Jurnal Teknik Informatika C.I.T Medicom, 14 (2) (2022) 01-04

Published by: Institute of Computer Science (IOCS)

Jurnal Teknik Informatika C.I.T Medicom

Journal homepage: www.medikom.iocspublisher.org

Integration of stochastic and robust optimization techniques into DEA model for more accurate and reliable efficiency estimation

Hengki Tamando Sihotang[1], Patricius Michaud Felix[2], Aisyah Alesha[3], Joan De Mathew[4]

[1)]Computer Science Doctoral Program, Universitas Sumatera Utara, Medan, Indonesia

[2)]Queen's University, Kingston, Canada

[3)]Institute of Computer Science (IOCS), Medan, Indonesia

[4)]Scuola Normale Superiore, Pisa PI, Italia

Article Info

Article history

Received : Feb 11, 2022
Revised : Mar 23, 2022
Accepted : Jun 26, 2022

Keywords:

Data Envelopment Analysis;
Efficiency estimation;
Optimization;
Stochastic and robust.

Abstract

Efficiency assessment is vital to assessing decision-making units (DMUs) in numerous sectors. DEA is a prominent non-parametric efficiency assessment tool. Traditional DEA models assume deterministic inputs and outputs, ignoring real-world uncertainties and variability. To improve efficiency estimation, stochastic and robust optimization approaches can be integrated into DEA models. To improve efficiency estimation, we present a stochastic and robust optimization framework incorporating DEA. Probabilistic inputs and outputs allow stochastic optimization to account for uncertainty. The model can capture data variability and create stochastic DMU efficiency scores by adding probability distributions. For data uncertainties and outliers, the DEA model uses robust optimization. Robust optimization considers worst-case scenarios and minimizes extreme observations on efficiency estimation. This makes efficiency scores more resilient to data outliers and noise. DEA models benefit from stochastic and resilient optimization. First, considering data uncertainties and fluctuations improves DMU efficiency representation. Second, eliminating outliers and extreme observations improves efficiency estimation. Third, efficiency scores help decision-makers make better, more informed choices. A case study in a specific industry shows the framework's efficacy. We compare classic and integrated stochastic-robust DEA model outcomes. The integrated model provides more accurate and dependable efficiency estimates, helping decision-makers understand DMU performance. DEA models with stochastic and resilient optimization increase efficiency estimation. By considering uncertainties and outliers, this paradigm helps decision-makers evaluate DMUs in many sectors more accurately and reliably.

Corresponding Author:

Hengki Tamando Sihotang,
Department of Operations Research and Informatics
Institute of Computer Science (IOCS)
Jl. Sei Mencirim Gg. Perkutut, Sei Mencirim, Kabupaten Deli Serdang, Sumatera Utara 20351, Indonesia
Email: hengkitamando@yahoo.com

1. Introduction

Data Envelopment Analysis (DEA) is a widely used methodology for measuring the efficiency of decision-making units (DMUs) in various industries [1], [2],[3]. DEA is a non-parametric technique that compares the input-output relationship of DMUs to generate efficiency scores, which can be used to identify best practices and benchmark performance [4],[5],[6]. DEA models often assume that the input and output data are deterministic and free of errors. In reality, the data are often uncertain, noisy, and incomplete, which can lead to biased and inefficient efficiency estimates [7],[8]. This limitation has led to the development of Robust Stochastic Data Envelopment Analysis (RSDEA), which integrates both stochastic and robust optimization techniques into DEA models [9]. Stochastic DEA allows for the consideration of uncertainties in input and output data, such as measurement errors, random fluctuations, and missing data [10],[11],[12],[13]. It uses probabilistic models to generate efficiency scores that reflect the likelihood of achieving optimal performance under different scenarios of uncertain data [14],[15].

Robust optimization, on the other hand, aims to provide solutions that are insensitive to uncertainties and deviations from the data's assumptions [16],[17]. It uses worst-case scenario analysis to generate efficient and robust solutions that are immune to any possible deviation from the assumed model. By integrating both stochastic and robust optimization techniques into DEA models, RSDEA can generate more accurate and reliable efficiency estimates that reflect the uncertainties and variability in the data [18]. RSDEA models can be used to evaluate the performance of DMUs in various industries, such as finance, healthcare, and energy, where the data are often uncertain and noisy. The application of RSDEA can provide decision-makers with more robust and accurate information, leading to more informed decision-making and improved performance.

