A survey and bibliometric analysis of scholarly literature in Data Envelopment Analysis*

Antonija Mandić, Katerina Fotova Čiković, and Damira Keček

Abstract—**The Data Envelopment Analysis (DEA) methodology is one of the most popular nonparametric linear programming methods. It was introduced in 1978 and has gained popularity throughout different industries and research areas. This bibliographic collection incorporates and provides an extensive analysis of DEA-related articles from its introduction in 1978 to February 2025. The main objective is to survey the Web of Science scientific database with the key phrase "DATA ENVELOPMENT ANALYSIS" and to present all the hotspots and applications DEA has inspired. The Bibliometrix R package provided a comprehensive set of techniques through its Biblioshiny interface. The findings of the bibliometric study reveal the annual scientific production, most relevant authors, top-cited and most relevant sources, most relevant affiliations, and thematic map of themes that are mainly tackled and associated with the overall application of DEA. The synthesis of knowledge and publications represents the impact of the DEA application and development since its inception.**

I. Introduction

Organizational efficiency and effectiveness are crucial for success in today's dynamic and rapidly changing business environment. Over time, there have been several definitions of efficiency in scholarly literature. Efficiency is a multifaceted concept with applications across various fields. Generally, it refers to using resources to maximize production [12]. Despite many definitions in scholarly literature, its main focus is optimizing resource utilization to effectively achieve desired outcomes.

This study aims to provide a comprehensive bibliographic analysis of DEA-related studies since the introduction of DEA methodology in 1978. For this purpose, the Web of Science scientific database was surveyed, and the data analysis was conducted using the R package Bibliometrix and Biblioshiny. The analysis of bibliometric results first presents the main bibliometric statistics, and subsequently, it considers authors, indicators, information, and the countries of research [29].

The findings of this study offer a holistic understanding of the DEA methodology and could inspire data analysts, managers, consultants, and academic members to incorporate this methodology at a larger scale in the future due to its large application potential.

The rest of this paper is structured as follows: Section 2 provides a theoretical background on the DEA. Section 3 thoroughly presents the research method, whereas Section 4 reveals the results. Section 5 discusses the most interesting findings and insights and concludes the paper.

II. Theoretical background on DEA

A. Introduction of DEA methodology

Data Envelopment Analysis (DEA) was introduced in 1978 as a nonlinear programming model to measure the efficiency of decision-making units in public programs [4]. Its earlier beginnings of application are related to the evaluation of the program follow-through experiment in U.S. public school education [5], to the evaluation of not-for-profit entities [6], and to the evaluation of the efficiency of a public education program [7]. The significance of DEA methodology was soon recognized in scientific circles, which led to a continuous and rapid growth in further development and application in the field of DEA-related research [14], [20]. The great interest of researchers and practitioners in the theoretical development and application of DEA as a versatile tool for measuring performance indicates the distinct popularity of the methodology based on numerous publications in both reputable journals and international conferences [19]. DEA is one of the most accepted and applied methodologies for evaluating the relative efficiency of entities or decision-making units with multiple inputs and outputs [1], [17], [26], [28]. DEA provides a scalar measure of the efficiency of the organizations being evaluated and provides their effective input and output levels [2].

B. DEA basic models

There are two basic DEA models which are named after their founders, CCR-Charnes, Cooper, and Rhodes and BBC-Banker, Charnes, and Cooper) models [4], [8], [10], [25]. The main feature by which the models differ is the set of feasible activities and frontiers [33]. The CCR model assumes constant returns to scale, meaning that output variables grow proportionally with input variables, while the BCC model assumes variable returns to scale, meaning that proportional changes in inputs do not necessarily imply proportional changes in outputs [10], [16], [21].

A number of extensions and modifications of the basic DEA models are presented in the literature [31]. Extensions of the DEA method in the context of multilevel, stochastic, and fuzzy DEA models are discussed in [20]. In [9] authors argued about desirable and undesirable outputs, while in [34] authors analysed variable returns to scale and non-discretionary inputs. A comparison of basic DEA models with extensions is

*Research supported by University North, Croatia.

