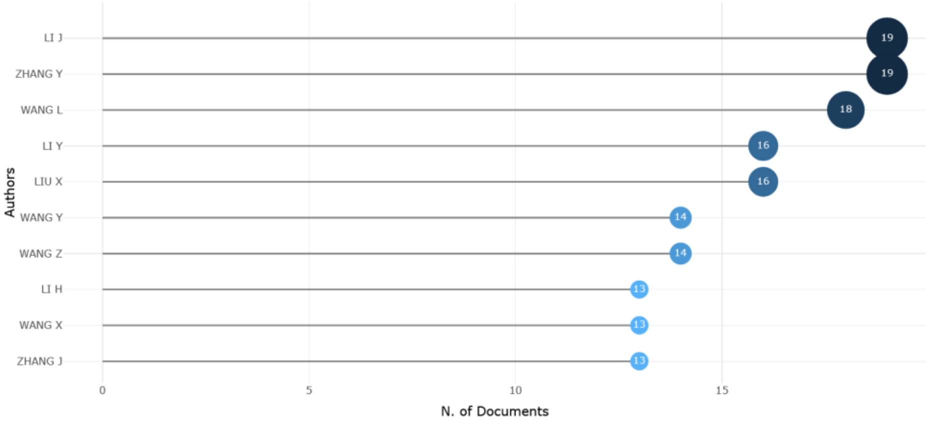


Figure 5. Top 10 most productive authors in big data preprocessing research (2000–2024).

Figure 5 shows the most prolific authors in the field of big data research. LI J and ZHANG Y lead with 19 publications each, followed by WANG L (18), and LI Y and LIU X (16 each). The top 10 authors, all publishing between 13 and 19 documents, highlight a concentration of academic output among a small group of researchers. Notably, all listed authors have Chinese-origin surnames, indicating China's leading role and geographic clustering in big data research. Overall, Figure 5 demonstrates that both individual productivity and regional concentration shape the research landscape in this field.

3.6 Most Relevant Affiliations

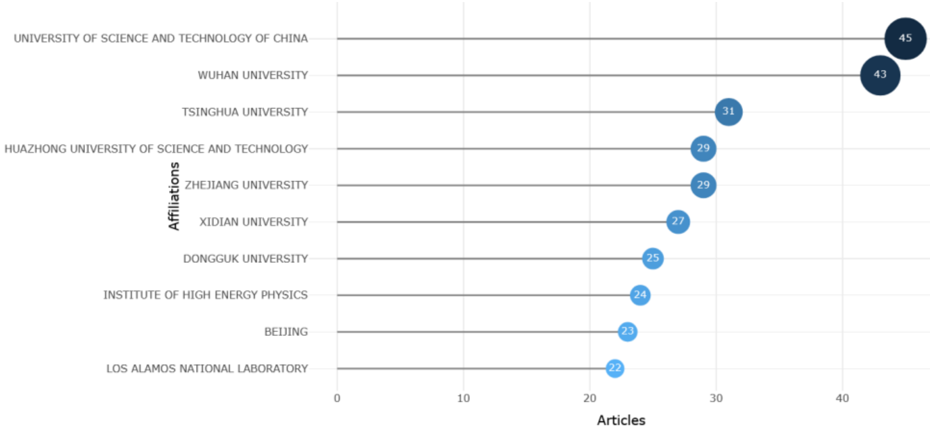


Figure 6. Most Relevant Affiliations

According to Figure 6, Chinese universities are the most active institutions in this field University of Science and Technology of China and Wuhan University are in the first two places. This is one of the factors that has contributed to China's rise in recent years. Only Los Alamos National Laboratory from USA is on the list, and there are no institutions from European or other Asian countries.

3.7 Countries' Scientific Production

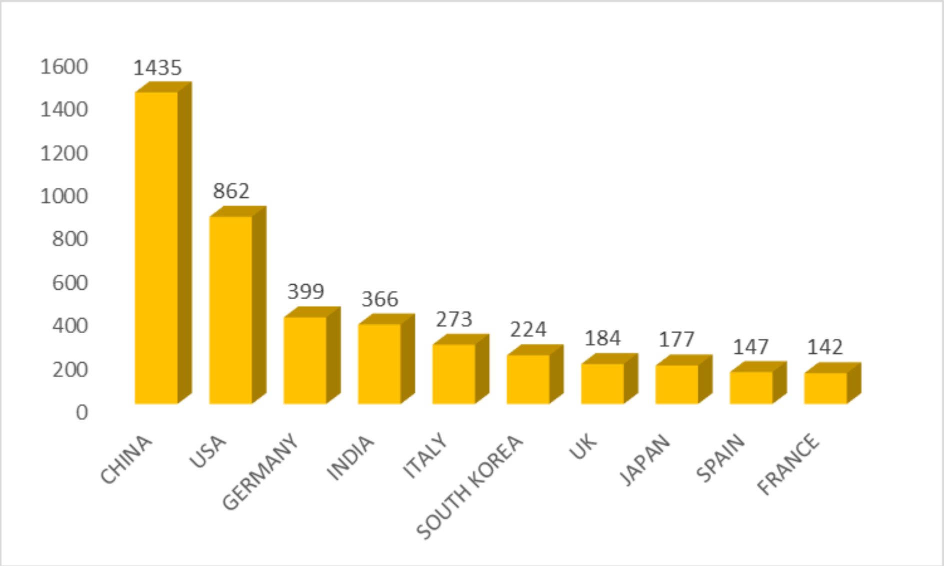


Figure 7. Top 10 Most Productive Countries in Big Data Research (2000–2025)

Figure 7 shows how big data publications are distributed among the top 10 most productive countries. China clearly leads with 1,435 publications, which can be attributed to its strong investments in big data and artificial intelligence. The United States follows with 862 publications, reflecting its solid academic infrastructure and technological leadership. Germany, India, and Italy demonstrate notable productivity, while South Korea, the United Kingdom, Japan, Spain, and France contribute with a more limited number of publications. Overall, **Figure 7** shows that big data research is not evenly distributed around the world; some countries lead the field, while others contribute at more modest levels.

3.8 Most Cited Countries

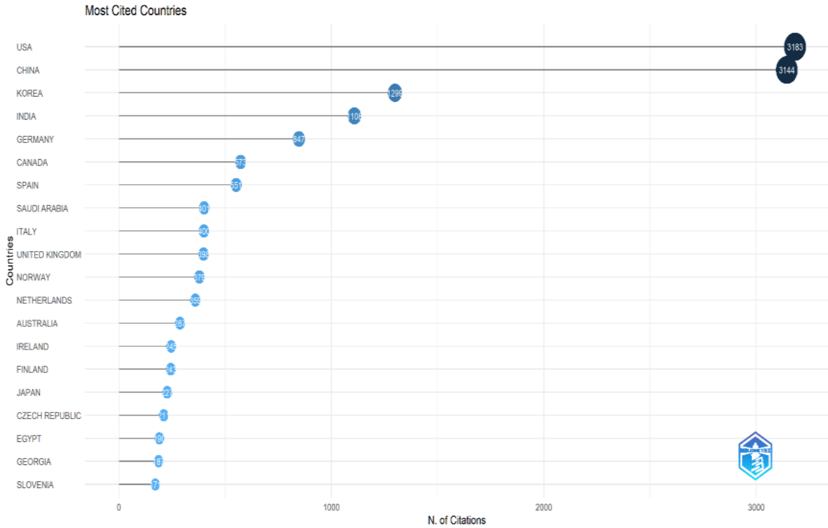


Figure 8. Most Cited Countries

In Figure 8, the United States and China are the most cited countries, with similar numbers, consistent with previous findings. South Korea ranks third here, down from previous rankings, indicating that its work has high impact value. Overall, three countries from Asia make the top five, with Europe being the most represented continent on the list.

3.9 Most Frequent Words

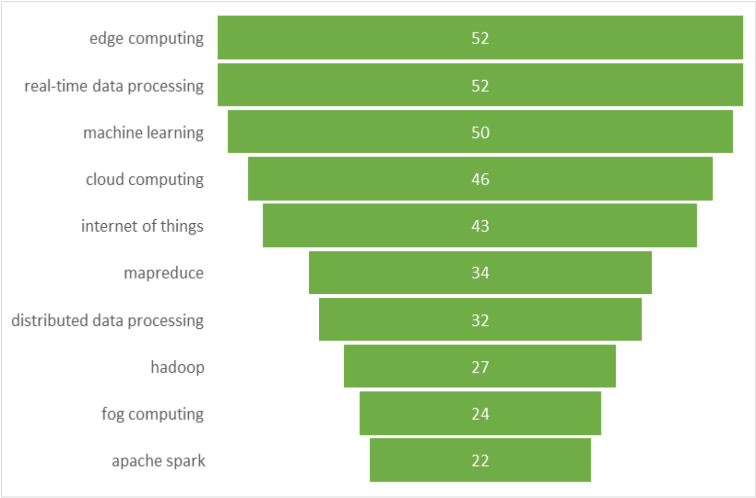


Figure 9. Most frequent terms in big data research and their occurrence counts (2000–2025).

Figure 9 shows the most frequently used terms in big data research and how often they appear. At the top of the list are “edge computing” and “real-time data processing,” which reflect the growing interest in processing information quickly and close to where it is collected. “Machine learning” appears often as well, showing that researchers are using smart systems to understand complex data. “Cloud computing” relates to storing and managing data through internet-based platforms. Other common terms like “internet of things,” “MapReduce,” “Hadoop,” and “Apache Spark” represent different tools and methods used in the field. Overall, **Figure 9** shows that big data research includes both well-established technologies and newer tools, highlighting how dynamic and diverse the field has become.

3.10 WordCloud

Figure 10. WordCloud

The word cloud in Figure 10 shows that Big Data Processing is a multidisciplinary field. While the most frequently occurring term is "Data Processing", terms such as "Real-Time Data Processing", "Internet of Things", "Machine Learning", "Deep Learning" and "Data Analytics" reveal the scope of this field. In addition, concepts such as "Information Management", "Decision Making", "Network Security" and "Digital Storage" show that the process includes not only software but also hardware and security dimensions.

3.11 Co-occurrence Network

Figure 11. Co-occurrence network of keywords in big data research (2000–2025)

Figure 11 shows how frequently used keywords in big data research are grouped based on their co-occurrence in the same documents. Terms that appear together more often are positioned closer in the network. "Big data" and "machine learning" appear at the center, indicating their central role in the field. They form a dense cluster with terms such as "real-time data processing," "edge computing," "cloud computing," and "internet of things," which are frequently mentioned together. Technical terms like "MapReduce," "Hadoop," and "Apache Spark" are found in more specific research areas, while some other keywords appear on the periphery, reflecting their limited usage. Overall, the network illustrates how key concepts in big data research are interconnected and highlights the main thematic structures shaping the field

3.12 Thematic Map

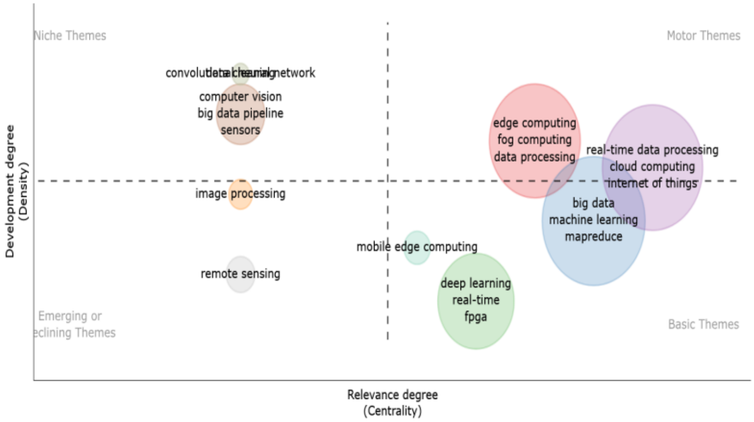


Figure 12. Thematic Map

Figure 12 shows the thematic map related to the subject. Concepts such as real-time data processing, cloud computing, edge computing, fog computing and IoT stand out as mature and highly important engine themes. Basic concepts such as big data, machine learning and MapReduce are basic themes that are open to development. While image processing and remote sensing are among the developing themes, topics such as convolutional neural networks, computer vision, big data pipeline and sensors are niche themes that may come to the fore in the future, despite their limited interest.

3.13 Author's Collaboration Network

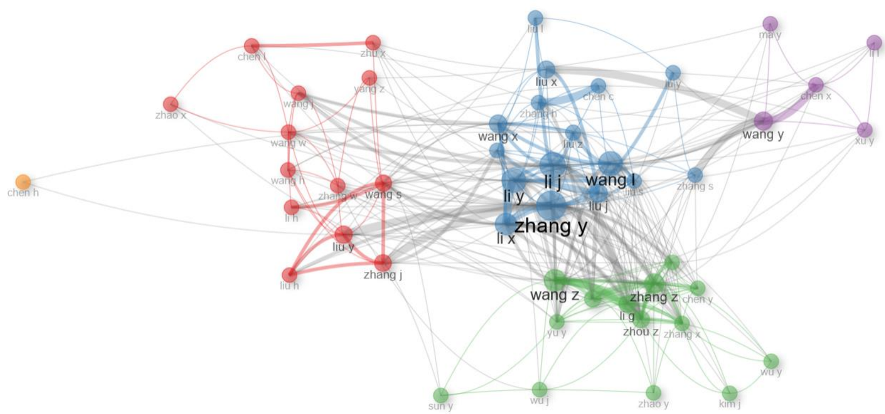


Figure 13. Author collaboration network in big data research (2000–2025)

Figure 13 illustrates the collaboration network among researchers in the field of big data. Central authors such as LI J, ZHANG Y, WANG L, LI Y, and LIU X show strong connections both within their group and with other researchers across the network. Authors positioned closer together represent tightly connected research groups, while those on the periphery tend to have fewer collaborations or weaker links to the broader research community. Overall, Figure 13 suggests that academic output in big data is structured around a few strong collaboration groups, with certain authors playing a central and connecting role within the research network.

3.14 Countries

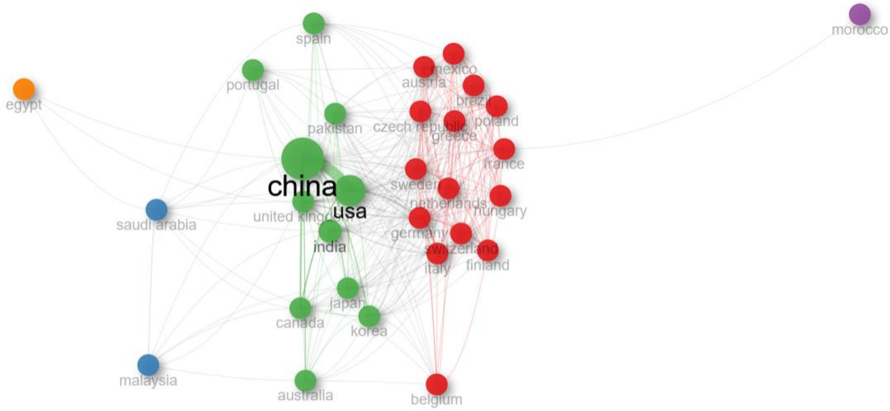


Figure 14. International Collaboration Map

As shown in **Figure 14**, the countries with the most extensive networks are depicted. The green countries, such as China and the United States, are the primary nations in this field. Countries like India and Pakistan, which are significantly developing in the software sector, are also represented. Additionally, countries like Spain and Portugal from the European Union, although part of the broader group of core countries, work closely with the primary nations but differ from other countries within the red zone, such as Saudi Arabia.

The red countries primarily consist of European Union nations that have formed networks with each other. Among these, Belgium occupies a slightly distinct position, while France stands out as the only country collaborating with Morocco. Egypt, while collaborating with countries in the green and blue zones, remains relatively isolated from international cooperation.

4. DISCUSSION &COMMENTS

The field of Big Data Preprocessing Techniques on Data Analytics is generally interdisciplinary and has been explored by various countries. In recent years, the number of studies in this area has been increasing. Publications related to this field have been increasing in recent years, and in parallel with this, there has been an increase in citations. The competition and innovation between America, China and Germany, which are advanced countries in technology and R&D, is also encountered in the analyses conducted to follow the developments in the field of big data. China, in particular, stands out as the dominant country in

R&D activities in the field of big data, as in other R&D and technology fields. (Zhang, et al, 2020)

Studies on big data mainly focus on edge computing, real-time processing, machine learning and cloud computing. The prominence of concepts such as edge computing, real-time processing, machine learning and cloud computing, which form the infrastructure of artificial intelligence, is an indicator of the development of artificial intelligence and generative artificial intelligence, which are rapidly growing and used in many fields today (Hossain et al, 2023, Berigel at al,2025). The developing trends in Big Data Preprocessing Techniques on Data Analytics are convolutional neural networks, computer vision, big data pipeline, and sensors. This situation shows us new trends in big data preprocessing processes with new approaches. In this field, where China and the USA are dominant, studies and collaborations are observed more and more in these countries and with each other. As a result, Big Data Preprocessing Techniques on Data Analytics is a multidisciplinary and developing field, and following the study areas of trends and examining their historical development will make significant contributions to this field.

5. REFERANCES

Aria, M. & Cuccurullo, C. (2017). *bibliometrix: An R-tool for comprehensive science mapping analysis*. Journal of Informetrics, 11(4), 959–975.

Berigel, D. S., & ŞILBIR, L. (2024). A BIBLIOMETRIC ANALYSIS OF AI LITERACY: TRENDS, TOPICS AND FUTURE DIRECTIONS. *Near East University Online Journal of Education*, *7*(1), 42-53.

Berigel, D. S., Şilbir, L., & Şilbir, G. M. (2025). Integrating Artificial Intelligence (AI) Into Technical and Vocational Education and Training (TVET): A Prisma-Based Systematic Review. *Calitatea Vieții*, *36*(1), 1-27.

Chen, C. P. & Zhang, C. Y. (2014). *Data-intensive applications, challenges, techniques and technologies: A survey on Big Data*. Information Sciences, 275, 314–347.

Diodato, V. & Gellatly, P. (2013). *Dictionary of bibliometrics*. Routledge.

Gahi, Y., Guennoun, M. & El-Khatib, K. (2016). *Big data analytics: Security and privacy challenges*. 2016 IEEE Symposium on Computers and Communication (ISCC), 952–957.

García, S., Luengo, J. & Herrera, F. (2015). *Data preprocessing in data mining*. Springer.

García, S., Ramírez-Gallego, S., Luengo, J. & Benítez, J. M. (2016). *Big data preprocessing: Methods and prospects*. Big Data Analytics, 1(1), 1–22. (metin içi atıfı yok)

Golov, N., & Rönnbäck, L. (2017). Big data normalization for massively parallel processing databases. *Computer Standards & Interfaces*, *54*, 86-93.

Hossain, M. E., Tarafder, M. T. R., Ahmed, N., Al Noman, A., Sarkar, M. I. & Hossain, Z. (2023). *Integrating AI with Edge Computing and Cloud Services for Real-Time Data Processing and Decision Making*. International Journal of Multidisciplinary Sciences and Arts, 2(4), 252–261.

Hu, H., Wen, Y., Chua, T.-S. & Li, X. (2014). *Toward scalable systems for big data analytics: A technology tutorial*. IEEE Access, 2, 652–687.

Laney, D. (2001). *3D data management: Controlling data volume, velocity, and variety*. META Group.

Li, X., Lin, A., Young, C.-H., Dai, Y. & Wang, C.-H. (2019). *Energetic and economic evaluation of hybrid solar energy systems in a residential net-zero energy building*. Applied Energy, 254, 113709.