RSDEA can be applied in different ways depending on the specific characteristics of the data and the problem at hand. One common approach is to use interval data, which defines a range of possible values for each input and output variable [19]. This allows for a more realistic representation of the uncertainties in the data, and RSDEA models can be used to generate efficient and robust solutions that are immune to any possible deviation from the assumed model. Another approach is to use a Monte Carlo simulation to generate multiple scenarios of uncertain data and apply a robust optimization approach to generate efficient and robust solutions that are insensitive to any possible deviations from the assumed model [20],[21]. RSDEA has been applied in various industries, such as healthcare, energy, finance, and education, to evaluate the efficiency and performance of DMUs under uncertain and variable conditions. For example, in healthcare, RSDEA has been used to evaluate the efficiency of hospitals and healthcare systems under different scenarios of uncertain data, such as patient volume, staffing levels, and medical supply availability. In finance, RSDEA has been applied to evaluate the efficiency of investment portfolios under different scenarios of uncertain data, such as stock prices, interest rates, and currency fluctuations. This can help investors and fund managers make more informed decisions and optimize their investment strategies.

The integration of stochastic and robust optimization techniques into DEA models can lead to more accurate and reliable efficiency estimates, making RSDEA a promising approach for decision-makers in various fields. The application of RSDEA can provide decision-makers with more robust and accurate information, leading to more informed decision-making and improved performance.
We address this need by conducting an SLR on the use of stochastic and robust optimization techniques into DEA models to generate efficiency estimates. SLRs can provide a valuable summary of the current knowledge in the research area and allow the identification of existing knowledge gaps and, consequently, avenues for future research.

This study contributes to the current literature on how stochastic and robust optimization techniques into DEA models can generate efficiency estimates by adding SLR to the prior in two ways. First, this study provides a thematically organized and up-to-date classification of previous studies with respect to their application areas, limitations, and recommendations. Second, based on the findings of the SLRs, we propose a synthesis framework to detail potential themes that require scientific attention to advance current knowledge.

The field of Robust Stochastic Data Envelopment Analysis (RSDEA) is relatively new and rapidly evolving, with ongoing research aimed at improving the methodology and its applications. Here are some potential themes that require scholarly attention to advance the current body of knowledge in RSDEA: Model development: There is still much room for innovation in developing new RSDEA models that can handle more complex and dynamic decision-making problems. For example, the development of models that can handle non-linear relationships between inputs and outputs, dynamic changes in the data, and multiple decision-makers or stakeholders.

Model selection: The choice of RSDEA model and its parameters can have a significant impact on the efficiency estimates and decision-making outcomes. Further research is needed to develop guidelines and best practices for selecting the most appropriate RSDEA model for specific decision-making problems. Data quality: The accuracy and reliability of the efficiency estimates depend on the quality of the input and output data. There is a need for further research on data pre-processing techniques, such as imputation, cleaning, and normalization, to improve the quality of the data and reduce the impact of outliers and missing data.

Interpretability: RSDEA models generate complex and often high-dimensional results, making it challenging to interpret and communicate the efficiency estimates to decision-makers. There is a need for further research on visualization and communication techniques to improve the interpretability of RSDEA results and support informed decision-making. Application areas: While RSDEA has been applied in various industries, such as healthcare, finance, and energy, there is a need for further research on its applicability in other fields, such as environmental sustainability, social welfare, and public policy. Additionally, there is a need for comparative studies that evaluate the effectiveness and efficiency of RSDEA compared to other decision-making tools and techniques. Integration with other decision-making tools: RSDEA is often used in combination with other decision-making tools and techniques, such as Multi-Criteria Decision Analysis (MCDA), to provide more comprehensive and robust decision-making support. Further research is needed to explore the synergies and complementarities between RSDEA and other decision-making tools and techniques and develop integrated decision-making frameworks. These themes represent potential areas for future research that can contribute to advancing the current body of knowledge in RSDEA and improving its applications in real-world decision-making problems.

Announcements

The topic of this article is being carried out research on Computer Science Doctoral dissertation research at the University of North Sumatra, by Hengki Tamando Sihotang who will propose a new method, namely RSDEA (Robust Stochastic Data Envelopment Analysis).

References

[1] M. Izadikhah and R. Farzipoor Saen, "A new data envelopment analysis method for ranking decision making units: an application in industrial parks," *Expert Syst.*, vol. 32, no. 5, pp. 596–608, 2015.

[2] M. Farahmand, M. I. Desa, and M. Nilashi, "A combined data envelopment analysis and support vector regression for efficiency evaluation of large decision making units," *Int. J. Eng. Technol.*, pp. 2310–2321, 2014.

[3] S. V. Avilés-Sacoto, W. D. Cook, D. Güemes-Castorena, and J. Zhu, "Modelling efficiency in regional innovation systems: A two-stage data envelopment analysis problem with shared outputs within groups of decision-making units," *Eur. J. Oper. Res.*, vol. 287, no. 2, pp. 572–582, 2020.