A.M. Author is with the University North, Croatia (e-mail: amandic@ unin.hr).

K.F.Č. The author is with University North, Croatia. (e-mail: kcikovic@unin.hr).

D.K. Author is with the University North, Croatia (e-mail: dkecek@unin.hr).

analysed in [18]. An extension of deterministic to stochastic DEA is described in [24].

C. Most common applications of DEA methodology

There has been "an exponential growth in the number of publications related to theory and applications of Data Envelopment Analysis (DEA)" [14]. In [13] authors found that there was literally "exponential" growth in the number of publications between DEA's introduction in 1978 and 1995.

Some researchers have carried out different studies on bibliometric analysis on DEA. For instance, in [14] authors found that "energy, industry, banking, education and healthcare including hospital" are the most applied research areas of DEA. In [22] authors reveal that "agricultural production, dairy farms, industries, renewable energy systems, non-renewable energy systems, transportation systems, airlines" are the most common applications of DEA methodology.

There are other authors who focus solely on the applications of DEA in a certain industry. Namely, in [11] authors focused on DEA in banking; in [16] authors explored the DEA application in sustainable supplier selection process; in [15] authors focused on the DEA applications in multimedia; whereas the paper [23] investigated the DEA applications in forest engineering.

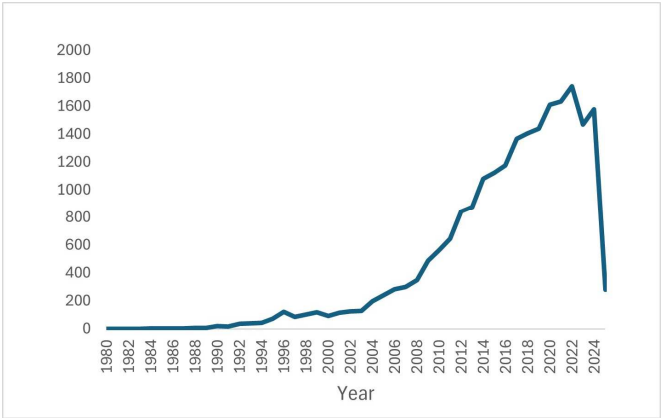
III. RESEARCH METHODOLOGY

This bibliometric study was conducted in the field of Data Envelopment Analysis (DEA), and all sources in the Web of Science database were eligible for inclusion. The data for this study was collected from the Web of Science (WoS) scientific database, with no restrictions on the language, date, or type of publications imposed. WoS from Thomson Reuters (ISI) was "the only citation database and publication which covers all domains of science for many years" and arose from the Science Citation Index in the 1960s. The main rationale for its survey was found in the notion that it has "strong coverage which goes back to 1990 and most of its journals written in English" [3]. Web of Science, which is known for its journal selection criteria, ensures that only high-quality, peer- reviewed papers are included. This database encompasses many disciplines, such as economics, sociology, and politics. The interdisciplinary approach is crucial for the context of an in-depth understanding of this topic and such an analysis.

The search term, i.e., keyword used, was "DATA ENVELOPMENT ANALYSIS". The main objective was to include all the published research thus far with the application of the DEA, and to perform several types of analyses, including collaboration networks, keyword analysis, trends over time, country, and affiliation-related analyses, etc.

Figure 1. Basic information regarding the survey. (Source: Authors work with Bibliometrix software)

Figure 2. Annual Scientific Production. (Source: Authors work with Bibliometrix software)



A datafile in BiB format was created totaling 21,894 papers, and this file was analysed using the Biblioshiny "web interfase of the Bibliometrix R package" [32] for the bibliometric analysis.

IV. RESULTS

This bibliometric study incorporates a total of 21,894 papers published in the period from 1980 to 2025 in the Web of Science database. Fig. 1 shows there are 28,709 authors to these publications, 388,305 used references and the annual growth rate is 13.37%. The average citations number per document was 24.7.

Focusing on the yearly research publication/growth, Fig. 2 illustrates the main trends regarding research interest in DEA applications. As shown in Fig. 2, the annual scientific production started increasing in 1980 (right after its introduction in 1978) and has grown steadily until 2004, when it experienced a rather steep rise until today. The application of DEA methodology is so wide, and the annual scientific production only confirms this notion. The scientific annual production reveals the "evolution of a particular activity in any field of research" [27], and in this case, in the DEA application.