Mayer-Schönberger, V. & Cukier, K. (2014). *Big data: A revolution that will transform how we live, work, and think*. John Murray.

O'Leary, D. E. (2015). *Big data and privacy: Emerging issues*. IEEE Intelligent Systems, 30(6), 92–96.

Oussous, A., Benjelloun, F. Z., Lahcen, A. A. & Belfkih, S. (2018). *Big data technologies: A survey*. Journal of King Saud University – Computer and Information Sciences, 30(4), 431–448.

Prakash, A., Navya, N. & Natarajan, J. (2018). *Big data preprocessing for modern world: opportunities and challenges*. In *International Conference on Intelligent Data Communication Technologies and Internet of Things* (pp. 335–343). Cham: Springer International Publishing.

Yaqoob, I., Hashem, I. A. T., Gani, A., Mokhtar, S., Ahmed, E., Anuar, N. B. & Vasilakos, A. V. (2016). *Big data: From beginning to future*. International Journal of Information Management, 36(6), 1231–1247.

Zhang, H., Yang, J., Ren, X., Wu, Q., Zhou, D. & Elahi, E. (2020). *How to accommodate curtailed wind power: A comparative analysis between the US, Germany, India and China*. Energy Strategy Reviews, 32, 100538.

Chapter 5

Tram Line Works in Batman

Ali Osman GENÇ[1] Ömer Faruk SEVİM[2], Halil İbrahim ERKAN[3], Seher DALGIÇ[4], Mevlüt KAÇMAZ[5], Abdulrezzak BAKIŞ[6]

1. WHAT IS A TRAM?

Tram systems are a rail public transportation vehicle that is frequently preferred in urban transportation. They are generally used for short and medium distances, in densely populated areas with high traffic congestion. The characteristics of trams may vary depending on their design and intended use. In general, however, most tram systems share some basic elements in common. More detailed characteristics of tram systems are as follows:

1.1 Rail System (Rail Transportation)

- Movement on Rails: Trams are vehicles that move on rails. This allows the tram to travel on a more stable and straight line. The rails are independent of vehicles on the roadway, so they are not affected by problems such as traffic jams.

- Rail Lines: Trams have specific routes. These routes are determined to cover the city and are usually organized as one-way or two-way tracks. The rails can be underground, above ground or at ground level. (Şahin, R. and Yılmaz, A.)

1

2

3

4

5

6 Assoc. Prof. Dr.

1.2 Working with Electricity (Electric Trams)

- Electric Based Movement: Most modern trams are powered by electricity. Electricity is usually supplied by energy from the city electricity grid. The tram moves by running its engines with the energy it receives over the power lines.

- Environmentally Friendly: Electric trams are environmentally friendly as they do not use fossil fuels. As a means of transportation that reduces carbon emissions and air pollution, it is an important transportation solution for sustainable cities. (Özdemir, E.)

1.3 Fixed Lines and Stops

- Fixed Routes: Trams operate on fixed routes. That is, a tram travels on a specific route and this route is usually fixed. The routes are determined to pass through densely populated areas in the city.

Stops Stopping points of trams, i.e. stops, are usually located in a visible and accessible way. They are usually located in areas with dense crowds of people, such as pedestrian paths, shopping centers or bus terminals. Stops can often have digital displays and timetables or ticketing systems.

- Express boarding and alighting: Stops are designed for fast boarding and alighting. This speeds up passenger traffic, especially during rush hours. (Karakas, M., & Aydin, N.)

1.4 High Carrying Capacity

- **High Passenger Capacity:** Trams are generally vehicles with a high carrying capacity. A tram can carry between 50 and 300 passengers. This capacity plays a major role in reducing traffic congestion in the city and making public transportation systems more efficient.

- Large Capacity Vehicles: Modern trams are usually equipped with long wagons and large interior volumes. This makes a tramway suitable for both short-distance and long-distance travel. (Çetin, R., & Aydın, N.)

1.5 Low Emission and Environmentally Friendly Transportation

- Zero Emission Systems: Electric trams are environmentally friendly means of transportation as they emit no carbon emissions. This contributes to reducing air pollution, especially in big cities.

- Zero Noise: Electric trams are silent in their internal mechanisms, an important feature that reduces urban noise pollution.

1.6 Fast Transportation in Slow Traffic

- Independent of Closing Traffic: Trams are independent of traffic as they usually run on separate lines. This allows trams to travel faster than road transportation. However, some tram lines may have intersections with the highway, which means that traffic may be affected from time to time.

- Short Duration Stops: Stops are usually limited to short stops, which increases the speed of the tram.

1.7 Efficient Transportation and Low Operating Costs

- Low Operating Costs: Trams are economical transportation vehicles with their high carrying capacity and low fuel costs. Especially the operating costs of electric trams are lower than fossil fuel vehicles.

- High Efficiency: Trams provide more regular and efficient transportation by operating on fixed lines. They consume less fuel than road transport vehicles.

1.8 Appropriate Speed and Safety

- **Safe Transportation at Low Speed:** Trams generally travel at low speed (around 20-40 km/h) in urban transportation. This speed makes trams a safe means of transportation for passengers. The risk of traffic accidents is lower.

- **Speed Restrictions:** Most tram lines are subject to certain speed limits, taking into account safety standards and pedestrian density in the city. (Kucuk, M., & Özdemir, E.)

2. ENVIRONMENTAL AND ECONOMIC ADVANTAGES OF THE TRAM

2.1 Environmental Advantages

2.1.1 Low Carbon Emission (Zero Emission)

Electric Transportation Vehicle: Trams do not emit carbon emissions as they are generally powered by electricity. Compared to fossil fuel vehicles, trams make a major contribution to reducing carbon dioxide (CO_2) levels in the atmosphere. They offer an important solution to combat air pollution, especially in city centers.

Green Transportation: Electric trams provide zero emissions, making them an environmentally friendly means of transportation. Their carbon footprint is much lower than fossil fuel powered vehicles.

2.1.2 Reduced Air Pollution

Tackling Air Pollution: Trams do not create air pollution compared to motorized vehicles. Electric trams clean the city's air, especially by removing air pollutants such as smog, exhaust gases and carbon monoxide within the city. This helps reduce health problems such as respiratory diseases and cancer.

Improved Air Quality in Cities: The use of electric systems improves air quality in cities and has a positive impact on public health.

2.1.3 Noise Pollution Reduction

Quiet Operation: Electric trams run much quieter than fossil fuel powered vehicles. This reduces noise pollution, especially in urban areas. Quiet operation makes the city more livable and provides psychological relief.

Low Noise Level: Trams allow people living in cities to live in a quieter environment, as they generally have low noise levels. This is a great advantage where noise pollution within the city is a serious problem.

2.1.4 Energy Efficiency

Electricity Consumption and Efficiency: Trams are an energy efficient means of transportation as they run on electricity. Trams become much more environmentally friendly, especially if electricity can be obtained from renewable energy sources (solar, wind, hydroelectric).

High Efficiency: Trams travel at lower energy costs because they are electric. Also, the energy consumption per passenger is much lower compared to other means of transportation such as buses and cars.

2.1.5 Protection of Natural Areas

Reduced Number of Vehicles: Trams relieve traffic, especially within the city. Reduced traffic density allows for the preservation of green areas. It also helps to conserve natural resources with less fossil fuel consumption.

Sustainable Cities: Trams have an important role in creating sustainable cities. Because, by reducing traffic congestion and air pollution, they contribute to making cities healthier and more sustainable. (Şahin, M., & Kucuk, H.)

2.2 Economic Advantages

2.2.1 Low Operating Cost

Low Maintenance and Operating Cost: Electric trams are low maintenance systems. Vehicles running on tram lines generally have fewer breakdowns and lower maintenance requirements. This significantly reduces the operating costs of trams.

Energy Efficiency: By offering high energy efficiency, electric trams also reduce energy costs in the long term. This allows both cities and tram operators to make economic gains.

2.2.2 High Passenger Capacity

Capacity Advantage: Trams are vehicles with a high carrying capacity. A tram can carry between 100 and 300 passengers, which significantly reduces traffic congestion. This high passenger capacity not only allows more passengers to be transported but also reduces the cost per passenger transported by a single vehicle.

Consistently High Demand: In areas close to city centers, trams often experience consistently high demand, which allows tram systems to operate more efficiently and generate revenue.

2.2.3 Reducing Traffic Congestion

Reducing Traffic Congestion: Tram systems do not interact or have limited interaction with road traffic as they operate on dedicated rail lines. This significantly reduces traffic congestion within the city, thereby shortening the transit times of vehicles and preventing the economic losses caused by traffic congestion.

Time Savings: Reduced traffic congestion reduces the time it takes people to get to work, which increases labor productivity. In addition, trams provide fast and regular transportation, reducing travel times and saving time.

2.2.4 Low Investment in Transportation Infrastructure

Low Investment in Infrastructure: Tram systems generally require lower cost investments than metro systems. This is because metro systems are much more expensive projects that need to be built underground. Trams, on the other hand, can be easily integrated with above-ground lines, so construction costs are lower.

Low Total Cost: The construction and operating costs of trams can be more economical in the long run because they consume less energy, require less maintenance and provide more efficient transportation.

2.2.5 Job Creation and Economic Growth

New Employment Opportunities: Building and operating tram systems creates new jobs. This provides employment opportunities, particularly in areas such as rail engineering, maintenance, ticketing and security.

Urban Economic Revitalization: Areas where trams pass through can experience growth in sectors such as commerce and housing. As tram lines generally make the city more accessible, economic vitality increases in these areas. (Şahin, M., & Karaca, T.)

3. SUCCESSFUL TRAM PROJECTS IN THE WORLD

Many cities around the world have invested heavily in tram systems in their transportation infrastructure and these projects have helped make urban transportation more efficient, environmentally friendly and economical. Successful tram projects not only improve the efficiency of public transportation, but also reduce environmental impacts, alleviate traffic congestion and increase the economic vitality of cities. Here are some successful tram projects around the world:

1. Porto (Portugal) - Metro do Porto

Porto's tram system is actually a light rail metro system. Metro projects started in the 2000s, aiming to make the city more accessible and environmentally friendly.

Success Factors:

Environmentally Friendly: Low emission as it works with electric system.

High Transportation Capacity: Porto's tram system carries hundreds of thousands of passengers every day.

Efficiency: Fast and regular services have significantly reduced traffic congestion.

Urban Accessibility: Fast connections between the city center and the suburbs.

Highlights: Porto's tram network has made the city greener and more sustainable. It has also become an important means of transportation for urban tourism.

2. Tram System, Zurich (Switzerland)

Zurich has one of the oldest tram systems in the world. The city's tram network dates back to 1882 and has become the backbone of Zurich's transportation infrastructure.

Success Factors:

Continuous Renovation: Zurich is constantly modernizing its tram system. New tram vehicles are equipped with energy efficient and environmentally friendly technologies.

System Integration: The Zurich tram system has an integrated transportation network with bus, train and ship transportation. This integration makes it possible for passengers to easily transfer and provide fast and cheap transportation.

High Utilization Rate: In Zurich, trams carry hundreds of thousands of passengers on a daily basis and are the most preferred means of public transportation.

Highlights: Zurich is notable for the continuous expansion of its tram lines and new projects to make them more environmentally friendly.

3. Melbourne (Australia) - Melbourne Tram Network

Melbourne is one of the cities with the largest tram network in the world. Launched in 1905, Melbourne trams play an important role in urban transportation today.

Success Factors:

World's Largest Tram Network: Melbourne has 2500 km of tram lines, reaching every part of the city.

High Passenger Capacity: Melbourne's trams have one of the highest carrying capacities in the world.

Environmentally Friendly: The electric trams provide zero emissions and do not harm the environment.

Continuous Renovation: The city is constantly modernizing its tram vehicles and lines and equipping them with new technologies.

Highlights: Melbourne trams are the main backbone of the city's transportation. They are also a popular option for sightseeing.

4. Düsseldorf (Germany) - Rheinbahn Tram Network

The Rheinbahn tram network in Düsseldorf is one of the most efficient tram systems in Germany and one of the main means of transportation in the city.

Success Factors:

High Integration: Rheinbahn trams are integrated with the bus and rail systems. This provides faster and more economical transportation in the city.

Environmentally Friendly Transportation: Trams are environmentally friendly as they run on electric power, reducing air pollution in the city.

Intensive Use: Düsseldorf carries hundreds of thousands of passengers daily, making trams the most preferred means of transportation.

Highlights: In Düsseldorf, trams provide improved integration into the city center and make the city environmentally friendly. In addition, the economic revitalization around the tram lines is a great advantage.

5. Istanbul (Turkey) - Istanbul Tram (T1 Line and Other Line Projects)

Istanbul has invested heavily in its tram system in recent years. With the T1 line and other new line projects, Istanbul has a more extensive tram network than before.

Success Factors:

Continuous Expansion: With the continuous expansion of the tram network in Istanbul, transportation has become more comfortable.

Traffic Relief: The tram line has helped to solve traffic and congestion problems, especially in the inner city.

Transportation Integration: Trams are integrated with metro and bus lines, allowing passengers to use more than one means of transportation.

Environment Friendly: Trams run on electric energy, contributing to the environment by reducing the city's carbon emissions.

Highlights: Tram projects in Istanbul offer an important solution in urban transportation, combining both historical and modern infrastructures.

6. Canberra (Australia) - Canberra Tram Network

Launched in 2019, the Canberra Tram Network is Australia's first modern tram system.

Success Factors:

Zero Emissions: The trams are electric and do not harm the environment.

Modern and Green Transportation: Canberra has strengthened its green city vision by offering environmentally friendly, sustainable transportation options.

Intra-City Connections: The tramway has alleviated traffic problems by speeding up transportation between the city center and the suburbs.

Highlights: This system sets an example for other cities looking to create sustainable cities.

7. Dubai (United Arab Emirates) - Dubai Tram System

Dubai launched its first tram system in 2014. The city is one of the fastest growing metropolises in the world and uses trams effectively for urban transportation.

Success Factors:

Environmentally Friendly: Electric trams provide zero emissions and keep the city cleaner.

Modern Technology: The Dubai tram system is equipped with state-of-the-art smart systems. **Luxury Transportation:** Dubai trams operate at high speed and with luxurious designs, making urban transportation comfortable.

Highlights: Dubai provides efficient transportation by integrating its tram network with other transportation systems in the city.(Pojani, D., & Stead, D.) (Ceder, A., & Wilson, N. H.)

4. TRAM TRAM PROJECTS IN TURKEY

In Turkey, tram applications have emerged as an important solution for urban transportation in recent years. In both large cities and medium-sized cities, tram systems have become an important component of public transportation. These projects have been implemented to reduce traffic congestion, promote environmentally friendly transportation and support economic development. Most of the successful tram projects in Turkey have been built as electric rail systems. Here are some important tram applications in Turkey:

4.1 Istanbul Tram

T1 Line (Bağcılar - Kabataş)

- About the Project: Istanbul's best known tram line, the T1 line, was launched in 1992 and was the first modern tram line in Istanbul. The T1 line runs from Bağcılar to Kabataş and offers an integrated transportation network with Marmaray.

- Success Factors:

o Intensive Use: The T1 line is one of the highest ridership lines in Istanbul. It carries hundreds of thousands of passengers daily.

o Traffic Reduction: Significant traffic relief has been achieved on the route from Kabataş to Bağcılar.

o Sustainable Transportation: As an electric system, the T1 line promotes environmentally friendly transportation.

o Integration: It works integrated with metro, metrobus, bus and sea transportation, which allows for transfer.

T2 Line (Topkapi - Masjid-i Selam)

- About the Project: Istanbul's T2 line is a tram line that started operations in 2011, providing transportation mainly between Topkapı and Masjid-i Selam. It connects the city from east to west and covers the Emin Ali Pasha neighborhood.

- Success Factors:

o Accessibility: This line is close to the city center and has close connections to tourist attractions.

o Reduced Travel Time: It speeds up transportation without wasting time in traffic.

T3 Line (Kadikoy - Moda)

- About the Project: The T3 line in Kadıköy was launched in the early 2000s to provide short distance transportation, especially between Kadıköy and Moda.

- Success Factors:

o Short Distance Transportation within the City: Provides easy access to the dense area of Kadıköy.

o High Passenger Capacity: It plays an important role in transporting the large population in and around Kadıköy.

New Tram Lines and Expansion:

- In Istanbul, the Istanbul Metropolitan Municipality is constantly expanding its tram lines. For example, tram line projects are underway in new areas such as Emin Ali Paşa and Beylikdüzü. In addition, new tram lines are planned to provide transportation to areas such as Sultanbeyli, Üsküdar and Kartal. Ministry of Transportation and Infrastructure.

4.2 Eskisehir Tram

- About the Project: Eskişehir's tram system is notable for Turkey's first locally built tram. The tram network in the city became operational in 2004. The Eskişehir tram line connects different parts of the city and is one of the busiest transportation lines within the city.

- Success Factors:

o Domestic Production Trams: Eskisehir was one of the first cities to produce domestic trams. These trams are produced by Tramvay A.Ş. and contribute to the economy.

o Efficient Transportation Network: The tram network in the city connects dense areas, especially Eskişehir Anadolu University and the city center.

o Environmentally Friendly: Electric trams reduce air pollution and provide environmentally friendly transportation.

o Tourism and Education: Eskişehir is a dense city, especially in terms of student population and tourism, and the tram provides easy transportation for these masses.

4.3 Konya Tram

- About the Project: Konya's tram system was started in 1992 by Konya Metropolitan Municipality and continues to expand today. The tram line in Konya extends from the center of the city to Selçuklu district.

- Success Factors:

o High Passenger Capacity: The tram network in Konya carries thousands of passengers every day.

o System Integration: The Konya tram line is integrated with buses and other public transportation systems, making transportation more efficient.

o Sustainable Development: As an electric system, trams meet the need for public transportation without harming the environment.

4.4 Antalya Tram

- **About the Project:** Antalya launched its tram system in the 2010s. The Antalya Tram Line connects key points in the city such as Karaalioglu Park and the Expo 2016 area.

- Success Factors:

o **Tourist Transportation:** Provides easy access to Antalya's touristic attractions and is an important transportation solution, especially between accommodation and tourism areas.

o Travel Time Reduction: In Antalya, where traffic density is high, trams provide fast transportation, especially to the city center.

o **Environmentally Friendly Transportation:** Electric trams reduce air pollution in the city and provide zero emissions.

4.5 Kayseri Tram

- About the Project: Kayseri is one of Turkey's medium-sized cities, which launched its tram system in 2009. The Kayseri Tram Line connects the city center with important locations such as the Organized Industrial Zone (OIZ).

- Success Factors:

o Intensive Use: Trams in Kayseri are the busiest public transport in the city, carrying hundreds of thousands of passengers every day.

o Efficient and Fast: Kayseri has chosen the tram system to solve its traffic problems and trams provide fast transportation.

o High Passenger Capacity: Trams play an important role in transporting the dense population of the city.

4.6. Bursa Tram

- About the Project: Bursa started to expand its tram system in the early 2010s. The tram line in Bursa has become an important means of transportation under the name Bursa Rail System.

- Success Factors:

o Modern Lines: Modern tram lines and locally produced trams are used in Bursa.

o Urban Connections: Different parts of Bursa are easily connected to each other thanks to tram lines. This system provides ease of transportation especially between the industrial zone and the city center.(Çelik, H., & Yıldız, İ)

5. SUGGESTIONS FOR TRAM LINE IN BATMAN

Establishing a tram system in Batman has a number of advantages in terms of solving transportation problems in the city and increasing environmental sustainability. The main advantages of establishing a tram in Batman are:

5.1 Reducing Traffic Congestion

Traffic relief on main arteries: Since Batman is a growing city, there is heavy traffic congestion on the main arteries within the city. Since the tram operates on special rail lines, it does not interact with road traffic, which leads to a decrease in traffic congestion.