[4] J. Stroobants and G. Bouckaert, "Benchmarking local public libraries using non-parametric frontier methods: A case study of Flanders," *Libr. Inf. Sci. Res.*, vol. 36, no. 3–4, pp. 211–224, 2014.

[5] A. Khoshroo, R. Mulwa, A. Emrouznejad, and B. Arabi, "A non-parametric Data Envelopment Analysis approach for improving energy efficiency of grape production," *Energy*, vol. 63, pp. 189–194, 2013.

[6] X. Dai and T. Kuosmanen, "Best-practice benchmarking using clustering methods: Application to energy regulation," *Omega*, vol. 42, no. 1, pp. 179–188, 2014.

[7] S. J. Sadjadi and H. Omrani, "A bootstrapped robust data envelopment analysis model for efficiency estimating of telecommunication companies in Iran," *Telecomm. Policy*, vol. 34, no. 4, pp. 221–232, 2010.

[8] T. Kuosmanen and M. Kortelainen, "Stochastic non-smooth envelopment of data: semi-parametric frontier estimation subject to shape constraints," *J. Product. Anal.*, vol. 38, pp. 11–28, 2012.

[9] S. J. Sadjadi, H. Omrani, S. Abdollahzadeh, M. Alinaghian, and H. Mohammadi, "A robust super-efficiency data envelopment analysis model for ranking of provincial gas companies in Iran," *Expert Syst. Appl.*, vol. 38, no. 9, pp. 10875–10881, 2011.

[10] B. E. El-Demerdash, I. A. El-Khodary, and A. A. Tharwat, "A stochastic data envelopment analysis model considering variation in input and output variables," *Int. J. Data Envel. Anal. Oper. Res.*, vol. 2, no. 1, pp. 1–6, 2016.

[11] A. Desai, S. J. Ratick, and A. P. Schinnar, "Data envelopment analysis with stochastic variations in data," *Socioecon. Plann. Sci.*, vol. 39, no. 2, pp. 147–164, 2005.

[12] M. Asgharian, M. Khodabakhshi, and L. Neralic, "Congestion in stochastic data envelopment analysis: An input relaxation approach," *Int. J. Stat. Manag. Syst.*, vol. 5, no. 1, pp. 84–106, 2010.

[13] R. Eslami, M. Khodabakhshi, G. R. Jahanshahloo, F. H. Lotfi, and M. Khoveyni, "Estimating most productive scale size with imprecise-chance constrained input–output orientation model in data envelopment analysis," *Comput. Ind. Eng.*, vol. 63, no. 1, pp. 254–261, 2012.

[14] C. Sutton, K. Rohanimanesh, and A. McCallum, "Dynamic conditional random fields: Factorized probabilistic models for labeling and segmenting sequence data," in *Proceedings of the twenty-first international conference on Machine learning*, 2004, p. 99.

[15] L. Getoor, N. Friedman, D. Koller, and A. Pfeffer, "Learning probabilistic relational models," *Relational data Min.*, pp. 307–335, 2001.

[16] A. Parisio, C. Del Vecchio, and A. Vaccaro, "A robust optimization approach to energy hub management," *Int. J. Electr. Power Energy Syst.*, vol. 42, no. 1, pp. 98–104, 2012.

[17] I. Sungur, F. Ordónez, and M. Dessouky, "A robust optimization approach for the capacitated vehicle routing problem with demand uncertainty," *Iie Trans.*, vol. 40, no. 5, pp. 509–523, 2008.

[18] M. Tavana, K. Khalili-Damghani, F. J. S. Arteaga, and A. Hosseini, "A fuzzy multi-objective multi-period network DEA model for efficiency measurement in oil refineries," *Comput. Ind. Eng.*, vol. 135, pp. 143–155, 2019.

[19] R. L. Iman, J. C. Helton, and J. E. Campbell, "An approach to sensitivity analysis of computer models: Part I—Introduction, input variable selection and preliminary variable assessment," *J. Qual. Technol.*, vol. 13, no. 3, pp. 174–183, 1981.

[20] C. Bohle, S. Maturana, and J. Vera, "A robust optimization approach to wine grape harvesting scheduling," *Eur. J. Oper. Res.*, vol. 200, no. 1, pp. 245–252, 2010.

[21] H.-G. Beyer and B. Sendhoff, "Robust optimization–a comprehensive survey," *Comput. Methods Appl. Mech. Eng.*, vol. 196, no. 33–34, pp. 3190–3218, 2007.