Fig. 3 reveals the most relevant sources (i.e. journals) where the DEA-related studies were published. These are the European Journal of Operational Research (with 979 documents), followed by Sustainability Journal with 608 documents and Journal of Cleaner Production with 502 documents. This is consistent with the findings of the seminal bibliometric study in [14] and in [30].

Figure 3. Most relevant sources. (Source: Authors work with Bibliometrix software)

Figure 4. Most relevant authors. (Source: Authors work with Bibliometrix software)

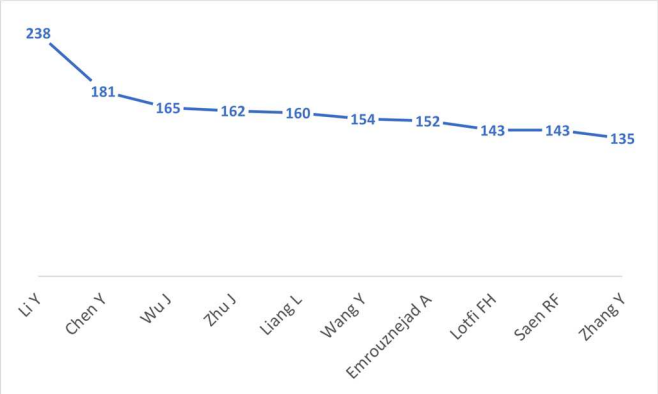


Fig. 4 shows the most relevant and contributing authors: Li, Y. (with 238 papers), Chen, Y. (with 181 papers), and Wu, J. (with 165 papers).

As for the most relevant documents and articles, logically, the paper of Banker et al. (1984), which introduces the BCC DEA model, is the most cited paper in the period 1980-2025, with a total of 9691 citations. Thereafter, Tone's (2001) paper has 3281 citations, and Andersen's (1993) paper has 2398 citations. The list of the most influential and cited works is presented in Fig. 5.

In Fig. 6, China appears to be the most cited country in the DEA-related papers, with 145,738 citations. It is followed by the USA (with 91,651 citations) and the UK (with 35,353 citations). This is partially in line with the study [30], where was found that authors from the USA and the UK are countries with the most researchers in the DEA publications.

Figure 5. Most global cited documents. (Source: Authors work with Bibliometrix software)

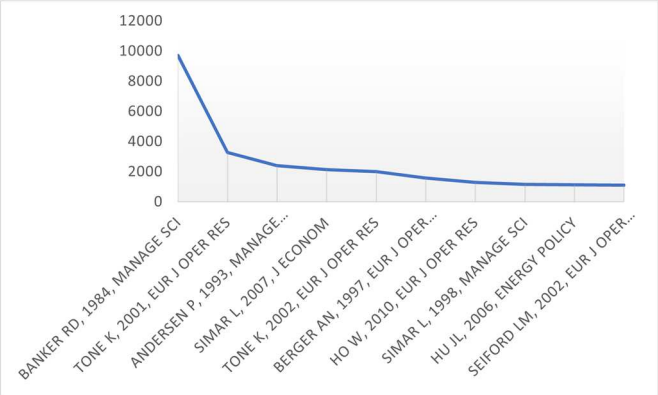


Figure 6. Most cited countries. (Source: Authors work with Bibliometrix software)