Fast and regular transportation: Trams largely solve the transportation problems in the city by carrying passengers on fast and regular trips, regardless of traffic congestion.

5.2 Environmentally Friendly Transportation

Zero emissions: Electric trams provide zero emissions since they do not use fossil fuels. This improves the air quality in Batman and reduces environmental pollution in the city.

Reducing air pollution: Especially the exhaust gases of motor vehicles in the city cause respiratory diseases. Trams positively impact public health by reducing carbon emissions and air pollution.

5.3 High Carrying Capacity

Large passenger capacity: Trams can carry 100-300 passengers at a time. This reduces the density in the city and allows the public transportation system to operate efficiently.

Efficiency during rush hours: During rush hours in Batman, trams can carry more passengers at the same time, which increases the efficiency of public transportation.

5.4 Fewer Traffic Accidents

Predictable routes: Trams reduce the risk of traffic accidents because they operate on fixed lines. Especially in heavy traffic, accidents with buses and personal vehicles may be more common due to the absence of tram lines.

Pedestrian safety: Trams allow pedestrians to cross more safely, especially in residential areas. The reduction in traffic accidents can increase safety in the city.

5.5 Faster and More Comfortable Transportation

Time saving: Reducing the time lost due to traffic congestion can make citizens' daily lives, such as work and school, more efficient. Trams provide fast transportation, especially for urban distances.

Comfortable travel: Trams are generally more comfortable than buses, offering travel without congestion and discomfort.

5.6 Sustainable Transportation Infrastructure

Long-term sustainability: Trams are a sustainable transportation solution in the long term. Being electric and requiring low maintenance, trams offer an economical and environmentally friendly transportation solution.

Integration with renewable energy sources: Electricity can be provided from renewable sources (solar, wind, hydroelectric). Thus, trams become a completely environmentally friendly means of transportation.

5.7 Transportation Integration

Integration of public transportation networks: Trams can provide uninterrupted transportation by integrating with existing bus and minibus lines. If the existing public transportation infrastructure in Batman is integrated with the tram, passengers can travel faster and more efficiently by transferring to different vehicles.

Advantage in tourist transportation: By establishing a tram line in areas where tourist areas are dense in Batman, tourists can travel around the city more easily. In addition, trams can make transportation more comfortable for tourists.

Considering all these conditions, the tram line routes that may be suitable for the city of Batman are visualized and given below with their reasons.

5.8 Tram Routes

Figure 5.1 Between the airport and the city centre

Batman Airport was opened to civil air traffic in 2010 and is an important transportation point contributing to the international air transportation of Batman province. This airport, which is one of the largest transportation infrastructures in the city, plays an important role especially for regional flights and domestic flights.

Passenger Capacity: Although Batman Airport has an annual passenger capacity of approximately 1 million, the airport serves mostly domestic flights.

Terminal Capacity: Batman Airport's current terminal was designed to meet this capacity, but new projects and capacity-increasing works are being carried out in line with increasing passenger demands over time.

The annual passenger number at Batman Airport generally ranged between 300,000 and 500,000 in the early 2020s. However, these figures may vary from year to year and there have been ups and downs in some years, considering the effects of the pandemic.

2021 Passenger Number: Batman Airport served approximately 400,000 passengers as of 2021.

Number of Passengers in 2022: This number increased in 2022 and approached 500,000. URL-1 URL-2 URL-3

Considering this data, the tram line should definitely start from the airport.

Figure 5.2 Bus Terminal-City Center-Municipality-Hasankeyf Road Route

Due to the increasing population in Batman, the currently used public transportation may be insufficient. In order to prevent this, connecting the places such as bus stations, governorships and municipalities that are used intensively by the citizens in the city center with a tram line will increase the popularity of the line.

Figure 5.3 Hasankeyf road-Stadium-University Route

The other stops where we plan for the tram line to pass are the Hasankeyf road-Stadium-University route, which will pass through the stadium where the matches of the city's growing team advancing towards the Super League are played, and will go all the way to the University, which will greatly ease transportation for the citizens.

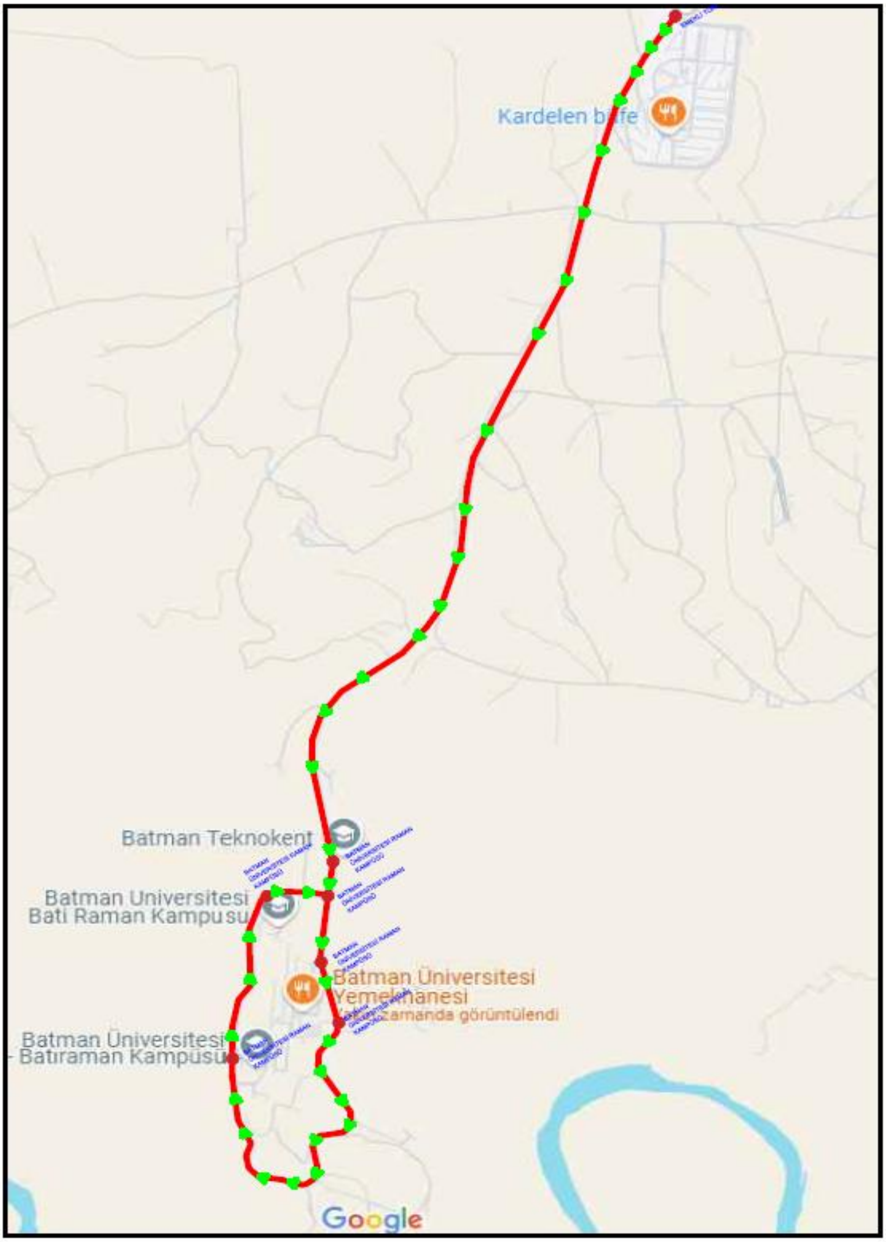


Figure 5.4 University Route

Finally, it is thought that the tram line, which is planned to start from the airport and end at the university, will be preferred by university students because it passes through the city center and has stops close to student dormitories.

6. APPROXIMATE COST OF TRAM LINE IN BATMAN

The cost of installing a tram line in a city like Batman may vary, similar to the cost of tram lines in other cities around the world. However, Batman's specific conditions (city size, infrastructure requirements, geographical challenges, construction costs, etc.) may affect these costs. Below are some of the main factors and approximate cost ranges that may affect the cost of installing a tram line for Batman.

6.1 Factors Affecting Batman Tram Line Cost:

6.1.1 Length of Line:

The length of the tram line is one of the most important factors that directly affects the cost. The length of a potential tram line in Batman may vary according to the city's transportation capacity.

Rail and Infrastructure Requirements:

Batman's urban structure, population density, and ground conditions may affect the requirements for track laying and infrastructure. While the cost may be lower in flat areas, complex infrastructures (e.g. viaducts, underground passages) will increase the cost.

Electrification and Energy Infrastructure:

Since trams are electric, infrastructure work requires the installation of power lines and power systems. This creates additional costs.

Stations and Stops:

The construction of each tram station or stop also affects the cost. In a growing city like Batman, the number and size of the stations can be determined by the frequency of use of the line.

Tram Vehicles:

The cost of the vehicles to be used in the tram system generally varies according to the size and capacity of the line. In Batman, 5-10 tram vehicles may be sufficient for a medium-sized tram system.

Intra-City Traffic and Expropriation:

In developing cities like Batman, expropriation may be required in some regions for rail laying and infrastructure work. This can create additional costs.

Labor and Construction Costs:

The construction labor costs in the city and the materials to be used in infrastructure construction also affect the cost. Labor costs in Batman may be lower compared to large metropolitan cities, but these regional differences still affect the total cost.

Cost Ranges:

The cost of a tram line is usually calculated per kilometer. Based on examples from Turkey and around the world, the average cost ranges for a 1 kilometer tram line installation can be as follows:

Cost of a Simple and Short Line:

The cost of a short and straight line can range from approximately 10 million USD to 20 million USD.

Mid-Level Tram Lines:

In a developing city like Batman, the cost of a medium-sized line can range from 20 million USD to 40 million USD.

Advanced and Integrated Systems:

The cost of a tram line using advanced systems such as electrification, smart ticketing systems, and energy-efficient trams can exceed 40 million USD. (URL-4)

Sample Calculation (for 1 km):

Including infrastructure, rails, stations, energy systems, and vehicles, the installation cost of a 1 kilometer tram line in Batman can be estimated to be approximately 15 million USD to 30 million USD.

Long-Term Economic Contributions:

Major infrastructure projects such as tram lines should not only consider their initial costs, but also their long-term economic contributions. Such projects can reduce traffic congestion, protect the environment, and create a more sustainable transportation network.

Factors such as time saving, fuel saving, and labor efficiency also provide indirect economic benefits. In addition, it should not be forgotten that the tram line will contribute to trade and tourism activities in the city.

7. CONCLUSION

The tram line project in Batman stands out as a strategic infrastructure investment that will enable the city to take a step towards a sustainable future in terms of both environmental and economic aspects. The advantages of tram systems such as low emissions, high carrying capacity and reducing traffic congestion were discussed in detail in the study. Considering Batman's increasing population and traffic density, it is anticipated that the tram system will greatly facilitate urban transportation and increase the quality of life.

In particular, the route proposal connecting the airport, city center and university aims to both meet the needs of the local population and increase the touristic potential. The figures presented in the cost analysis of the project show that the tram line can amortize itself in the long term by providing economic and environmental benefits.

As a result, the installation of a tram line in Batman will not only solve transportation problems, but will also make the city more livable, environmentally friendly and modern. This project is an important step for sustainable urbanization.

8. REFERENCES

Şahin, R. and Yılmaz, A. Tram and Light Rail Systems: Fundamentals and Applications. Transportation Technologies Publications, 2017.

Özdemir, E. (2021). Environmental Impacts and Future of Electric Tram Systems. Transportation Technologies Journal

Karakas, M., & Aydın, N. (2017). Efficiency and Station Design of Urban Rail Systems. Urban Planning Journal

Çetin, R., & Aydın, N. (2019). The Role of Large Capacity Vehicles in Urban Transportation Systems. Urban Planning and Transportation Journal

Kucuk, M., & Özdemir, E. (2019). Environmentally Friendly Transportation Systems: Trams and Electric Transportation Technologies. Transportation Technologies Publications.

Şahin, M., & Kucuk, H. (2019). Green Transportation: Electric Systems and Zero Emission Technologies. Environmental Publications.

Sahin, M., & Karaca, T. (2019). Economic Impacts of Trams and Urban Transportation. Environmental and Transportation Publications.

Pojani, D., & Stead, D. (2015). "The Role of Public Transport in Achieving Sustainability in Cities: The Case of Tram Systems." Urban Studies

Ceder, A., & Wilson, N. H. (2007). "Tramway Systems: Design, Planning, and Operations." Transportation Research Part A: Policy and Practice

Çelik, H., & Yıldız, İ. (2018). Development of Rail Systems in Istanbul and Their Impacts on Traffic. Transportation Magazine

Url 1 www.batman.dhmi.gov.tr Access Date :07.01.2025

Url 2 www.uab.gov.tr Access Date :07.01.2025

Url 3 www.tuik.gov.tr Access Date :07.01.2025

Url 4 www.ibb.gov.tr Access Date :07.01.2025

Chapter 6

Drought Monitoring with Non-Parametric Kernel Density Estimation and The Lebesgue Method

Hasan TATLI[1]

1. INTRODUCTION

Drought is a significant environmental hazard with profound socio-economic and ecological impacts, making its assessment and monitoring a critical focus in climate and water sciences. Characterized by prolonged periods of insufficient precipitation relative to normal climatic conditions, droughts disrupt ecosystems, agriculture, and water resources management, necessitating reliable tools for their evaluation (McKee et al., 1993; Hayes et al., 2002).

To address the complexity of drought phenomena, various indices have been developed. Among these, the Standardized Precipitation Index (SPI) is widely used due to its simplicity and ability to quantify drought intensity over multiple timescales (McKee et al., 1993). The SPI focuses solely on precipitation data, making it adaptable for regions with limited hydrological records. More comprehensive indices, such as the Standardized Precipitation-Evapotranspiration Index (SPEI) and the Palmer Drought Severity Index (PDSI), incorporate evapotranspiration and soil moisture balance, providing a broader perspective on hydrological and agricultural drought (Vicente-Serrano et al., 2010; Rajsekhar et al., 2015; Baringo and Vicente-Serrano, 2018). The Reconnaissance Drought Index (RDI), on the other hand, integrates precipitation and potential evapotranspiration, emphasizing drought severity under changing climatic conditions (Pashiardis and Michaelides, 2008; Türkeş and Tatli, 2009; Öz *et al.,* 2024).

Traditional drought indices rely heavily on statistical assumptions, particularly the need to fit precipitation data to predefined parametric probability distributions, such as Gamma, Normal, or Pearson distributions (Zargar et al.,

[1] Prof. Dr, Department of Geography, Çanakkale Onsekiz Mart University, 17020, Çanakkale, Turkey
Orcid: https://orcid.org/0000-0002-1960-0618

2011; Beguería et al., 2014). These assumptions may not accurately represent precipitation variability, especially in regions with non-standard or skewed precipitation distributions, as often observed in arid and semi-arid climates (Tatli and Türkeş, 2011; Zarch et al., 2015; Tatli, 2021). Consequently, parametric methods can introduce biases in estimating drought severity, frequency, and duration, potentially compromising their utility in decision-making processes.

To overcome these limitations, non-parametric methods such as Kernel Density Estimation (KDE) have gained traction for their flexibility in modeling data distributions without assuming a specific parametric form (Chen, 2017). The Lebesgue method (Weir, 1973; Hartman and Mikusinski, 2014; Deliu and Liseo, 2023), which leverages KDE, provides a robust framework for estimating drought indices by directly capturing the empirical distribution of precipitation data. Unlike parametric approaches, the KDE-based method adapts to the actual shape of the data, making it particularly useful for analyzing extreme events or highly skewed distributions.

The primary objective of this study is to employ the Lebesgue-KDE method for calculating various drought indices, including SPI, SPEI, PDSI, and RDI. This innovative approach is applied to a case study spanning the Mediterranean region, from 1948 to 2024, covering the geographical bounds of 20°W to 45°E and 30°N to 50°N. In the initial phase, traditional SPI values are estimated using parametric methods. Subsequently, the same precipitation dataset is used to compute SPI values through the Lebesgue-KDE approach. The comparative analysis highlights the advantages of the KDE-based method, particularly in regions with complex precipitation patterns.

By offering a flexible, non-parametric alternative to traditional methods, this study aims to advance drought monitoring techniques. The findings have significant implications for improving the accuracy and reliability of drought assessments in diverse climatic regions, ultimately contributing to more effective climate adaptation and resource management strategies.

2. METHODOLOGY AND DATA

This section delves into the use of the Lebesgue method for the analysis of drought indices, with particular focus on the mathematical formulation of KDE, bandwidth selection, and the application to SPI, SPEI, PDSI, and RDI. We also discuss the computational steps involved, the advantages of the Lebesgue method over traditional approaches, and the necessary considerations for accurate results.

2.1 Data Sources

The first stage of our study involves gathering and processing monthly rainfall data for the Mediterranean region, utilizing the National Centers for Environmental Prediction and National Center for Atmospheric Research (NCEP/NCAR) reanalysis dataset (Kalnay et al., 1996) spanning from January 1948 to December 2024. This dataset, which covers longitudes from 20°W to 45°E and latitudes from 30°N to 50°N, is accessible through NOAA's Physical Sciences Laboratory. It is a widely recognized resource in climate research (Zhu et al., 2021a, b; Kerkar and Seelam, 2024), offering key global atmospheric variables including temperature, humidity, wind speed, and precipitation.

Since the original dataset provides precipitation rates, the dataset is preprocessed by converting these rates into daily values, multiplying by 86,400 (the number of seconds in a day). Monthly precipitation totals are further adjusted for leap years, guaranteeing accurate representation of precipitation patterns. This thorough preprocessing ensures our analysis captures the full spectrum of monthly precipitation variations throughout the study period.

2.2 The Lebesgue Kernel Density Estimation Method for Drought Indices

This study suggests the Lebesgue method (Weir, 1973; Hartman and Mikusinski, 2014; Deliu and Liseo, 2023) as a non-parametric approach for the estimation of drought indices, including SPI and other related indices. The core idea behind the Lebesgue method is the estimation of the CDF using KDE, followed by the transformation of this CDF into a standardized normal variable to obtain the drought index.

The KDE approach is a statistical technique used to estimate the probability density function (PDF) of a random variable based on observed data points (Sheather and Jones, 1991; Pulkkinen, 2016; Ji et *al.,* 2022). Given a set of data points $\boldsymbol{X} = \{\boldsymbol{x_1}, \boldsymbol{x_2}, \dots, \boldsymbol{x_n}\}$, KDE provides a smooth estimate of the underlying distribution. The KDE at any point $\boldsymbol{x}$ is given by:

$$\hat{\boldsymbol{f}}(\boldsymbol{x}) = \frac{\boldsymbol{1}}{\boldsymbol{nh}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{K}\left(\frac{\boldsymbol{x-x_i}}{\boldsymbol{h}}\right) \tag{1}$$

where:

- $\hat{f}(x)$ is the estimated density at point x,
- $\boldsymbol{n}$ is the number of data points,
- $\boldsymbol{h}$ is the bandwidth parameter (which controls the smoothness of the estimate),
- $\boldsymbol{K}(\cdot)$ is the kernel function, often chosen as a Gaussian kernel, i.e.,

$$\boldsymbol{K(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}} \tag{2}$$

The kernel function is typically symmetric and non-negative, with the most common choice being the Gaussian kernel. The bandwidth $\boldsymbol{h}$ plays a crucial role in determining the smoothness of the estimated density. A small bandwidth leads to a "wigglier" estimate, while a larger bandwidth results in a smoother estimate. The appropriate choice of bandwidth is critical for obtaining reliable results in drought index estimation.