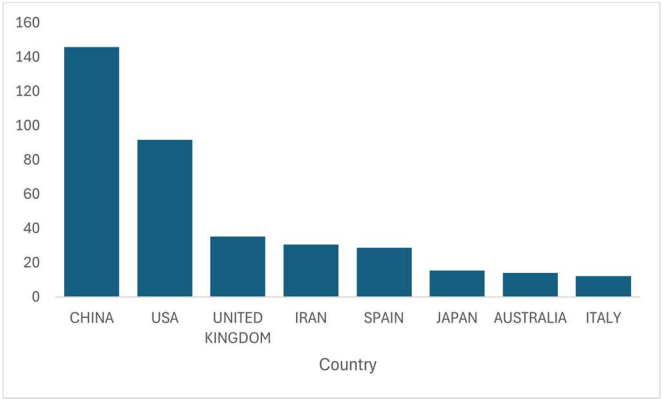
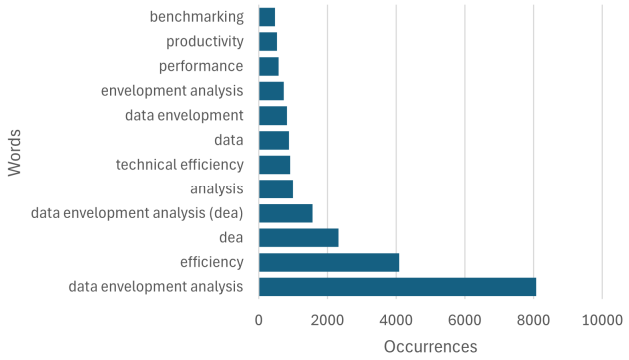


Figure 7. Most relevant words. (Source: Authors work with Bibliometrix software)



As presented in Fig. 7, the most relevant words used in the studies published ever since the introduction of DEA are “data envelopment analysis” with 8077 occurrences, “efficiency” with 4093 occurrences and “DEA” with 2328 occurrences. This is in line with the findings in [30].

In Fig. 8, the thematic map is presented. It reveals that the central bubble corresponds to "data envelopment analysis," associated with concepts like "efficiency" and "performance." Its position near the center indicates it has moderate centrality and development, highlighting its significance while suggesting potential for further research. This type of analysis is crucial for identifying research gaps and strategic opportunities within the scientific community, allowing researchers to focus on high-impact areas or explore niche topics with specialized insights.

Figure 8. Thematic map. (Source: Authors work with Bibliometrix software)

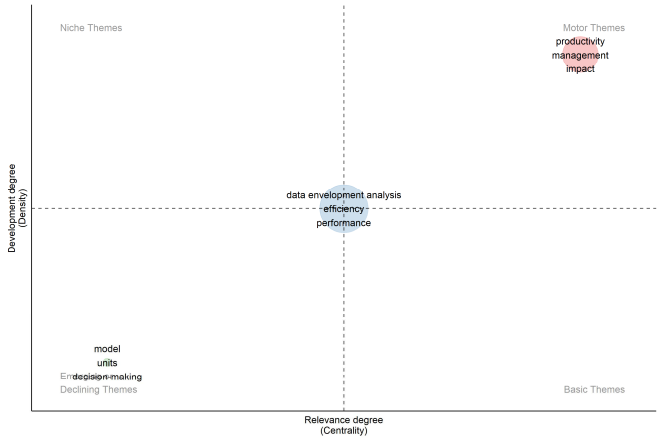
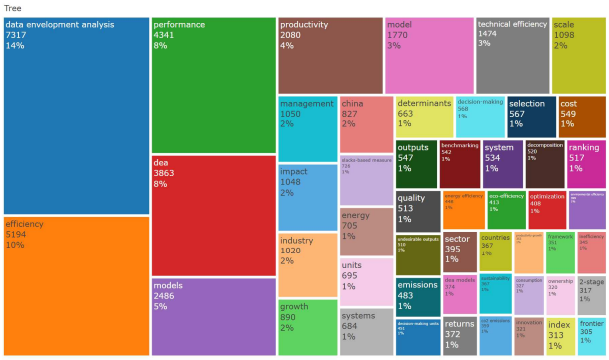


Figure 9. The treemap of DEA-related documents in the Web of Science database.



The tree map in Fig. 9 reveals rather interesting insights. Namely, it highlights the combination of possible keywords used in the DEA-related papers. Most of the papers in the WoS database use the terms "data envelopment analysis" (14%), “efficiency” (10%), “DEA” (10%), and "performance” (8%).