Once the KDE is obtained, to approximate the CDF using a range-based approach inspired by Lebesgue integration, we partitioned the range of the PDF values $\boldsymbol{\hat{f}(x)}$ into discrete density levels. For each density level $\boldsymbol{\hat{f}(x_k)}$, the corresponding domain interval $[\boldsymbol{a_k}, \boldsymbol{b_k}]$ was identified, representing the range of precipitation values $\boldsymbol{x}$ where $\boldsymbol{f^{(x)}} \geq \boldsymbol{\hat{f}(x)} \geq \boldsymbol{\hat{f}(x_k)}$. The measure of this interval, denoted as $\boldsymbol{\mu_k}$, was computed as:

$$\boldsymbol{\mu_k = b_k - a_k} \tag{3}$$

where $\boldsymbol{b_k}$ and $\boldsymbol{a_k}$are the upper and lower bounds of the interval, respectively. The contribution of each density level $\boldsymbol{\hat{f}(x_k)}$ to the cumulative probability was then calculated as the product of the density value and the corresponding interval measure:

$$\boldsymbol{\Delta P_k = \hat{f}(x_k) \cdot \mu_k} \tag{4}$$

By summing the contributions $\boldsymbol{\Delta P_k}$ for all density levels $\boldsymbol{\hat{f}(x_k)}$ that are less than or equal to the density at a given point $\boldsymbol{x}$, the cumulative probability $\boldsymbol{F(x)}$ was obtained. This can be expressed mathematically as:

$$\boldsymbol{F(x) = \sum_{\hat{f}(x_k) \leq \hat{f}(x)} \hat{f}(x_k) \cdot \mu_k} \tag{5}$$

Boundary conditions were applied to ensure that $\boldsymbol{F(x) = 0}$ for precipitation values below the minimum observed value $(\boldsymbol{x} < \mathbf{min}(\boldsymbol{X})$) $\boldsymbol{and\ F(x) = 1}$ for values above the maximum observed value $(\boldsymbol{x} > \mathbf{max}(\boldsymbol{X}))$. The Drought Index (DI) then is calculated using the inverse of the normal distribution function (the quantile function):

$$\boldsymbol{DI(x) = \Phi^{-1}\left(\hat{F}(x)\right)} \tag{6}$$

where $\boldsymbol{\Phi^{-1}}(\cdot)$ denotes the inverse CDF of the standard normal distribution. The result is a standardized drought index that can be interpreted in the same way as traditional SPI (McKee et al., 1993; Baringo and Vicente-Serrano, 2018) values: values below zero indicate drought conditions, while values above zero indicate wet conditions.

The bandwidth parameter $\boldsymbol{h}$ is one of the most critical components in the KDE (Chen, 2017; Wand and Jones, 1995). It controls the degree of smoothing and directly influences the accuracy of the density estimate. If the bandwidth is too small, the estimate will be overly sensitive to fluctuations in the data, leading to a "noisy" or highly variable density estimate. Conversely, a large bandwidth will over-smooth the data, potentially ignoring important features of the underlying distribution.

In the context of precipitation data for drought index analysis, the choice of bandwidth is especially important due to the potential skewness and heavy tails of precipitation distributions. Therefore, the selection of an optimal bandwidth is essential to ensure that the estimated density reflects the true characteristics of the precipitation data.

There are several methods for selecting the optimal bandwidth $\boldsymbol{h}$ in KDE (Wand and Jones, 1995). These include:

Rule-of-Thumb Method: This is a common, simple approach where the bandwidth is set based on the standard deviation $\boldsymbol{\sigma}$ of the data (Silverman, 1986; Groß and Rendtel, 2016; Harpole et al., 2014). For a Gaussian kernel, the rule-of-thumb bandwidth is given by:

$$\boldsymbol{h = 1.06\sigma n^{-1/5}} \tag{7}$$

where $\boldsymbol{\sigma}$ is the standard deviation of the data, and $\boldsymbol{n}$ is the number of data points. While this method is easy to implement, it may not perform well for skewed or multimodal distributions.

Cross-Validation (CV): Cross-validation involves minimizing the integrated squared error between the true density and the estimated density (Wahbah et al., 2022). This method is more computationally intensive but often provides better results in practice. The bandwidth is chosen as:

$$\boldsymbol{h_{CV} = \arg\min_{h} \int \left(f(x) - \widehat{f_h}(x)\right)^2 dx} \tag{8}$$

where $\boldsymbol{\widehat{f_h}(x)}$ is the KDE with bandwidth $\boldsymbol{h}$, and $\boldsymbol{f(x)}$ is the true density (unknown). The bandwidth that minimizes this error is considered optimal.

Silverman's Method: A well-known approach for selecting bandwidth in KDE is Silverman's method, which is based on the normal reference rule (Silverman, 1986). The bandwidth is given by:

$$\boldsymbol{h}_{\text{Silverman}} = \left(\frac{\mathbf{4}}{\mathbf{3}\boldsymbol{n}}\right)^{\mathbf{1/5}} \tag{9}$$

where $\boldsymbol{\sigma}$ is the standard deviation of the data. This method assumes that the data is unimodal and normally distributed, which may not be the case for precipitation data. However, it provides a useful starting point.

Skewness-Adjusted Bandwidth: Given that precipitation data are often skewed, we propose an adjusted bandwidth formula that incorporates the skewness of the data. The bandwidth $\boldsymbol{h}_{\text{skew}}$ can be defined as:

$$\boldsymbol{h}_{\text{skew}} = \boldsymbol{\sigma}(\mathbf{1} + \boldsymbol{\gamma})^{\mathbf{1/3}}\boldsymbol{n}^{-\mathbf{1/5}} \tag{10}$$

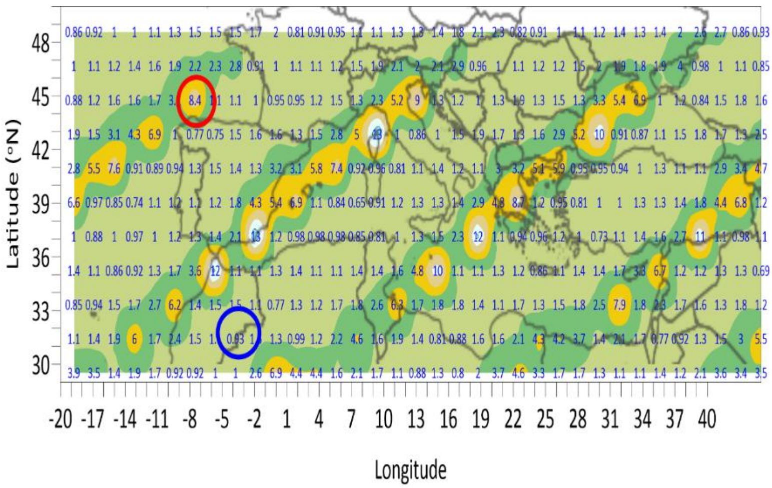
where $\boldsymbol{\gamma}$ is the skewness of the data, and $\boldsymbol{\sigma}$ is the standard deviation. This method adjusts the bandwidth based on the degree of skewness in the data, allowing for more appropriate smoothing in the presence of skewed distributions. This approach has been shown to improve the quality of density estimates for data with significant skewness, as is typical for precipitation data.

3. RESULTS

3.1 Skewness Distribution Across the Mediterranean Basin

Figure 1 provides a detailed spatial representation of skewness coefficients calculated for monthly precipitation data over the Mediterranean basin. High positive skewness values, which dominate arid and semi-arid regions, indicate precipitation regimes characterized by infrequent but intense rainfall events. These findings align with the established climatic understanding that arid regions experience sporadic precipitation episodes interspersed with prolonged dry periods. The spatial patterns observed emphasize the interplay between regional climatic dynamics and topographic influences.

Figure 1. Distribution of skewness coefficients for monthly precipitation across the Mediterranean Basin, derived from NCEP/NCAR Reanalysis data spanning 1948 to 2024.



Iberian Peninsula: The western coast of the Iberian Peninsula also exhibits elevated skewness, influenced by the Azores High, which often suppresses precipitation, limiting it to episodic Atlantic storms and occasional atmospheric rivers. These intense but rare rainfall events result in high positive skewness coefficients, typical of semi-arid to arid climates (Luterbacher et al. 2006; Hurrell, 1995; Hurrell and Deser, 2010; Bell et al., 2022).

Balkans and Southeastern Europe: A prominent cluster of high skewness extends across the Balkans, including Croatia, Serbia, and Bosnia-Herzegovina, and into Eastern Europe. This pattern is likely driven by a combination of convective storms during summer, Mediterranean cyclones (including cut-off lows) in the cool season, and orographic effects due to the Dinaric Alps (Trigo et al. 2002; Papadopoulos and Varlas, 2020; Milošević et al., 2021).

Eastern Mediterranean and Southeastern Turkey: High positive skewness is particularly pronounced in the Eastern Mediterranean, encompassing Cyprus, Israel, Lebanon, and Syria, extending into southeastern Turkey. This region is strongly influenced by the interaction of warm Mediterranean waters and arid continental air masses, leading to infrequent but high-intensity rainfall events. The Taurus Mountains further enhance precipitation extremes (Cook et al., 2016; Zittis et al., 2022).

North Africa to Central Europe (Gibraltar Corridor): A wave-like pattern of skewness coefficients extends from North Africa (including Morocco and Algeria), passing through the Strait of Gibraltar and reaching regions like France and Germany. This corridor represents the pathway of Mediterranean cyclones and Saharan dust-laden storms, which contribute to sporadic but intense precipitation events (Schilling et al., 2020; Deidda et al. 2021). While

Mediterranean cyclones do pass through this region, the primary driver of the skewness pattern in North Africa is the Saharan climate. The influence of Mediterranean cyclones becomes more pronounced as you move north towards Europe. Rephrasing to emphasize the Saharan influence in North Africa and the increasing role of Mediterranean cyclones further north would be more accurate

Central Mediterranean to Eastern Europe: Starting in the central Mediterranean, over regions like Sicily and southern Italy, the skewness pattern intensifies over the Balkans before extending into Eastern Europe (including Romania and Ukraine). This pattern is shaped by the convergence of moist Mediterranean air and continental air masses, leading to high variability in precipitation intensity (Zolina et al., 2010; Capozzi et al., 2024).

Northwestern Middle East (Mesopotamian Basin): While high skewness is apparent in the northwestern Middle East, including parts of Iraq and Iran, data limitations, primarily due to sparse station coverage in this region, make it challenging to precisely define the pattern's extent. This area's precipitation is driven by synoptic systems moving from the Mediterranean and occasional moisture incursions from the Persian Gulf (Zittis et al., 2022; Cook et al., 2016).

The red and blue circles in Figure 1-mark specific grid points selected as test precipitation data values for our proposed method. The red circles represent locations with the highest observed skewness, indicative of highly variable precipitation regimes, while the blue circles mark locations with the lowest skewness, representing more consistent rainfall patterns. These points were chosen to represent the extremes of the skewness distribution and to test the method's ability to handle diverse precipitation characteristics.

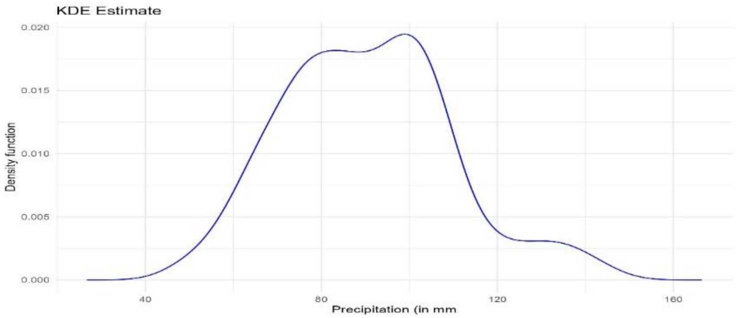
These patterns demonstrate a strong correlation with regional geography, including mountain ranges (e.g., the Alps, Taurus Mountains, Dinaric Alps) and large bodies of water (e.g., the Mediterranean Sea, Atlantic Ocean). Seasonal variability in climatic drivers—such as the North Atlantic Oscillation (NAO), Mediterranean Oscillation (MO), and interactions between subtropical and mid-latitude systems—likely further modulates these skewness patterns. The observed skewness distribution underscores the Mediterranean basin's vulnerability to extreme weather events, including flash floods and droughts, highlighting the importance of accurate precipitation modeling for risk assessment and water resource management.

3.2 Kernel Density Estimation and SPI Comparisons

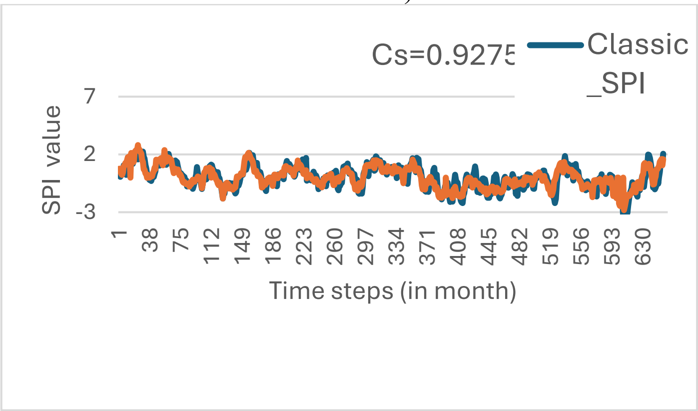
Figures 2 and 3 showcase the efficacy of the KDE in accurately capturing precipitation variability across grid points with contrasting skewness values. By utilizing the Lebesgue method, this study demonstrates improved representation of both low and high-skewness precipitation distributions in SPI calculations.

Figure 2a illustrates the KDE results for a grid point with a low skewness coefficient (Cs = 0.927514). For this grid point, the precipitation distribution is nearly symmetric, and the KDE output aligns closely with the Gamma CDF employed in classical SPI calculations. The similarity between these distributions validates the applicability of parametric methods in regions with near-normal precipitation behavior.

Figure 2. (a) Kernel Density Estimation (KDE) results for the analyzed precipitation data, illustrating the probability density function for the selected grid point. (b) Standardized Precipitation Index (SPI) simulations at the same grid point.



a)



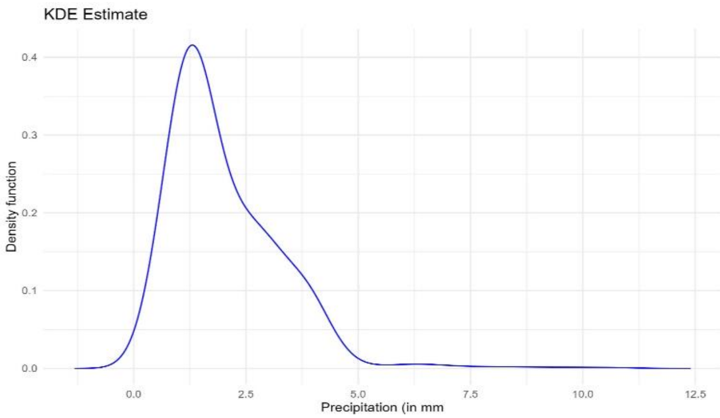
b)

The KDE-based Lebesgue SPI further refines the results by directly modeling empirical data without assuming a predefined distribution, capturing subtle deviations from symmetry. Figure 2b compares SPI values derived using the classical Gamma-based approach and the non-parametric Lebesgue method. The results demonstrate minor discrepancies, particularly during transitions from drought to wet conditions. While the classical SPI slightly smooths over these transitions, the Lebesgue SPI captures finer temporal variations, reflecting its enhanced sensitivity even in low-skewness contexts.

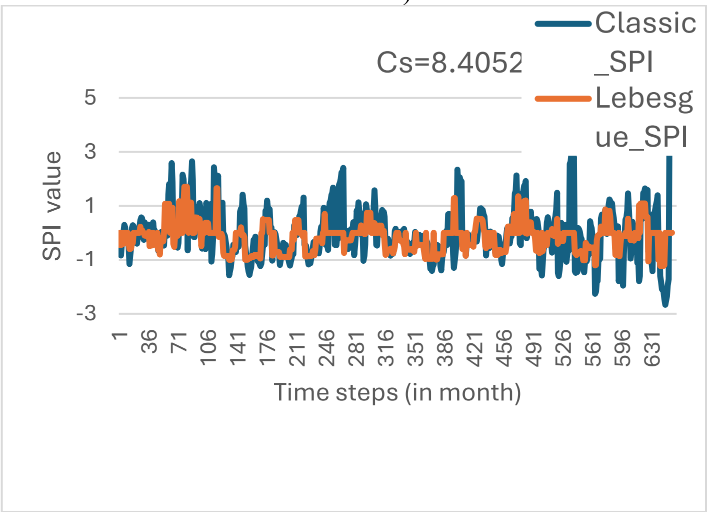
Figure 3a presents the KDE results for a grid point with a high skewness coefficient (Cs = 8.405206). This grid point is characterized by a heavily skewed precipitation distribution, indicative of an arid climate where extreme rainfall events dominate. The Gamma CDF used in classical SPI calculations fails to adequately represent the heavy tail of the precipitation data, leading to a mischaracterization of extreme events.

In contrast, the KDE accurately models the asymmetric and heavy-tailed distribution, which is crucial for capturing the variability inherent in such data. Figure 3b compares SPI values obtained using classical and Lebesgue methods. The classical SPI underestimates drought severity during dry periods and overestimates wetness during extreme rainfall events due to the limitations of its parametric assumptions. The Lebesgue SPI, leveraging the flexibility of KDE, provides a more accurate and responsive representation of precipitation extremes. This advantage is particularly evident during periods of anomalous precipitation, where the Lebesgue method aligns closely with observed data and effectively highlights the impact of extreme events.

Figure 3. Kernel Density Estimation (KDE) results for the analyzed precipitation data, showing the probability density function for the selected grid point, (a). Standardized Precipitation Index (SPI) simulations at the same grid point (b).



a)



b)

According to the seasonal and spatial analyses of SPI-6 values underscore the limitations of classical methods in highly skewed climates and highlight the advantages of the Lebesgue method.

4. CONCLUSIONS

This study provides a comprehensive evaluation of drought index estimation methodologies over the Mediterranean basin, emphasizing the critical role of non-parametric approaches in addressing the limitations of traditional parametric methods. The results underscore the region's complex climatic and meteorological dynamics, particularly the impact of high skewness in precipitation distributions on drought assessment.

The skewness distribution analysis reveals pronounced positive skewness in arid and semi-arid regions, such as North Africa, the Levant, and parts of southeastern Europe. These findings highlight the episodic and intense nature of precipitation in these areas, which is driven by interactions between regional climatic drivers such as the Azores High, Mediterranean cyclones, and subtropical influences. The strong correlation between high skewness coefficients and arid climates underscores the Mediterranean basin's vulnerability to extreme weather events, including flash floods and droughts. This spatial variability necessitates advanced modeling techniques that accurately reflect the unique precipitation patterns of each subregion.

By incorporating KDE into the calculation of SPI, the Lebesgue method offers a significant advancement in the representation of precipitation variability. In low-skewness contexts, such as those observed in certain Mediterranean coastal areas, the classical Gamma-based SPI performs adequately, as demonstrated by the alignment of its results with KDE outputs. However, the limitations of classical methods become evident in high-skewness scenarios, where the Gamma CDF fails to account for the heavy tails of precipitation distributions. The Lebesgue SPI effectively addresses these challenges, capturing the full spectrum of precipitation variability and providing a more nuanced depiction of drought severity and duration.

The seasonal and spatial analyses further illustrate Lebesgue SPI's superior performance. Seasonal transitions, particularly during winter cyclonic activity and summer convective storms, are better represented by the KDE-based approach. This method's ability to integrate episodic heavy rainfall into drought assessments ensures greater precision in identifying periods of drought and wetness. Spatially, the Lebesgue SPI outperforms classical methods in highlighting drought-prone regions, particularly those with irregular precipitation patterns, such as the arid zones of North Africa and the Levant.

The success of the Lebesgue method is largely attributed to the use of a skewness-adjusted bandwidth in KDE. This innovation ensures that the density estimates reflect the true variability of the precipitation data, avoiding the pitfalls of over-smoothing or noise amplification. By accurately modeling heavy-tailed distributions, the Lebesgue SPI provides a robust framework for monitoring extreme precipitation events, which are increasingly critical in the context of climate change. Key findings from this study include:

Enhanced Accuracy in Arid Climates: The Lebesgue SPI demonstrates its capacity to capture the asymmetry in precipitation distributions, offering a more reliable assessment of drought and wetness extremes.

Improved Seasonal Sensitivity: Seasonal variations in precipitation, particularly during transitional periods, are better captured by the KDE-based approach, enabling a finer resolution of drought dynamics.

Adaptability Across Regions: The flexibility of the Lebesgue method ensures its applicability across diverse climatic zones, from semi-arid to temperate regions, making it a versatile tool for global drought monitoring.

The implications of these findings extend beyond the Mediterranean basin, offering a template for improving drought assessment in regions with similar climatic complexities. As climate change continues to exacerbate precipitation variability and extremes, the adoption of advanced non-parametric methods such as the Lebesgue SPI will be essential for developing adaptive management strategies. Future research should focus on integrating additional climatic variables, such as temperature and soil moisture, to further enhance the robustness of drought indices. Additionally, expanding the application of the Lebesgue method to multi-scalar drought indices, including SPEI and PDSI, will provide a more holistic understanding of hydrological and agricultural droughts.

5. REFERENCES

1. Baringo, L., & Vicente-Serrano, S. M. (2018). A review of the Standardized Precipitation Index (SPI) and its applications. *Atmospheric Research*, 204, 1-13.
2. Beguería, S., Vicente-Serrano, S. M., Reig, F., & Latorre, B. (2014). Standardized precipitation evapotranspiration index (SPEI) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *International Journal of Climatology*, *34*(10), 3001-3023.
3. Bell, B. A., Hughes, P. D., Fletcher, W. J., Cornelissen, H. L., Rhoujjati, A., Hanich, L., & Braithwaite, R. J. (2022). Climate of the Marrakech High Atlas, Morocco: Temperature lapse rates and precipitation gradient from piedmont to summits. *Arctic, Antarctic, and Alpine Research*, *54*(1), 78-95.
4. Capozzi, V., Serrapica, F., Rocco, A., Annella, C., & Budillon, G. (2024). Historical snowfall measurements in the Central and Southern Apennine Mountains: climatology, variability and trend. *EGUsphere*, *2024*, 1-43.
5. Chen, Y. C. (2017). A tutorial on kernel density estimation and recent advances. *Biostatistics & Epidemiology*, *1*(1), 161-187.
6. Cook, B. I., Anchukaitis, K. J., Touchan, R., Meko, D. M., & Cook, E. R. (2016). Spatiotemporal drought variability in the Mediterranean over the last 900 years. *Journal of Geophysical Research: Atmospheres*, *121*(5), 2060-2074.
7. Deidda, C., Rahimi, L., & De Michele, C. (2021). Causes of dependence between extreme floods. *Environmental Research Letters*, *16*(8), 084002.
8. Deliu, N., & Liseo, B. (2023). Alternative approaches for computing highest-density regions. *arXiv preprint arXiv:2401.00245*.
9. Groß, M., & Rendtel, U. (2016). Kernel density estimation for heaped data. *Journal of Survey Statistics and Methodology*, *4*(3), 339-361.
10. Harpole, J. K., Woods, C. M., Rodebaugh, T. L., Levinson, C. A., & Lenze, E. J. (2014). How bandwidth selection algorithms impact exploratory data analysis using kernel density estimation. *Psychological Methods*, *19*(3), 428.
11. Hartman, S., & Mikusinski, J. (2014). *The Theory of Lebesgue Measure and Integration*. Elsevier.
12. Hayes, M. J., Alvord, C., & Lowrey, J. (2002). *Drought Indices*. National Drought Mitigation Center, University of Nebraska.
13. Hurrell, J. W. (1995). Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. Science, 269 (5224), 676–679. https://doi.org/10.1126/science.269.5224.676.

14. Hurrell, J. W., & Deser, C. (2010). North Atlantic climate variability: The role of the North Atlantic Oscillation. Journal of Marine Systems, 79, 231–244.

15. Ji, Y., Li, Y., Yao, N., Biswas, A., Chen, X., Li, L., ... & Liu, F. (2022). Multivariate global agricultural drought frequency analysis using kernel density estimation. *Ecological Engineering*, *177*, 106550.

16. Kalnay, E., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77, 437-470.

17. Kerkar, J. P., & Seelam, J. K. (2024). Trends in the surface wind-wave heights in the North Indian Ocean based on 50-year NCEP/NCAR reanalysis wind-wave data. *Regional Studies in Marine Science*, 72, 103430.

18. Luterbacher, J., Xoplaki, E., Casty, C., Wanner, H., Pauling, A., Küttel, M., ... & Ladurie, E. L. R. (2006). Mediterranean climate variability over the last centuries: *A review. Developments in Earth and Environmental Sciences*, 4, 27-148.

19. McKee, T. B., Doesken, N. J., & Kleist, J. (1993, January). The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th Conference on Applied Climatology* (Vol. 17, No. 22, pp. 179-183).

20. Milošević, D., Stojsavljević, R., Szabó, S., Stankov, U., Savić, S., & Mitrović, L. (2021). Spatio-temporal variability of precipitation over the Western Balkan countries and its links with the atmospheric circulation patterns. *Journal of the Geographical Institute" Jovan Cvijic", SASA*, *71*(1), 29-42.

21. Öz, F. Y., Özelkan, E., & Tatli, H. (2024). Comparative analysis of SPI, SPEI, and RDI indices for assessing spatio-temporal variation of drought in Türkiye. *Earth Science Informatics*, 1-33.

22. Papadopoulos, A., & Varlas, G. (2020). Weather systems affecting the meteorological conditions over the Aegean Sea. In *The Aegean Sea Environment: The Geodiversity of the Natural System* (pp. 141-165). Cham: Springer International Publishing.

23. Pashiardis, S., & Michaelides, S. (2008). Implementation of the standardized precipitation index (SPI) and the reconnaissance drought index (RDI) for regional drought assessment: a case study for Cyprus. *European Water*, *23*(24), 57-65.

24. Pulkkinen, S. (2016). Nonlinear kernel density principal component analysis with application to climate data. *Statistics and Computing*, *26*, 471-492.

25. Rajsekhar, D., Singh, V. P., & Mishra, A. K. (2015). Multivariate drought index: An information theory-based approach for integrated drought assessment. *Journal of Hydrology*, *526*, 164-182.
26. Schilling, J., Hertig, E., Tramblay, Y., & Scheffran, J. (2020). Climate change vulnerability, water resources and social implications in North Africa. *Regional Environmental Change*, *20*, 1-12.
27. Sheather, S. J., & Jones, M. C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society: Series B (Methodological)*, *53*(3), 683-690.
28. Silverman, B. W. (1986). Density estimation for statistics and data analysis. CRC press.
29. Tatli, H. (2021). Multivariate-drought indices—case studies with observations and outputs of NCAR CCSM-4 ensemble models. *Theoretical and Applied Climatology*, *146*(1), 257-275.
30. Tatli, H., & Türkeş, M. (2011). Empirical orthogonal function analysis of the Palmer drought indices. *Agricultural and Forest Meteorology*, *151*(7), 981-991.
31. Trigo, I. F., Bigg, G. R., & Davies, T. D. (2002). Climatology of cyclogenesis mechanisms in the Mediterranean. *Monthly Weather Review*, *130*(3), 549-569.
32. Türkeş, M., & Tatli, H. (2009). Use of the standardized precipitation index (SPI) and a modified SPI for shaping the drought probabilities over Turkey. *International Journal of Climatology*, *29*(15), 2270-2282.
33. Vicente-Serrano, S. M., et al. (2010). A multi-scalar drought index for evaluating drought impacts: SPI, SPEI, and PDSI. *Geophysical Research Letters*, 37(7), L07701.
34. Wahbah, M., Mohandes, B., EL-Fouly, T. H., & El Moursi, M. S. (2022). Unbiased cross-validation kernel density estimation for wind and PV probabilistic modelling. *Energy Conversion and Management*, *266*, 115811.
35. Wand, M. P., & Jones, M. C. (1995). *Kernel Smoothing*. CRC Press.
36. Weir, A. J. (1973). *Lebesgue Integration and Measure* (Vol. 1). Cambridge University Press.
37. Zarch, M. A. A., Sivakumar, B., & Sharma, A. (2015). Droughts in a warming climate: A global assessment of Standardized precipitation index (SPI) and Reconnaissance drought index (RDI). *Journal of Hydrology*, *526*, 183-195.
38. Zargar, A., Sadiq, R., Naser, B., & Khan, F. I. (2011). A review of drought indices. *Environmental Reviews*, *19*(NA), 333-349.

39. Zhu, K., Guan, X., Huang, J., Wang, J., Guo, S., & Cao, C. (2021a). Precipitation over semi-arid regions of North Hemisphere affected by Atlantic Multidecadal Oscillation. *Atmospheric Research*, 262, 105801 https://doi.org/10.1016/j.atmosres.2021.105801
40. Zhu, X., Dong, T., Zhao, S., & He, W. (2021b). A comparison of global surface air temperature over the oceans between CMIP5 models and NCEP reanalysis. *Frontiers in Environmental Science*, 9, 656779. https://doi.org/10.3389/fenvs.2021.656779
41. Zittis, G., Almazroui, M., Alpert, P., Ciais, P., Cramer, W., Dahdal, Y., ... & Lelieveld, J. (2022). Climate change and weather extremes in the Eastern Mediterranean and Middle East. *Reviews of geophysics*, *60*(3), e2021RG000762.
42. Zolina, O., Simmer, C., Gulev, S. K., & Kollet, S. (2010). Changing structure of European precipitation: Longer wet periods leading to more abundant rainfalls. *Geophysical Research Letters*, *37*(6).

Chapter 7

Revisiting The Canadian Forest Fire Weather Index (FWI) Using A System of Ordinary Differential Equations

Hasan TATLI[1]

1. INTRODUCTION

The increasing frequency and intensity of wildfires worldwide, driven by climate change and expanding human activities in fire-prone areas, have become a pressing environmental concern (Pausas and Keeley, 2021; Shi and Touge, 2022; Jones et al., 2024). Accurate fire danger rating systems are essential for predicting fire risks and informing prevention strategies. One of the most widely adopted systems globally is the Canadian Forest Fire Weather Index (FWI) system (Lawson and Armitage, 2008; Wang et al., 2023), which has proven effective across diverse environmental conditions in assessing the potential for fire ignition and spread.

The FWI system, developed by Van Wagner (1987), is an empirical framework that utilizes daily weather inputs such as temperature, relative humidity, wind speed, and precipitation to estimate moisture levels in various layers of forest fuels. The system computes indices including the Fine Fuel Moisture Code (FFMC), Duff Moisture Code (DMC), and Drought Code (DC), which correspond to surface fuels, decomposing organic material, and deeper organic layers, respectively. Derived indices like the Initial Spread Index (ISI), Build-Up Index (BUI), and the final Fire Weather Index (FWI) collectively provide a comprehensive assessment of fire danger based on short- and long-term meteorological and fuel conditions.

Despite its robustness, the standard methodology for calculating FWI indices is inherently discrete, relying on daily updates that do not fully capture the continuous and dynamic nature of environmental processes governing fire weather (Lawson and Armitage, 2008). Weather variables such as temperature,

[1] Prof. Dr, Department of Geography, Çanakkale Onsekiz Mart University, 17020, Çanakkale, Turkey
Orcid: https://orcid.org/0000-0002-1960-0618

wind, and humidity fluctuate continuously throughout the day, and discrete models may overlook subtle yet critical variations in fire risk. Additionally, phenomena like the rapid drying of fine fuels during peak solar hours or sudden increases in wind speed may require more frequent updates than the traditional daily FWI calculations allow.

To address these limitations, this study proposes an alternative approach that reformulates the FWI system as a set of Ordinary Differential Equations (ODEs). ODEs describe the rate of change of variables over time, making them ideal for modeling continuously evolving processes. By representing the evolution of FWI indices as ODEs, this approach allows for a more dynamic and responsive fire danger model that continuously updates indices based on real-time meteorological data.

While the ODE-based model successfully captures fine-scale variations in fire weather conditions, certain long-term indices, such as the Drought Code (DC), exhibit minor deviations from traditional FWI calculations due to the influence of cumulative precipitation. Despite these challenges, the proposed model aligns closely with the classical FWI framework and offers enhanced flexibility and precision in fire danger assessment.

This paper details the mathematical formulation of the ODE-based FWI system, outlines the numerical methods used for solving the equations, and compares the results with those from the conventional FWI method. By demonstrating the strengths and limitations of the continuous-time approach, this study provides valuable insights into the potential of dynamical systems in improving fire danger predictions.

2. METHODOLOGY

This section describes the reformulation of the FWI system as a set of ODEs, providing a continuous-time framework for modeling fire danger. Numerical integration techniques are employed to solve these equations, enabling dynamic updates of fire danger indices over time.

2.1 Revisiting the Fine Fuel Moisture Code (FFMC) Formulation

The Fine Fuel Moisture Code (FFMC) reflects the moisture content in fine surface fuels, which directly affects fire ignition and spread. The temporal evolution of FFMC is modeled as:

$$\frac{dM_f}{dt} = -D\left(M_f - M_e\right) \qquad (1)$$

where M_f represents the fine fuel moisture content, M_e is the equilibrium moisture content influenced by temperature T, relative humidity H, and wind speed W, and D denotes the drying rate. Precipitation P modifies M_f as follows:

- For ($P > 0$):

$$M_f = M_f + C \cdot P \tag{2}$$

- Otherwise:

$$\frac{dM_f}{dt} = f(T, H, W) \tag{3}$$

where C is a precipitation-dependent coefficient, and $f(T, H, W)$ represents the drying function.

2.2 Revisiting the Duff Moisture Code (DMC Formulation

The DMC quantifies moisture in decomposing organic material. Its rate of change is given by:

$$\frac{dD}{dt} = -k(T, H) \cdot D + g(P) \tag{4}$$

Here, $k(T, H)$ represents the temperature and humidity-dependent drying rate, and $g(P)$ is a precipitation adjustment function.

2.3 Revisiting the Drought Code (DC) Formulation

The DC measures moisture in deep organic layers, indicative of long-term drying. Its evolution follows:

$$\frac{dC}{dt} = -r(T) \cdot C + h(P) \tag{5}$$

where $r(T)$ is a temperature-driven drying rate, and $h(P)$ adjusts for precipitation effects.

2.4 Revisiting the Fire Spread and the Initial Spread Index (ISI) Formulation

The ISI estimates fire spread potential, derived as:

$$\frac{dI}{dt} = W \cdot g(M_f) \tag{6}$$

where (W) is wind speed, and $(g(M_f))$ is a moisture response function based on FFMC.

2.5 Reformulation of the Cumulative Fuel Potential in the Build-Up Index (BUI)

The BUI integrates DMC and DC to estimate available fuel for combustion. It evolves as:

$$\frac{dB}{dt} = h(D, C) - \gamma \cdot B \tag{7}$$

where $h(D, C)$ represents the fuel accumulation rate, and γ is a damping coefficient.

2.6 Reformulation of the Integrating Fire Risk: The Fire Weather Index (FWI)

The FWI integrates ISI and BUI to assess fire intensity:

$$\frac{dF}{dt} = \alpha \cdot I \cdot B - \beta \cdot F \tag{8}$$

where α is a growth coefficient, and β is a limiting factor that accounts for fire suppression or environmental constraints.

Numerical solutions for these equations were implemented using the Euler method, applying daily meteorological data such as temperature T, relative humidity H, wind speed W, and precipitation P from the Canadian Forest Fire Weather system dataset.

2.7 Existence, Uniqueness, and Stability of ODE Solutions

The mathematical soundness of the proposed ODE-based model relies on ensuring the existence and uniqueness of solutions for the system of differential equations representing the evolution of FWI indices. These properties guarantee that for any set of well-defined initial conditions and continuous meteorological inputs, the model produces a consistent and valid prediction over a specified time interval. The Picard-Lindelöf theorem (also known as the Cauchy-Lipschitz theorem) serves as the theoretical foundation for analyzing these properties. Consider a system of ordinary differential equations of the form:

$$\frac{dy}{dt} = f(t, y), \quad y(t_0) = y_0 \tag{9}$$

where $t \in [t_0, t_1]$ represents time, $y(t)$ in R^n is the state vector of the system, and $f: [t_0, t_1] \times R^n \to R^n$ is a vector field. The existence and uniqueness of solutions depend on the following conditions:

1. Continuity: $f(t, y)$ must be continuous in both t and y.
2. Lipschitz Condition: There exists a constant $L > 0$ such that for all $t \in [t_0, t_1]$ and $y_1, y_2 \in R^n$:

$$|f(t, y_1) - f(t, y_2)| \leq L|y_1 - y_2| \tag{10}$$

where $|\cdot|$ denotes the Euclidean norm in R^n.

If these conditions are satisfied, the Picard-Lindelöf theorem guarantees the existence of a unique solution $y(t)$ on some interval $[t_0, t_0 + \delta]$ $where\ \delta > 0$.

2.8 Application to the FWI System

The ODE-based formulation of the FWI indices—FFMC, DMC, DC, ISI, BUI, and FWI—can be expressed as:

$$\frac{dM}{dt} = f_M(t, M), \quad M(t_0) = M_0 \tag{11}$$

where $M(t) \in R^m$ represents the vector of fire indices, and $f_M: R \times R^m \to R^m$ defines the governing equations. Specifically, each index evolves as:

$$\frac{dM_i}{dt} = g_i(t, M_i) + h_i(t, P), \quad i = 1, \dots, m \tag{12}$$

where $g_i(t, M_i)$ represents the drying term dependent on temperature *T*, humidity *H*, and wind speed *W*, and $h_i(t, P)$ accounts for precipitation adjustments.

Continuity: The functions g_i and h_i involve smooth operations such as exponentials, logarithms, and linear transformations of *T, H, W,* and *P*. These meteorological inputs, provided by numerical weather prediction models, are assumed to be continuous in t. Thus, $f_M(t, M)$ is continuous in both t and M.

Lipschitz Condition: Each component of f_M satisfies a bounded derivative condition due to the physical constraints on meteorological inputs. For example, $(T), (H), and\ (W)$ are bounded by realistic atmospheric ranges, and precipitation P is non-negative and finite. These bounds ensure that:

$$|f_M(t, M_1) - f_M(t, M_2)| \leq L|M_1 - M_2| \tag{13}$$

where L is a finite constant determined by the physical parameters of the system.