V. Conclusion

The main inspiration for this bibliometric study was the DEA methodology itself and its major popularity across different industries and research areas. The main objective of this bibliometric study is to provide a comprehensive bibliographic analysis of DEA-related studies ever since the introduction of DEA methodology in 1978. For this purpose, the Web of Science scientific database was surveyed, and the data analysis was conducted with the use of the R package of Bibliometrix and Biblioshiny.

This study provided a comprehensive survey and bibliometric analysis of scholarly literature in Data Envelopment Analysis (DEA), highlighting key trends, influential authors, prominent journals, and evolving research themes. The analysis revealed a steady growth in DEA-related publications since its inception in 1978, with significant contributions from both academic and applied research domains. The most prolific authors and institutions have played a critical role in shaping the field, and citation analysis helped identify foundational works that continue to influence current research.

The thematic evolution of DEA research indicates a transition from methodological advancements to more application-driven studies, particularly in healthcare, education, energy efficiency, and environmental assessment. The integration of DEA with other techniques such as machine learning, multi-criteria decision-making, and fuzzy logic reflects the field’s growing interdisciplinarity.

This bibliometric study incorporates a total of 21,894 papers published in the period from 1980 to 2025 in the Web of Science database. Moreover, there are 28,709 authors in these publications, 388,305 used references, and the annual growth rate is 13.37%. The average citation number per document was 24.7. The annual scientific production started increasing in 1980 (right after its introduction in 1978), and has grown steadily until 2004, when it experienced a rather steep rise until today. The application of DEA methodology is so wide, and the annual scientific production only confirms this notion. The most relevant sources (i.e., journals) where the DEA-related studies were published are the European Journal of Operational Research (with 979 documents), followed by the Sustainability Journal with 608 documents, and the Journal of Cleaner Production with 502 documents. The most relevant and contributing authors are Li, Y. (with 238 papers), Chen, Y. (with 181 papers), and Wu, J. (with 165 papers). As for the most relevant documents and articles, logically, the paper of Banker et al. (1984), where the BCC DEA model is introduced, is the most cited paper in the period 1980-2025, with a total of 9691 citations. Thereafter, Tone's (2001) paper has 3281 citations, and Andersen's (1993) paper has 2398 citations.

The most relevant words used in the studies published ever since the introduction of DEA are "data envelopment analysis" with 8077 occurrences, "efficiency" with 4093 occurrences, and "DEA" with 2328 occurrences. Most of the findings are consistent with the thus far bibliometric findings in [14] and in [30]. This bibliometric study reveals the often-overlooking sectoral potential in the context of DEA applications. While DEA is widely used in banking, healthcare, education, agriculture, and energy, there are several underutilized sectors where DEA could have significant impact such as creative industries, startups and innovation ecosystems, tourism and hospitality etc.

Considering the study only surveyed exclusively the Web of Science database, this is considered its significant methodological limitation. Namely, there may have been relevant studies published in other scientific databases, but were not included in this bibliometric study. This is planned to be overcome by future work of authors, which incorporates various scientific databases as an addition to the Web of Science (such as the Scopus and Google Scholar databases).

The findings provide a systematization of thus far gained knowledge in the DEA applications and represent a solid basis for future work with DEA applications. While DEA remains a robust and versatile tool for efficiency analysis, our review suggests areas for future research, including handling dynamic and network structures, incorporating big data and AI-driven methods, and enhancing decision-making under uncertainty.

This study offers many scientific and practical contributions. First of all, the scientific contributions could be summarized as follows: it provides a knowledge synthesis,

evolution and research trends of the DEA ever since its introduction in 1978; it highlights research gaps and it provides methodological rigor. Secondly, the practical contributions of the study revolve around the opportunity to inform, educate and inspire practitioners in regards to DEA; it represents a valuable tool for policy evaluation, benchmarking and public-sector efficiency analysis and it provides a basis for strategic planning for organizations.

Overall, this study not only maps the intellectual landscape of DEA but also provides a valuable reference for researchers seeking to understand its development and explore new frontiers in efficiency and productivity analysis.