2.9 Physical Interpretation

Existence: The existence of solutions reflects the continuous nature of the processes underlying fire indices. For instance, fine fuel drying, moisture absorption, and energy exchange processes are inherently smooth and governed by atmospheric continuity. Thus, small changes in $T, H, W, or\ P$ lead to smooth adjustments in M.

Uniqueness: The uniqueness of solutions ensures deterministic behavior of the FWI system. Given identical initial conditions and inputs, the system evolves along a single trajectory, avoiding ambiguities in fire danger predictions.

While the Picard-Lindelöf theorem ensures local existence and uniqueness, global solutions over extended time intervals are guaranteed if $f_M(t, M)$ is globally Lipschitz or if the solutions remain bounded. In practice, the bounded nature of M—constrained by physical limits of moisture content and fire spread indices—ensures that the ODE system remains well-posed for operational fire forecasting.

3. RESULTS

This section presents the numerical results obtained from implementing the proposed ODE-based framework. A rigorous analysis of the computed indices and their behavior is performed using mathematical reasoning and physical principles in meteorology. The ODE-based model was implemented using the explicit Euler method, solving the system of equations with daily meteorological inputs temperature T, relative humidity H, wind speed W, and precipitation P. The temporal evolution of the indices was tracked for each time step, and their dynamics were analyzed against traditional FWI outputs. The system of ODEs for FFMC, DMC, DC, ISI, BUI, and FWI is solved numerically with the same daily meteorological data inputs as in used in Van Wagner (1987) that one can freely download our developed software from (https://github.com/tatliHasan/FWI_ODEs). The functions in our software are designed to perform iterative updates on the fire weather indices based on meteorological data.

Figure 1: Comparison of FWI outputs: Traditional discrete method (blue) vs. proposed ODE-based method (orange). The ODE framework captures subtle intra-day variations in fire danger influenced by dynamic meteorological conditions

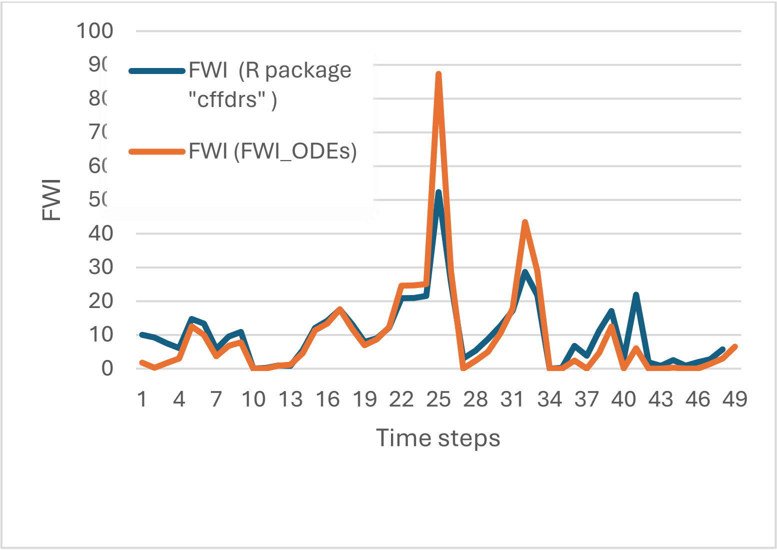
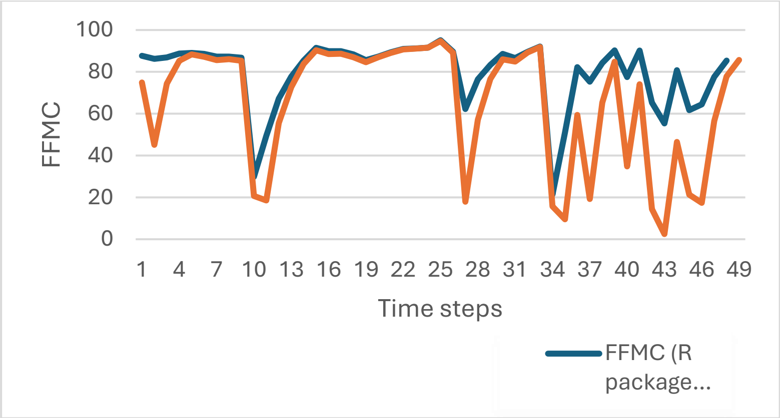


Figure 1 illustrates the FWI computed using the traditional discrete Van Wagner method and the proposed ODE-based approach. Both methods capture the general trends in fire danger assessment, validating the accuracy of the ODE-based framework. However, the ODE model demonstrates its ability to capture finer-scale temporal variations caused by meteorological fluctuations. For example, rapid changes in relative humidity or sudden increases in wind speed, which may occur within hours, are reflected in the continuous evolution of the ODE model, whereas the traditional method, with its daily update structure, smooths out these variations. This enhanced responsiveness is particularly evident during transitional weather periods, such as the onset of dry spells or sharp increases in temperature. By offering more precise and dynamic assessments of fire danger, the ODE approach provides critical insights for operational fire management, enabling better preparedness during high-risk periods.

Figure 2 highlights the temporal evolution of the FFMC, comparing results from the traditional discrete method and the ODE-based framework. The FFMC, representing the moisture content in fine surface fuels, is highly sensitive to rapid environmental changes, and this sensitivity is well-captured by the ODE-based approach. For instance, during midday hours, when solar radiation peaks and temperatures rise sharply, the ODE model shows significant reductions in FFMC due to accelerated drying.

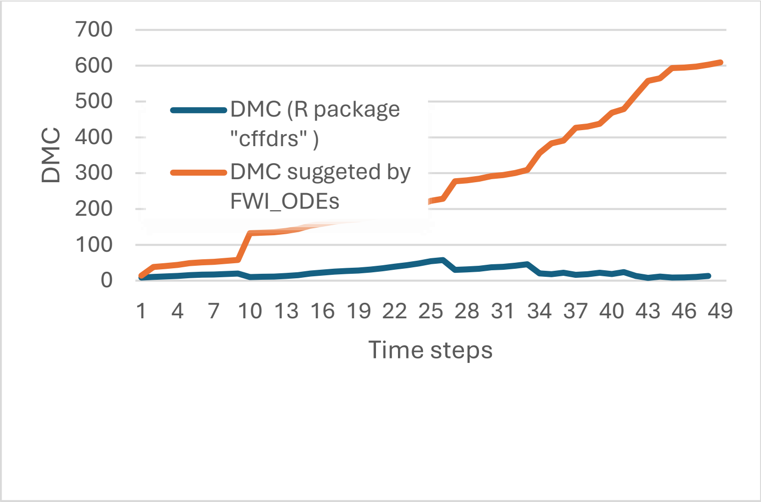
Figure 2. Temporal evolution of FFMC: Comparing the traditional discrete approach (blue) with the ODE-based method (orange). The proposed method highlights rapid responsiveness to meteorological changes



Conversely, periods of rainfall lead to immediate increases in FFMC, accurately modeled as the system incorporates real-time precipitation data. The underlying dynamics of FFMC are governed by the equilibrium moisture content (M_e), which itself depends on temperature, humidity, and wind speed. The ODE model's ability to continuously track changes in M_e highlights periods of heightened fire ignition potential, such as those following rapid drying of fine fuels during sunny conditions. These results emphasize the importance of high-frequency updates in fire danger models for predicting ignition risks.

Figure 3 shows the DMC, which tracks the moisture content in decomposing organic material. Unlike FFMC, DMC exhibits a slower response to environmental changes, reflecting the deeper moisture reservoir it represents. The ODE-based model effectively captures the gradual accumulation of moisture during extended wet periods and its depletion during prolonged dry spells. For instance, during multi-day precipitation events, the ODE model demonstrates smoother transitions compared to the discrete method, which tends to overestimate moisture recovery due to its stepwise adjustments. This continuous integration of precipitation effects in the ODE framework highlights its advantage in modeling long-term moisture dynamics.

Figure 3. DMC dynamics under varying precipitation conditions: The ODE-based approach (orange) captures moisture accumulation trends more effectively compared to the discrete method (blue).



Additionally, the drying term in the governing equations dominates during dry periods, allowing the ODE model to provide a realistic depiction of moisture depletion under high temperatures. These findings are particularly important for fire spread prediction in ecosystems with thick organic layers, such as coniferous forests, where duff fuels play a significant role in sustaining fires.

Figure 4 focuses on the Drought Code (DC), an index representing moisture trends in deep organic layers. The DC evolves more gradually than both FFMC and DMC, as it reflects long-term drying processes and cumulative precipitation effects. The ODE-based model provides a continuous and smooth representation of DC, particularly during periods of frequent, low-intensity rainfall. In contrast, the discrete method shows abrupt shifts that may misrepresent the gradual recovery of moisture in deep organic layers. The drying term, driven by temperature, dominates during extended dry periods, leading to a slow depletion of moisture reserves. Meanwhile, precipitation adjustments in the ODE model ensure that rainfall effects are integrated over time, providing a realistic depiction of moisture dynamics. These results underscore the importance of monitoring DC trends for fire danger assessments in ecosystems with deep organic layers, such as peatlands, where prolonged dry conditions can lead to elevated fire risks.

Figure 4. Continuous representation of DC trends: The ODE model (orange) offers smoother transitions during rainfall events than the discrete method (blue), highlighting cumulative precipitation effects.

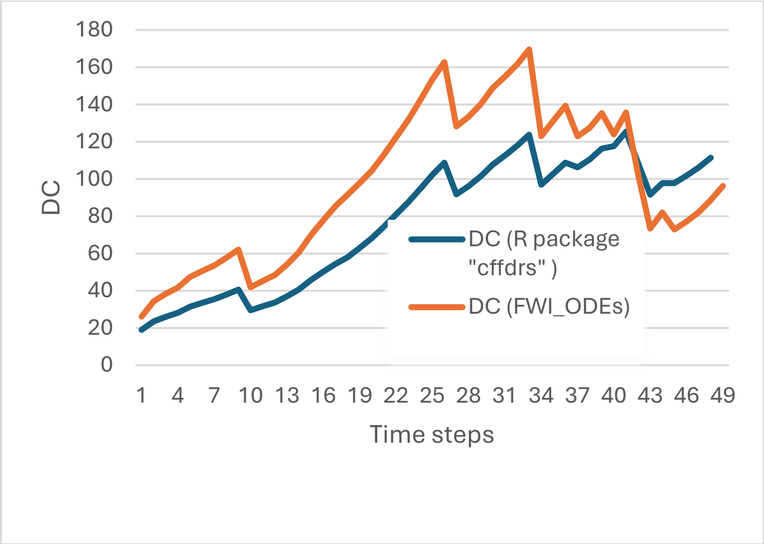


Figure 5 examines ISI, which estimates the potential rate of fire spread based on wind speed and fine fuel dryness. The ODE-based model captures rapid fluctuations in ISI caused by temporal variations in wind speed, particularly during windy conditions. For example, the model highlights spikes in ISI during periods of increased wind, which can significantly enhance fire spread potential. The traditional discrete method, on the other hand, fails to capture such short-term variability, offering a smoothed representation that may underestimate fire risks during critical periods. This capability of the ODE model to incorporate real-time wind variability is particularly valuable for operational fire management, as it enables more accurate assessments of fire spread potential during high-risk conditions.

Figure 5. Temporal evaluation of ISI: The ODE-based method (orange) reveals finer temporal details and nonlinear interactions compared to the traditional method (blue).

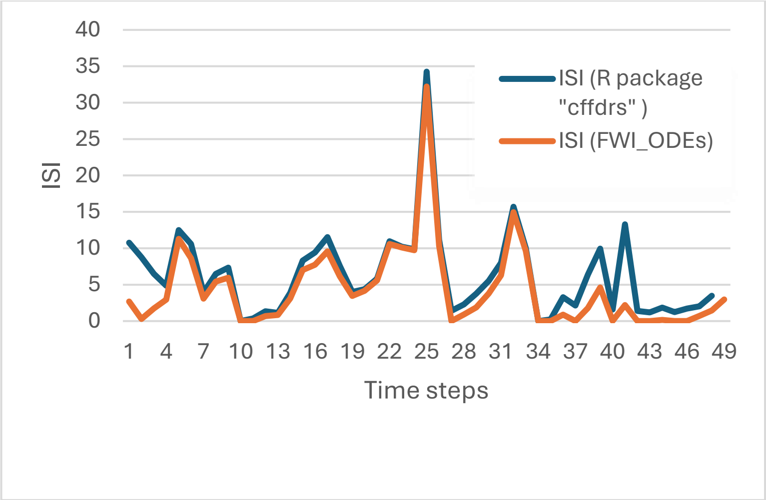
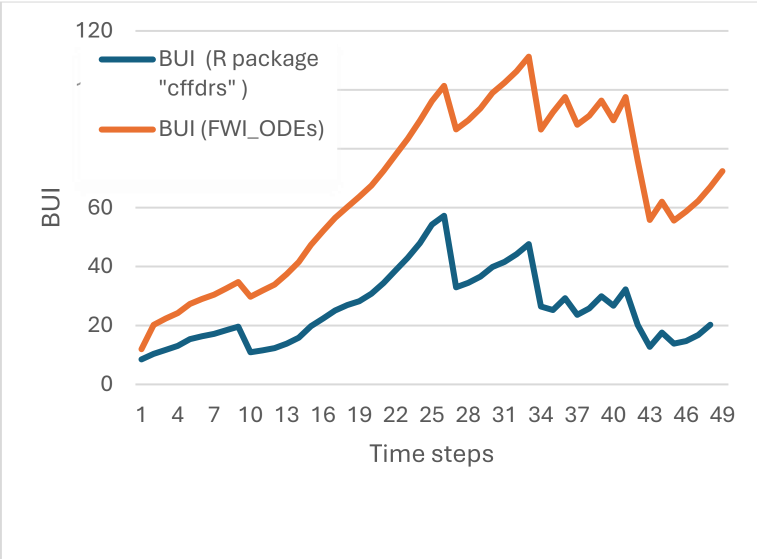


Figure 6 explores the BUI, which integrates the DMC and DC to estimate the total amount of fuel available for combustion. The ODE-based model offers a more continuous representation of BUI dynamics compared to the discrete method, which tends to overestimate fuel availability during periods of drying. The nonlinear interaction between ISI and BUI is critical for determining fire intensity, as high ISI values can exponentially amplify FWI during peak fire danger periods. The ODE model captures this interaction effectively, providing a more nuanced understanding of how short-term meteorological conditions (via ISI) and long-term fuel moisture trends (via BUI) combine to shape fire behavior. This interaction highlights the dual importance of monitoring both immediate weather conditions and cumulative fuel availability for effective fire management.

Figure 6. Comparison of the BUI computed with the R package "cffdrs" using both the traditional discrete method (blue) and the proposed ODE-based method (orange).



The ODE-based model consistently outperforms the traditional discrete method by capturing subtle, short-term meteorological fluctuations while maintaining alignment with long-term fire danger trends. Its ability to provide continuous updates enhances the precision of fire danger assessments, making it a valuable tool for operational fire forecasting and resource allocation.

4. CONCLUSIONS

This study introduces an innovative reformulation of the Canadian Forest Fire Weather Index system using an Ordinary Differential Equations-based framework. The ODE model addresses the limitations of the traditional discrete approach by enabling continuous-time integration of meteorological inputs, such as temperature, wind, relative humidity, and precipitation. Through this dynamical systems approach, the model captures subtle intraday variations in fire danger indices, offering a more responsive and precise tool for fire management.

The results demonstrate that the ODE-based framework aligns closely with the traditional FWI system in terms of overall trends while providing significant improvements in representing rapid changes in meteorological conditions. For instance, the ODE model captures midday drying of fine fuels and immediate responses to rainfall more effectively than the discrete method. These advantages are particularly evident in indices like the FFMC and ISI, which are highly sensitive to short-term fluctuations. The model's ability to continuously adjust long-term indices, such as the DC, ensures a realistic representation of cumulative moisture trends over time, further enhancing its utility in fire danger assessment.

From an operational perspective, the ODE model's capacity for real-time integration with meteorological systems positions as a valuable tool for early warning and resource allocation during fire events. By providing dynamic updates and improved resolution, the model can enhance situational awareness and decision-making for fire managers. Moreover, the ODE framework's compatibility with advanced numerical weather prediction (NWP) systems creates opportunities for seamless integration into existing fire forecasting platforms.

Despite its advantages, the ODE-based model has some limitations. Minor discrepancies were observed in indices like the DC under conditions of prolonged rainfall or low-intensity precipitation, likely due to the inherent assumptions in the numerical methods. Additionally, the computational demands of real-time implementation require further optimization to ensure operational feasibility. Future research should explore alternative numerical techniques to improve stability and accuracy, as well as hybrid approaches that integrate machine learning with the ODE model to enhance predictive performance.

In conclusion, the ODE-based framework represents a significant advancement in fire danger modeling, offering both scientific and practical benefits. Its ability to address the dynamic nature of fire weather conditions and its potential for integration with modern meteorological systems pave the way for more effective fire risk management. This study lays the groundwork for further innovations in fire danger assessment, contributing to improved resilience against wildfire threats in a changing climate.

5. REFERENCES

1. Evans, L. C. (2010). Partial Differential Equations. Graduate Studies in Mathematics. American Mathematical Society.
2. Feng, Z., Li, F., Lv, Y., & Zhang, S. (2016). A note on Cauchy–Lipschitz–Picard theorem. Journal of Inequalities and Applications, 2016, Paper No. 271, 6 pp. https://doi.org/10.1186/s13660-016-1214-x.
3. Jones, M. W., Kelley, D. I., Burton, C. A., Di Giuseppe, F., Barbosa, M. L. F., Brambleby, E., ... & Xanthopoulos, G. (2024). State of wildfires 2023–24. Earth System Science Data Discussions.
4. Kolmogorov, A. N., & Fomin, S. V. (1970). Introductory Real Analysis (R. A. Silverman, Trans. & Ed.). Dover Publications, Inc.
5. Lawson, B. D., & Armitage, O. B. (2008). Weather guide for the Canadian Forest Fire Danger Rating System. Natural Resources Canada, Canadian Forest Service. http://cfs.nrcan.gc.ca/pubwarehouse/pdfs/29152.pdf.
6. Lindelöf, E. (1900). Démonstration de quelques théorèmes sur les équations différentielles. Journal de Mathématiques Pures et Appliquées, 6, 423–441.
7. Lewy, H. (1957). An example of a smooth linear partial differential equation without solution. Annals of Mathematics, Second Series, 66(1), 155–158.
8. Pausas, J. G., & Keeley, J. E. (2021). Wildfires and global change. Frontiers in Ecology and the Environment, 19(7), 387–395.
9. Shi, K., & Touge, Y. (2022). Characterization of global wildfire burned area spatiotemporal patterns and underlying climatic causes. Scientific Reports, 12(1), 644.
10. Van Wagner, C. E. (1987). Development and structure of the Canadian Forest Fire Weather Index System. Forestry Technical Report 35. Canadian Forest Service.
11. Wang, X., Oliver, J., Swystun, T., Hanes, C. C., Erni, S., & Flannigan, M. D. (2023). Critical fire weather conditions during active fire spread days in Canada. Science of the Total Environment, 869, 161831.

Chapter 8

Recent Trends in Automatic Radiology Report Generation from Medical Images

İlhami SEL [1]

INTRODUCTION

Radiology, a cornerstone of modern healthcare, is a visual discipline that plays a critical role in the diagnosis and monitoring of diseases. With the advent of advanced imaging technologies, it has become possible to detect diseases at early stages, monitor the effectiveness of treatments, and support clinical decision-making. One of the most crucial components of this process is the radiology report, which conveys imaging findings in a structured and interpretive format, thereby guiding clinicians through treatment planning [1,2]. However, generating a radiology report is not merely a technical task; it requires attention to detail, clinical experience, and considerable time. The literature reports that approximately 3–5% of radiology reports contain errors, and up to 35% include ambiguous language [2]. Such errors can directly impact clinical decision-making and pose significant risks to patient safety. Additionally, many countries are experiencing a shortage of radiologists, leading to delays in reporting and increased workloads. For instance, a study in the United Kingdom found that 97% of imaging departments were unable to meet demand in a timely manner [3].

In this context, the integration of artificial intelligence (AI), particularly deep learning (DL)-based methods, into healthcare has gained significant momentum. The healthcare sector, which produces large volumes of data—over 90% of which are related to radiology—offers a fertile ground for AI research [4]. Recent advances in computer vision and natural language processing have enabled automated systems to extract meaningful insights from medical images and generate corresponding textual reports. As a result of these technological developments, the field of Automatic Radiology Report Generation (ARRG) has emerged.

[1] Dr., Inonu University Faculty of Engineering Department of Computer Engineering
Orcid: 0000-0003-0222-7017

ARRG is an interdisciplinary field aimed at producing meaningful, accurate, and clinically useful textual reports directly from medical images. These systems integrate both image analysis and natural language generation capabilities. While ARRG tasks resemble image captioning problems, they are significantly more complex in the medical domain [5,6]. For example, radiological images often exhibit high visual similarity; pathological findings may appear only in small regions of the image, yet these findings can drastically alter the content of a report. Therefore, traditional image captioning models fall short in this context. The first ARRG studies were initiated in 2015 by Shin et al., who utilized the Indiana University chest X-ray dataset [7]. Initially, visual features were extracted using convolutional neural networks (CNNs), and text generation was performed using recurrent neural networks (RNNs). However, these early systems typically produced templated, repetitive sentences with limited variability. Subsequent developments introduced the Transformer architecture [8] and large language models (e.g., BERT, GPT), which have since become widely adopted in this domain.

ARRG systems are not only relevant for text generation but also serve as clinical decision support tools. They can help reduce the workload of radiologists, assist in the preliminary assessment of routine cases, and offer second opinions to less experienced practitioners [9]. Moreover, the development of systems capable of highlighting potential abnormalities and generating personalized, template-based reports could enhance the overall quality of healthcare delivery. In recent years, there has been a marked increase in the number of publications focused on ARRG. Cutting-edge deep learning techniques such as contrastive learning, curriculum learning, knowledge graphs, and multimodal inputs have further enriched the diversity of research in this area [8].

Nevertheless, existing reviews in the ARRG literature often focus solely on chest X-rays or cover only a narrow range of training strategies and model architectures [1,6,8]. This book chapter aims to provide a comprehensive overview of the most recent studies published in the past three years. It explores a wide array of topics including datasets, training strategies, model architectures, knowledge integration, evaluation methods, and future research directions within a unified framework. Ultimately, Automatic Radiology Report Generation is not merely a technological advancement but a multidimensional and interdisciplinary research domain with the potential to transform clinical workflows. A deeper understanding and further development of this field promise substantial contributions to both academic research and clinical practice.

2. FUNDAMENTALS OF AUTOMATIC RADIOLOGY REPORT GENERATION

Automatic Radiology Report Generation (ARRG) is an artificial intelligence task that aims to generate textual outputs containing radiological interpretations from one or more medical images. The success of such systems depends heavily on the effective integration of image processing and natural language generation techniques. ARRG systems hold the potential to accelerate and standardize diagnostic workflows by mimicking human expertise [5,6]. The fundamental structure of ARRG can be broadly summarized as the extraction of visual features from an image, the transformation of these features into a contextual vector, and the generation of meaningful, structured text based on this vector. This process is typically implemented through an encoder–decoder architecture, where the encoder analyzes the image to generate meaningful representations, and the decoder uses these representations to generate textual descriptions.

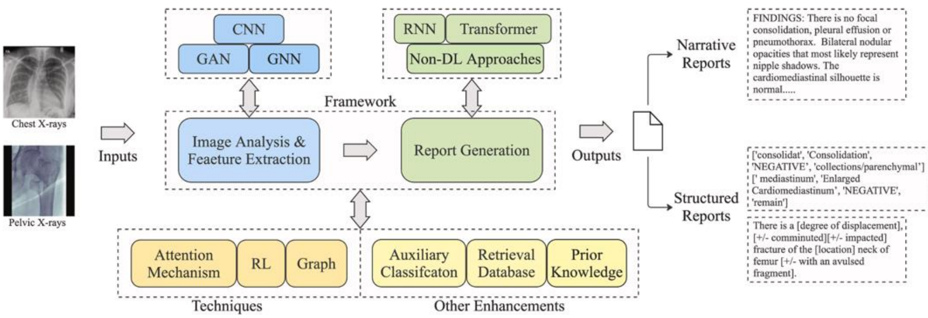


Fig. 1. Overview of the ARRG system workflow[6].

2.1 Input Data

ARRG systems generally rely on the following types of input data [8]:

- Medical images: Radiological modalities such as chest X-rays, computed tomography (CT), magnetic resonance imaging (MRI), and mammography.
- Metadata: Additional clinical context such as patient age, gender, medical history, and referral notes.
- Previous images and reports: Prior examinations to enable longitudinal analysis.

This multimodal data enables the system to produce more accurate and context-aware reports.

2.2 Visual Feature Extraction

The extraction of meaningful visual features from medical images is the first technical step in ARRG systems. Common visual encoders include [1,8]:

- CNN-based architectures: Deep convolutional neural networks like DenseNet and ResNet are used to generate low-dimensional vector representations from images.
- Vision Transformer (ViT): Processes the image as a sequence of patches using transformer architecture. Domain-specific variants such as Swin Transformer and MedViT are increasingly adopted in medical imaging [5].
- Segmentation networks: Architectures such as U-Net are used to detect and highlight abnormal regions [10].

These techniques allow the model to focus on clinically relevant visual regions.

2.3 Textual Representation and Language Generation

Radiology reports are typically structured into distinct sections such as Findings and Impression [11]. Generating these sections meaningfully requires advanced language modeling capabilities.

Commonly used decoders include:

- RNN / LSTM / GRU-based decoders: Widely used in early ARRG models [12].
- Transformer-based decoders: Modern models such as GPT, BERT, and BART offer improved contextual understanding and are capable of managing longer dependencies [6,13].
- Hierarchical decoders: Architectures that first generate sentence-level structures and then decompose them into words. These are particularly effective for multi-sentence reports [14].

2.4 Implementation of the Encoder–Decoder Architecture

The encoder–decoder architecture serves as the backbone of ARRG systems. The encoder analyzes the medical images, and the decoder transforms the resulting representations into natural language. Initial implementations used CNN–LSTM models, but transformer-based encoder–decoder frameworks have since become dominant [1].

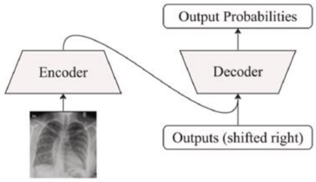


Fig. 2. Overview of the encoder-decoder framework [6].

Several studies have also proposed auxiliary modules to enhance this architecture, including:

- Contrastive attention mechanisms: Enable knowledge transfer from similar cases [15].
- Memory modules: Facilitate the reuse of previously acquired knowledge [16].
- Knowledge graphs: Provide semantic context through structured medical ontologies [17].

2.5 Multimodal Learning

Due to its inherent complexity, ARRG is a multimodal task. Advanced systems process not only visual data but also clinical directives, patient history, and prior reports. These additional inputs can be incorporated in the following ways:

- Early fusion: All data sources are merged at the encoder stage.
- Late fusion: Visual and non-visual data are processed separately and integrated at the decoder stage [8].

Such integrations enhance contextual awareness and enable the generation of more personalized and clinically relevant reports.

2.6 Report Generation Strategies

Two primary strategies are used for generating radiology reports:

- Fully generative systems: The model generates each word from scratch, as in GPT-based architectures.
- Hybrid systems (retrieval + generation): The model can retrieve sentences from past reports and modify them when necessary [1].

These strategies have a direct impact on the clinical validity, originality, and coherence of the generated reports.

2.7 Evaluation Process

ARRG outputs must be evaluated from both linguistic and clinical perspectives. Key evaluation metrics include:

- NLP metrics: BLEU, ROUGE, CIDEr, and METEOR assess word-level similarities between generated and reference texts [1].
- Clinical metrics: Tools like CheXbert and RadGraph evaluate medical correctness by assessing the alignment of clinical entities and relationships [18].
- Qualitative evaluation: Blind reviews by medical experts assess the interpretability and reliability of the generated reports [19].

These evaluation mechanisms are critical not only for assessing model performance but also for determining clinical applicability.

3. DATASETS

One of the most critical requirements for developing deep learning-based systems in the field of Automatic Radiology Report Generation (ARRG) is the availability of high-quality, annotated, and multimodal (image + text) medical datasets. However, due to concerns related to privacy, ethics, and restricted access, open-source medical datasets are limited in number. This creates a significant bottleneck in the model development process [1,8]. This section reviews the most commonly used datasets in the ARRG domain, along with their characteristics, advantages, and limitations.

3.1 Indiana University X-ray Dataset (IU-Xray)

Also known as the OpenI dataset, IU-Xray is one of the first publicly accessible datasets in the ARRG field. Published in 2016, it was the first to combine chest X-ray images with their corresponding radiology reports [6].

The structure of IU-Xray loosely aligns with the report format recommended by the European Society of Radiology. Furthermore, it is enriched with labels automatically generated by the Medical Text Indexer (MTI), enhancing its utility for various research tasks.

3.2 MIMIC-CXR

MIMIC-CXR is currently the largest publicly available, multimodal, and richly annotated ARRG dataset. Sourced from Beth Israel Deaconess Medical Center (Boston, USA), the data was collected between 2011 and 2016 and released through the PhysioNet platform.

3.3 MIMIC-Derived Datasets

The comprehensive content of MIMIC-CXR has led to the creation of several derivative datasets, each designed to test ARRG systems under specific scenarios [1]:

- MIMIC-ABN: Consists of reports containing at least one abnormal finding.
- CXR-PRO: Applies filtering to remove references to previous reports.
- Chest ImaGenome: Provides scene graphs that annotate 29 anatomical regions per image.
- MS-CXR and MS-CXR-T: Target tasks involving image–text alignment and time series analysis.

3.4 Other Open-Access Datasets

- PadChest: Collected in Spain, it includes 160,868 images and 109,931 reports in Spanish. Adoption has been limited due to the language barrier [1].
- COV-CTR: Contains COVID-19-specific CT reports and includes both English and Chinese data [6].
- CX-CHR: Comprises 33,236 Chinese chest X-ray reports; however, access is restricted [8].
- CT-RATE: The first large-scale 3D thoracic CT dataset published on Hugging Face.
- JLiverCT, CH-Xray: These are specialized datasets focusing on brain, liver, and stroke MRI images, respectively [1].

3.5 General Evaluation

The success of ARRG systems is not solely determined by architecture or training methodology, but also by the scope, quality, and realism of the training data. Datasets such as IU-Xray are useful for rapid prototyping due to their small size and accessibility, whereas MIMIC-CXR and its derivatives support the development of more realistic systems. However, the lack of standardization across datasets presents challenges for fair model comparison [20]. As a result, the literature emphasizes the need for more diverse datasets—such as those covering brain MRI or abdominal CT scans—as well as the development of multilingual corpora [1,6].

4. TRAINING METHODOLOGIES

The performance of Automatic Radiology Report Generation (ARRG) systems depends not only on model architecture but also heavily on the training strategies employed. The training process enables the model to learn how to interpret medical images, generate semantically rich text, and produce clinically accurate content. This section categorizes the main training methods used in ARRG into five major approaches: supervised learning, unsupervised learning, contrastive learning, triplet loss, and reinforcement learning.

4.1 Supervised Learning

Supervised learning is the most commonly applied training strategy in ARRG systems. In this paradigm, the model is trained on paired data consisting of input images and their corresponding radiology reports [1,6,8]. The model learns from these examples with the goal of generating new reports for similar cases. One notable approach within supervised learning is curriculum learning, where the model is first trained on simpler examples and then progressively exposed to more complex instances. A multimodal version of this strategy was proposed by designing a difficulty scoring system that accounts for both visual and textual complexity [21]. Wang et al. [5] enhanced model performance by incorporating auxiliary tasks such as image-text alignment, image classification using MeSH labels, and text generation enriched with medical knowledge. Similarly, Li et al. [22] augmented their model with auxiliary signals based on medical knowledge derived from both images and text. This included segmentation of abnormal regions in images and incorporation of disease–symptom relationships from textual data.

4.2 Unsupervised Learning

Fully unsupervised learning remains relatively rare in ARRG research. Nonetheless, several studies have attempted to reduce reliance on labeled data. Liu et al. [23] created a shared embedding space between visual and textual modalities using pre-constructed knowledge graphs, facilitating alignment between image and text features. Yan et al. [24] introduced a variational autoencoder (VAE)-based framework, where features extracted via a CNN–Transformer architecture were modeled as Gaussian distributions for use in text generation. Li et al. [25] employed HDBSCAN clustering to extract information clusters from reports and transferred knowledge based on these clusters. Wang et al. [26] applied K-means clustering to visual and textual features to create memory matrices, which were then integrated into the model's memory components.

4.3 Contrastive Learning

Contrastive learning is a powerful method for learning robust representations, particularly in low-resource settings. In this technique, positive pairs (similar examples) are brought closer together in the embedding space, while negative pairs (dissimilar examples) are pushed apart [1].

- Different contrastive learning strategies have been applied in ARRG systems:
- Image-based contrast: Positive samples are generated through augmented versions of the same image.
- CLIP-based approaches: Large pre-trained models such as CLIP and ALBEF, which jointly process image and text, are widely used.
- Hard negative mining: Clustering or semantic similarity-based techniques are used to select challenging negative examples.

4.4 Triplet Loss

Triplet loss is a specific form of contrastive learning that involves three-part training samples:

- Anchor (reference example),
- Positive (similar example),
- Negative (dissimilar example).

The objective is to minimize the distance between the anchor and positive while maximizing the distance between the anchor and negative. Wang et al. [27] implemented a dual-branch network with an auxiliary image-text matching task and applied triplet loss. Yang et al. [28] proposed an inverted triplet structure where abnormal examples were chosen as anchors to improve the model's sensitivity.

4.5 Reinforcement Learning (RL)

Reinforcement learning enables a model to learn optimal behaviors through reward signals. In ARRG systems, task-specific reward functions—beyond conventional NLP metrics—are often defined to reflect clinical relevance. Some studies have directly used BLEU or ROUGE scores as reward functions [1]. However, due to the limited clinical validity of these metrics, semantic and medically-informed reward functions are generally preferred.

4.6 General Evaluation

Current trends in ARRG training methodologies can be summarized as follows:

- Supervised learning remains the dominant paradigm, increasingly enriched through difficulty-based sampling and multimodal curriculum strategies.
- Contrastive learning enables effective representation learning in data-scarce scenarios.
- Reinforcement learning allows for the generation of clinically meaningful reports by optimizing for domain-specific reward signals.
- Unsupervised and triplet loss-based methods are less explored but hold significant potential for advancing ARRG capabilities.

In conclusion, the diversification of training strategies and the adoption of specialized techniques are contributing to the advancement of more reliable and high-performing ARRG systems.

5. ARCHITECTURAL APPROACHES

The architectural design of Automatic Radiology Report Generation (ARRG) systems determines how medical images are processed, what information is emphasized, and the strategy by which meaningful textual reports are produced. A successful ARRG model must accurately capture both visual and linguistic representations and integrate multimodal information to generate clinically relevant narratives. This section categorizes the primary architectural frameworks employed in ARRG into four major groups: CNN–LSTM-based models, Transformer architectures, pretrained language models, and attention mechanisms.

5.1 CNN–LSTM-Based Architectures

Early ARRG systems were inspired by classical image captioning models and primarily adopted CNN–RNN architectures [6]. In such frameworks:

- Encoder (CNN): Extracts low-dimensional feature representations from visual input.
- Decoder (LSTM or GRU): Generates sequential text based on the encoded features.

Jing et al. [12] implemented visual feature extraction using CNNs and used a hierarchical LSTM for report generation. Similarly, Zhang et al. [29] employed

DenseNet-121 for feature extraction and integrated knowledge graphs via graph convolutional networks (GCNs) to guide an LSTM-based decoder.

Advantages:

- Simpler implementation and lower computational cost.
- Can be trained effectively on smaller datasets.

Limitations:

- Tend to produce repetitive and templated sentences.
- Struggle with maintaining contextual coherence in long reports [5].

5.2 Encoder–Decoder Transformer Architectures

Transformer architectures, which revolutionized natural language processing (NLP), have also become prevalent in ARRG systems [8]. These architectures can be implemented in two main ways:

- Hybrid Approach: Visual features are first extracted via a CNN and then processed by a Transformer encoder.
- Fully Transformer-Based: Visual inputs are directly encoded using Vision Transformers (ViT).

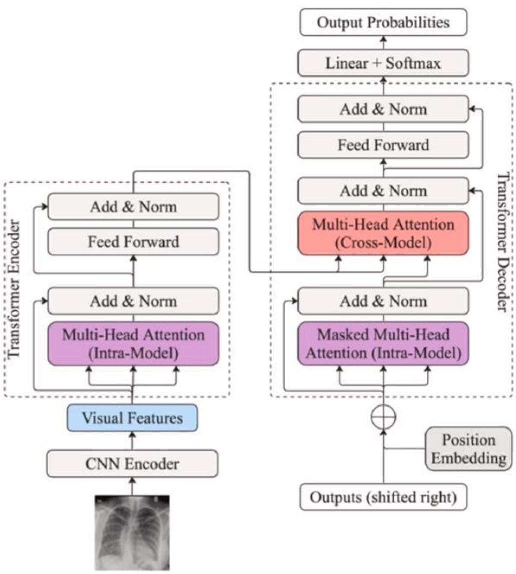


Fig. 3. Overview of the transformer encoder-decoder framework [6].

Nooralahzadeh et al. [30] proposed a two-stage system combining Meshed-Memory Transformer for high-level concept extraction and BART for fluent text generation. Dalla Serra et al. [31] integrated both images and knowledge triples (entity–relation–entity) to produce semantically enriched representations. Wang et al. [32] developed a "self-adaptive fusion gate" to combine global and regional features effectively.

5.3 Pretrained Language Models (PLMs)

In recent years, the integration of large pretrained language models into ARRG systems has significantly improved both the fluency and contextual accuracy of generated reports. PLMs have been used in both encoder and decoder roles to enhance language modeling capabilities in ARRG tasks.

5.4 Attention Mechanisms

In radiological images, pathological findings are often confined to small, localized regions. Therefore, attention mechanisms play a crucial role by enabling the model to focus on the most clinically relevant areas. These mechanisms help in enhancing interpretability and precision, particularly in high-resolution medical imaging tasks.

5.5 Hybrid and Specialized Architectures

Beyond conventional generative models, several innovative approaches have been developed in ARRG systems:

- Retrieval-Based Systems: Wang et al. [33] proposed a model capable of retrieving prewritten sentences when appropriate, improving both efficiency and accuracy.
- Topic-Guided Generation: Najdenkoska et al. [34] introduced a variational topic inference (VTI) mechanism that guides sentence-level generation based on semantic themes.
- Segmentation-Guided Control: Tanida et al. [35] used region segmentation via Faster R-CNN to determine which regions should guide sentence generation.