REFERENCES

[1] A. Amirteimoori, B. K. Sahoo, V. Charles, and S. Mehdizadeh, "An Introduction to Data Envelopment Analysis," in *Stochastic Benchmarking. International Series in Operations Research & Management Science,* vol. 317. Springer, Cham, 2022. https://doi.org/10.1007/978-3-030-89869-4_2

[2] W. F. Bowlin, "Measuring Performance: An Introduction to Data Envelopment Analysis (DEA)", *The Journal of Cost Analysis*, vol. 15, pp. 3–27, 1998. https://doi.org/10.1080/08823871.1998.10462318

[3] A. A. Chadegani, H. Salehi, M. M. Yunus, H. Farhadi, M. Fooladi, M. Farhadi, and N. A. Ebrahim, "A comparison between two main academic literature collections: Web of Science and Scopus databases," *Asian Social Science*, vol. 9, 2013.

[4] A. Charnes, W. W. Cooper, and E. Rhodes, "Measuring the efficiency of decision-making units," *European Journal of Operational Research*, vol. 2, pp. 429-444, 1978. https://doi.org/10.1016/0377-2217(78)90138-8.

[5] A. Charnes, W. W. Cooper, and E. Rhodes, "A Data Envelopment Analysis Approach to Evaluation of the Program Follow-through Experiment in US Public School Education", Carnegie-Mellon Univ Pittsburgh Pa Management Sciences Research Group, 1978. https://apps.dtic.mil/sti/pdfs/ADA068710.pdf

[6] A. Charnes, and W. W. Cooper, "Auditing and accounting for program efficiency and management efficiency in not-for-profit entities," *Accounting, Organizations and Society*, vol. 5, pp. 87-107, 1980. https://doi.org/10.1016/0361-3682(80)90025-2.

[7] A. Charnes, W. W. Cooper, and E. Rhodes, "Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through," *Management Science*, vol. 27, pp. 668–697, 1981. http://www.jstor.org/stable/2631155

[8] A. Charnes, W. W. Cooper, A. Y. Lewin, and L. M. Seiford, "Basic DEA Models," in *Data Envelopment Analysis: Theory, Methodology, and Applications*, Springer, Dordrecht. https://doi.org/10.1007/978-94-011-0637-5_2

[9] L. Cherchye, B. De Rock, and B. Walheer, "Multi-output efficiency with good and bad outputs," *European Journal of Operational Research*, vol. 240, pp. 872-881, 2015. https://doi.org/10.1016/j.ejor.2014.07.028.

[10] W. W. Cooper, L. M. Seiford, and K. Tone, *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, 2nd ed.; Springer: New York, NY, USA, 2007.

[11] V. Cvetkoska, and G. Savic, "DEA in banking: Analysis and visualization of bibliometric data," *Data Envelopment Analysis Journal*, vol. 5, pp. 455-485, 2021.

[12] L. W. Day, and D. Belson, "Studying and incorporating efficiency into gastrointestinal endoscopy centers," *Gastroenterology Research and Practice*, vol. 2015, 764153, 2015.

[13] A. Emrouznejad, B. R. Parker, and G. Tavares, "Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA," *Socio-economic planning sciences*, vol. 42, pp. 151-157, 2008.

[14] A. Emrouznejad, and G-I. Yang, "A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016," *Socio-Economic Planning Sciences*, vol. 61, pp. 4-8, 2018. https://doi.org/10.1016/j.seps.2017.01.008.

[15] K. Fotova Čiković, D. Keček, and J. Lozić, "A PRISMA-compliant review of DEA applications in multimedia," *Tehnički vjesnik*, vol. 30, pp. 655-659, 2023.

[16] K. Fotova Čiković, I. Martinčević, and J. Lozić, "Application of Data Envelopment Analysis (DEA) in the Selection of Sustainable Suppliers: A Review and Bibliometric Analysis," *Sustainability*, vol. 14, 6672, 2022. https://doi.org/10.3390/su14116672

[17] H. Fukuyama, "Radial Efficiency Measures in Data Envelopment Analysis," in *Encyclopedia of Business Analytics and Optimization*, J. Wang, Ed. IGI Global Scientific Publishing, 2014, pp. 1967-1976 https://doi.org/10.4018/978-1-4666-5202-6.ch177

[18] S. Gupta, V. K. Yadav, S. Ghosh, and D. Saravanan, "Non-parametric Frontier Analysis Models for Relative Performance Evaluation," in: *Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering*, vol 661, M. N. Favorskaya, S. Mekhilef, R. K. Pandey, N. Singh, Eds. Singapore: Springer, 2021. https://doi.org/10.1007/978-981-15-4692-1_34

[19] J. Jablonský, A. Emrouznejad, and M. Toloo, "Editorial: Special issue on data envelopment analysis," *Central European Journal of Operations Research*, vol. 26, pp. 809-812, 2018.