5.6 General Evaluation

As ARRG architectures become increasingly complex and multimodal, they offer greater potential for producing clinically accurate and coherent reports. Architectural choices influence not only the quality of the output but also the model's explainability, reliability, and practical applicability in clinical settings.

6. INTEGRATION OF KNOWLEDGE AND MULTIMODALITY

In Automatic Radiology Report Generation (ARRG), incorporating clinical context and auxiliary information into the model is just as crucial as processing visual representations. This section discusses the strategies for integrating multimodal information into ARRG models and enriching them semantically through structures such as knowledge graphs.

6.1 Multimodal Inputs

Given the nature of the ARRG task, models are often required to process not only images but also additional information such as clinical directives, patient history, prior reports, and laboratory results. Multimodal inputs are generally integrated using two primary strategies:

- Early Fusion: All data modalities—images, clinical text, demographic data—are merged at the encoder level into a shared representation.
- Late Fusion: Visual and non-visual modalities are processed independently and then combined during the decoding stage or the final reasoning layer.

These strategies enable the model to generate more targeted and context-aware outputs by simultaneously considering both visual evidence and auxiliary clinical context.

6.2 Temporal and Longitudinal Information Utilization

For monitoring chronic conditions or evaluating treatment response, the analysis of sequential imaging examinations is essential. Accordingly, some ARRG systems incorporate longitudinal data, allowing the model to consider not only the current findings but also the temporal evolution of the patient's condition. This leads to reports that are more clinically meaningful and informative.

6.3 Knowledge Graphs

Knowledge graphs represent semantic relationships between medical concepts using nodes and edges. In ARRG systems, incorporating knowledge graphs into the model architecture adds a clinical knowledge layer to the textual output. These graphs enable the model to better understand medical terminology, recognize relationships among clinical concepts, and reflect these associations in the generated report content.

6.4 Clinical Knowledge Integration

Beyond knowledge graphs, the integration of other forms of structured or semi-structured clinical data can also enhance model performance. These approaches help the model better comprehend complex clinical scenarios by providing richer context beyond the image alone.

6.5 General Evaluation

Multimodal information integration and the use of knowledge graphs contribute to improving the clinical consistency, semantic richness, and reliability of ARRG systems. However, these approaches also introduce additional computational costs, data preparation complexity, and ethical/privacy concerns. Future directions in the field of multimodality and knowledge-based modeling may include:

- Establishing standardized representations,
- Promoting concept-level annotation in datasets,
- Developing lightweight and interpretable multimodal architectures [1].

7. EVALUATION METHODS

Evaluating the performance of Automatic Radiology Report Generation (ARRG) systems is a multidimensional process that encompasses not only linguistic accuracy but also clinical relevance. Unlike general natural language generation tasks, ARRG evaluation must consider the output's contribution to disease diagnosis and clinical decision-making. Therefore, ARRG models are assessed from several perspectives, including linguistic metrics, clinically informed metrics, and qualitative evaluation strategies [1,6,8].

7.1 Natural Language Processing (NLP) Metrics

Early evaluations of ARRG systems primarily used general-purpose NLP metrics to measure textual similarity between model outputs and reference reports [36]. These metrics typically operate at the word level:

- BLEU (Bilingual Evaluation Understudy): Measures n-gram precision; tends to favor short, template-like outputs.
- ROUGE (Recall-Oriented Understudy for Gisting Evaluation): Emphasizes recall relative to the reference; useful for longer texts.
- CIDEr (Consensus-based Image Description Evaluation): Based on TF-IDF-weighted n-gram vectors, originally developed for image captioning.
- METEOR: Incorporates synonymy and morphological variations for better semantic matching.

- SPICE: Converts sentences into scene graphs and focuses on conceptual similarity.

Advantages: Easy to compute.

Limitations: Do not reflect clinical correctness. Penalize linguistic diversity due to reliance on single reference texts [1].

7.2 Clinical Relevance-Based Metrics

To evaluate ARRG outputs in terms of clinical validity, several specialized metrics have been developed:

7.2.1 CheXpert Labeler & CheXpert F1: Developed at Stanford, this labeler extracts sentence-level annotations for 14 thoracic conditions. The F1 score is calculated based on alignment between labels in the generated and reference reports [37].

7.2.2 RadGraph F1: Converts entities and relations in reports into semantic graphs. The Entity F1 and Relation F1 scores reflect the structural similarity between generated and reference graphs [38].

7.2.3 FactENT / FactENT-NLI: Proposed by Miura et al. [39], this method evaluates both the overlap and logical consistency of clinical entities between generated and gold-standard reports.

These metrics are especially crucial when evaluating outputs from large language models like GPT-3 and GPT-4V, which may produce grammatically correct but clinically flawed statements [40].

7.3 Qualitative Evaluation Methods

In addition to quantitative metrics, expert-based evaluation is commonly employed. This typically includes the following criteria [36]:

- Accuracy: How well does the report reflect actual findings?
- Usefulness: How helpful is it for clinical decision-making?
- Fluency: Is the report grammatically and syntactically coherent?
- Clarity: Are the findings clearly expressed, with minimal ambiguity?

Such evaluations are often conducted through blinded review procedures. For instance, Delbrouck et al. [38] tested whether experts could distinguish between real and generated reports to assess model realism. Wu et al. [41] measured inter-rater agreement among different expert groups.

7.4 Emerging and Proposed Metrics

To bridge the gap between traditional NLP metrics and clinical relevance, the following new evaluation metrics have been proposed [1]:

- BLEU-Rad / ROUGE-Rad: Apply domain-specific weighting to clinical terms.
- RadCliQ: A composite metric that jointly considers linguistic and clinical accuracy.
- AI-based evaluators: Systems based on ChatGPT or GPT-4 are being explored to emulate human judgment. However, these remain controversial and are still under investigation.

7.5 Comparative Performance Summary

ARRG performance varies significantly across datasets and evaluation metrics, depending on architectural and training choices. Meta-analysis results reported in [1] indicate the following trends:

- Transformer-based models outperform CNN–LSTM structures by 10–15% in ROUGE and CIDEr scores.
- CheXpert F1 and RadGraph F1 scores typically range between 0.50 and 0.65.
- The highest clinical accuracy is achieved by systems incorporating knowledge graphs and reinforcement learning [38].

7.6 General Evaluation

Evaluating ARRG systems requires a layered and multidimensional approach. This diversity is essential to ensure that ARRG models can be deployed safely and effectively in real-world clinical settings. In the future, it is expected that hybrid evaluation systems, combining multiple metrics and aligning with medical guidelines, will become more widespread.

8. FUTURE DIRECTIONS AND OPEN RESEARCH QUESTIONS

Automatic Radiology Report Generation (ARRG) represents one of the most clinically impactful applications of medical artificial intelligence. However, the current literature emphasizes that these technologies are still at the research stage, and numerous technical, ethical, and practical challenges must be addressed before they can be deployed in real-world clinical settings. This section outlines the key future research trends, knowledge gaps, and open questions highlighted in three comprehensive review studies.

8.1 Larger and More Diverse Datasets

Most existing ARRG models have been trained on chest X-ray datasets. Yet, a significant portion of clinical imaging involves other modalities such as CT, MRI, and ultrasound. Public, high-quality, report-linked datasets for these modalities remain scarce [1,6].

Research Questions:

- How can ethically accessible and balanced open datasets be constructed for CT and MRI?
- How can multimodal ARRG datasets covering anatomical diversity be standardized?

8.2 Standardized Evaluation Protocols

Due to variations in metrics, dataset splits, and preprocessing pipelines, comparing ARRG systems remains difficult.

Research Questions:

- Can the creation of benchmark subsets and fixed data splits lead to more equitable model comparisons?
- How can evaluation metrics that reflect clinical accuracy be standardized?

8.3 Integration with Expert Systems

For real-world implementation, ARRG systems must be interoperable with clinical decision support systems and capable of facilitating human–machine collaboration.

Research Questions:

- How can ARRG systems be designed to support radiologists rather than replace them?
- How can semi-automated models be trained to incorporate user feedback in real time?

8.4 Explainability and Trustworthiness

To ensure safe clinical use, ARRG models must offer explainability—answers to "Why did the model generate this report?"—yet most current models function as black boxes [42].

Research Questions:

- How can explainable models be built to link each generated sentence to its corresponding visual evidence (e.g., segmentation maps)?
- How can systems be designed to alert users when the model is uncertain or detects ambiguous findings?

8.5 Role of Large Multimodal Models (Multimodal LLMs)

Massive multimodal models like GPT-4V, PaLM-E, and Med-PaLM have the potential to revolutionize image-to-text generation. However, their applicability to medical domains remains underexplored [8].

Research Questions:

- How can large multimodal language models be fine-tuned specifically for ARRG tasks?
- How can ethical concerns, data privacy, and misinformation risks be mitigated during training?

8.6 Integration of Medical Knowledge and Guideline-Based Reporting

Incorporating knowledge graphs, medical coding systems (e.g., ICD, SNOMED CT), and rule-based inference engines can improve clinical validity [20].

Research Questions:

- How can structured report generation be aligned with clinical guidelines?
- Can ARRG outputs be mapped to standard disease classification systems like ICD-10?

8.7 Ethics, Safety, and Regulation

Clinical deployment of ARRG systems requires not only technical robustness but also ethical and regulatory compliance.

Research Questions:

- Who holds responsibility for automatically generated reports?
- How can errors be detected and corrected in a systematic way?
- How can international standards for data anonymization and patient privacy be ensured?

8.8 Human–Machine Collaboration and Semi-Automated Systems

Rather than fully automated systems, semi-automated models that allow user feedback may offer greater safety and adoption potential.

Research Questions:

- How can hybrid systems be designed so that model-generated reports can be edited and approved by human experts?
- How can models be continuously improved through active learning from user feedback?

8.9 General Perspective

In the near future, ARRG systems are expected to become significantly more robust through:

- Larger and more diverse datasets,
- Standardized evaluation metrics,
- Explainable hybrid models built for human collaboration,
- Ethically aligned, clinically grounded reporting systems.

9. CONCLUSION

Automatic Radiology Report Generation (ARRG) is a pioneering field in medical AI, notable for its technical complexity and potential clinical impact. These systems aim to generate accurate, interpretable, and clinically useful reports from medical imaging data—thereby reducing radiologists' workload and accelerating diagnostic workflows.

The body of literature analyzed in this chapter shows that ARRG has undergone a significant transformation in the past five years. Early systems based on simple CNN–LSTM architectures have evolved into complex multimodal frameworks incorporating Vision Transformers, large language models (e.g., BERT, GPT-4V), knowledge graphs, and reinforcement learning.

The following key enablers have driven this evolution:

- The proliferation of large-scale, multimodal datasets like MIMIC-CXR,
- Successful adaptation of Transformer and PLM architectures to medical data,
- Enhancement of clinical relevance through contrastive and reinforcement learning strategies,
- Semantic enrichment via knowledge graphs,
- Improved evaluation with clinically informed metrics like RadGraph and CheXpert F1.
- Nevertheless, clinical integration remains limited, due to:
- Data scarcity for modalities beyond chest X-ray,
- Lack of standardization in benchmarks and evaluation protocols,
- Insufficient model explainability and unclear clinical reliability,
- Ongoing ambiguity in legal and ethical frameworks.

Future research is expected to prioritize:

- Larger and more diverse datasets (e.g., CT, MRI),
- Explainable, human-in-the-loop hybrid systems,

- Tailored fine-tuning of large multimodal LLMs (e.g., GPT-4V, PaLM-E) for ARRG,
- Semi-automated, guideline-aligned report generation,
- Embedding ethical safety layers into ARRG systems.

These advances will shift ARRG beyond mere text generation to becoming an integral component of clinical decision support. Achieving this goal will require interdisciplinary collaboration—not only from engineers but also from clinical scientists, ethicists, legal experts, and user experience designers.

REFERENCES

[1] Sloan, Phillip, et al. "Automated radiology report generation: A review of recent advances." IEEE Reviews in Biomedical Engineering (2024).

[2] A. P. Brady, “Error and discrepancy in radiology: Inevitable or avoid-able?,” Insights Into Imag., vol. 8, no. 1, pp. 171–182, 2017.

[3] A. Rimmer, “Radiologist shortage leaves patient care at risk, warns royal college,” BMJ (Online), vol. 359, 2017, Art. no. j4683.

[4] S. K. Zhou et al., “A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises,” Proc. IEEE, vol. 109, no. 5, pp. 820–838, May 2021.

[5] Z. Wang et al., “Automated radiographic report generation purely on transformer: A multicriteria supervised approach,” IEEE Trans. Med Imag., vol. 41, no. 10, pp. 2803–2813, Oct. 2022.

[6] Hamma, Moustapha Boubacar, et al. "Automatic Radiology Report Generation: A Comprehensive Review and Innovative Framework." 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2024.

[7] H.-C. Shin, L. Lu, L. Kim, A. Seff, J. Yao, and R. M. Summers, ‘Interleaved Text/Image Deep Mining on a Large-Scale Radiology Database for Automated Image Interpretation’, May 04, 2015, arXiv: arXiv:1505.00670. doi: 10.48550/arXiv.1505.00670.

[8] Liao, Yuxiang, Hantao Liu, and Irena Spasić. "Deep learning approaches to automatic radiology report generation: A systematic review." Informatics in Medicine Unlocked 39 (2023): 101273.

[9] Royal College of Radiologists, “The Standards for interpretation and reporting of imaging investigations, 2nd Ed,” 2018. [Online] Avail-able from: https://www.rcr.ac.uk/publication/standards-interpretation-and-reporting-imaging-investigations-second-edition

[10] W. Nimalsiri et al., “Automated radiology report generation us-ing transformers,” in Proc. Int. Conf. Adv. Res. Comput., 2023, pp. 90–95.

[11] European Society of Radiology (ESR), “Good practice for radiological reporting. Guidelines from the european society of radiology (ESR),” Insights Imag., vol. 2, pp. 93–96, 2011.

[12] ‘Good practice for radiological reporting. Guidelines from the European Society of Radiology (ESR)’, Insights Imaging, vol. 2, no. 2, pp. 93–96, Feb. 2011, doi: 10.1007/s13244-011-0066-7.

[13] A. Radford et al., “Language Models are Unsupervised Multitask Learn-ers,” 2019.

[14] Y. Zhang et al., “When radiology report generation meets knowledge graph,” in Proc. AAAI Conf. Artif. Intell., 2020, pp. 12910–12917.
[15] X. Song et al., “Cross-modal contrastive attention model for medi-cal report generation,” in Proc. Int. Conf. Comput. Linguistics, 2022, pp. 2388–2397.
[16] Z. Chen et al., “Generating radiology reports via memory-driven transformer,” in Proc. Conf. Empirical Methods Natural Lang. Process., 2020, pp. 1439–1449.
[17] D. Zhang et al., “Improving medical X-ray report generation by using knowledge graph,” Appl. Sci., vol. 12, no. 21, 2022, Art. no. 11111.
[18] J. Delbrouck et al., “Improving the factual correctness of radiology report generation with semantic rewards,” in Proc. Findings Empirical Methods Natural Lang. Process., 2022, pp. 4348–4360.
[19] W. Chaoyi et al., “Can GPT-4V(ision) serve medical applications? case studies on GPT-4 V for multimodal medical diagnosis,” 2023.
[20] W. Hou et al., “ORGAN: Observation-guided radiology report gener-ation via tree reasoning,” in Proc. Assoc. Comput. Linguistics, 2023, pp. 8108–8122.
[21] F. Liu et al., “Competence-based multimodal curriculum learning for medical report generation,” in Proc. Int Jt. Conf. Nat. Lang. Process., 2021, pp. 3001–3012.
[22] M. Li et al., “Auxiliary signal-guided knowledge encoder-decoder for medical report generation,” World Wide Web, vol. 26, pp. 253–270, 2023.
[23] F. Liu et al., “Auto-encoding knowledge graph for unsupervised medical report generation,” in Proc. Adv. Neural Inf. Process. Syst., 2021, vol. 34, pp. 16266–16279.
[24] B. Yan et al., “Prior guided transformer for accurate radiology reports generation,” IEEE J. Biomed Health Inform., vol. 26, no. 11, pp. 5631–5640, Nov. 2022.
[25] J. Li et al., “A self-guided framework for radiology report generation,” in Proc. Int. Conf. Med. Image Comput. Comput. Assist. Interv., 2022, vol. 13438, pp. 588–598.
[26] J. Wang et al., “Cross-modal prototype driven network for radiology report generation,” in Proc. Eur. Conf. Comput. Vis., 2022, vol. 13695, pp. 563–579.
[27] Z. Wang et al., “A self-boosting framework for automated radiographic report generation,” in Proc. IEEE/CVF Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 2021, pp. 2433–2442.
[28] Y. Yang et al., “A contrastive triplet network for automatic chest x-ray reporting,” Neurocomputing, vol. 502, pp. 71–83, 2022.

[29] Y. Zhang et al., “When radiology report generation meets knowledge graph,” in Proc. AAAI Conf. Artif. Intell., 2020, pp. 12910–12917.
[30] F. Nooralahzadeh et al., “Progressive transformer-based generation of radiology reports,” in Proc. Findings Conf. Empir. Methods Natural Lang. Process., 2021, pp. 2824–2832.
[31] F. Dalla Serra et al., “Multimodal generation of radiology reports using knowledge-grounded extraction of entities and relations,” in Proc. Conf. Int. Joint Conf. Nat. Lang. Process., 2022, pp. 615–624.
[32] Y. Wang et al., “Self adaptive global-local feature enhancement for radiology report generation,” in Proc. Int. Conf. Image Process., 2023, pp. 2275–2279.
[33] F. Wang et al., “Unifying relational sentence generation and retrieval for medical image report composition,” IEEE Trans. Cybern., vol. 52, no. 6, pp. 5015–5025, Jun. 2022.
[34] I. Najdenkoska et al., “Variational topic inference for chest X-ray report generation,” in Proc. Int. Conf. Med. Image Comput. Comput. Assist. Interv., 2021, pp. 625–635.
[35] T. Tanida et al., “Interactive and explainable region-guided radiology report generation,” in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 2023, pp. 7433–7442.
[36] Sel, İlhami, and Davut Hanbay. "Fully attentional network for low-resource academic machine translation and post editing." Applied Sciences 12.22 (2022): 11456.
[37] J. A. Irvin et al., “CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison,” in Proc. AAAI Conf. Artif. Intell., 2019, pp. 590–597.
[38] J. Delbrouck et al., “Improving the factual correctness of radiology report generation with semantic rewards,” in Proc. Findings Empirical Methods Natural Lang. Process., 2022, pp. 4348–4360.
[39] Y. Miura et al., “Improving factual completeness and consistency of image-to-text radiology report generation,” in Proc. Assoc. Comput. Linguist: Hum. Lang. Tech., 2021, pp. 5288–5304.
[40] W. Chaoyi et al., “Can GPT-4V(ision) serve medical applications? case studies on GPT-4 V for multimodal medical diagnosis,” 2023.
[41] Y. Wu et al., “Token imbalance adaptation for radiology report genera-tion,” in Proc. Conf. Health, Inference, Learn., 2023, vol. 209, pp. 72–85.
[42] Z. Chen et al., “Generating radiology reports via memory-driven trans-former,” in Proc. Conf. Empirical Methods Natural Lang. Process., 2020, pp. 1439–1449.