[20] C. T. Kuah, K. Y. Wong, and F. Behrouzi, "A Review on Data Envelopment Analysis (DEA)," presented at the 2010 Fourth Asia International Conference on Mathematical/Analytical Modelling and Computer Simulation, Kota Kinabalu, Malaysia, pp. 168–173, doi: 10.1109/AMS.2010.45.

[21] M. Martić, M. Novaković, and A. Baggia, "Data Envelopment Analysis - Basic Models and their Utilization," *Organizacija*, vol. 42, pp. 37-43, 2009.

[22] Z. Mohtashami, M. Khedmati, and K. Eshghi, "Applications of Data Envelopment Analysis (DEA) for Optimizing Energy Consumptions," *Handbook of Smart Energy Systems*, pp. 1–41, 2022.

[23] O. F. Obi, L. Lebel, and F. Latterini, "A review of applications of data envelopment analysis in forest engineering," *Current Forestry Reports*, vol. 9, pp. 171-186, 2023.

[24] O. B. Olesen, and N. C. Petersen, "Stochastic Data Envelopment Analysis—A review," *European Journal of Operational Research*, vol. 251, pp. 2-21, 2016. https://doi.org/10.1016/j.ejor.2015.07.058.

[25] A. Panwar, M. Olfati, M. Pant, et al. "A Review on the 40 Years of Existence of Data Envelopment Analysis Models: Historic Development and Current Trends," *Arch Computat Methods Eng*, vol. 29, pp. 5397–5426, 2022. https://doi.org/10.1007/s11831-022-09770-3

[26] S. C. Ray, "Data Envelopment Analysis: An Overview," Working papers 2014-33, University of Connecticut, Department of Economics, 2014.

[27] K. Saini, M. Saini, A. Kumar, and D. K. Saini, "Performance analysis and optimization in renewable energy systems: a bibliometric review," *Discover Applied Sciences*, vol. 7, pp. 1-25, 2025.

[28] S. Samoilenko, "Overview of Data Envelopment Analysis," in *Advances in Research Methods for Information Systems Research*, vol 34, K. M. Osei-Bryson, O. Ngwenyama, O. Eds. Boston: Springer, 2014. https://doi.org/10.1007/978-1-4614-9463-8_11

[29] S. Secinaro, V. Brescia, D. Calandra, and P. Biancone, "Employing bibliometric analysis to identify suitable business models for electric cars," *Journal of cleaner production*, vol. 264, 121503, 2020.

[30] G. Tavares, "A bibliography of data envelopment analysis" (1978e2001). RUTCOR, Rutgers University; 2002. http://rutcor.rutgers.edu/ pub/rrr/reports2002/1_2002.pdf.

[31] E. Thanassoulis, "Extensions to Basic DEA Models," in *Introduction to the Theory and Application of Data Envelopment Analysis*, Boston: Springer, 2001. https://doi.org/10.1007/978-1-4615-1407-7_9

[32] E. M. Vătămănescu, G. Dominici, V. E. Ciuciuc, A. Vițelar, and F. G. Anghel, "Connecting smart mobility and car sharing using a systematic literature review. An outlook using Bibliometrix," *Journal of Cleaner Production*, 144333, 2024.

[33] Li, X. and G. Wang, "A Review Study on Two Basic Models of DEA," *Journal of Xi'an University of Arts and Science,* 2009, pp. 67-71.

[34] M. Xue, and P. T. Harker, "Extensions of Modified DEA" Center for Financial Institutions Working Papers 99-07, Wharton School Center for Financial Institutions, University of Pennsylvania, 1999.