2020 International Conference on Contemporary Computing and Applications (IC3A)
Dr. A.P.J. Abdul Kalam Technical University, Lucknow. Feb 05-07, 2020

Combined Shannon's Entropy and Data Envelopment Analysis Model: Methods and Applications

2020 International Conference on Contemporary Computing and Applications (IC3A) 978-1-7281-5432-9/20/$31.00 ©2020 IEEE 10.1109/IC3A48958.2020.233310

Shubham Gupta
Department of Electrical Engineering
IIT(ISM)
Dhanbad, India
shubham.gupta6691@gmail.com

Tanushree
Department of Computer Science Engineering
Galgotias College of Engineering & Technology
Greater Noida, India
tanushree2k7@gmail.com

Amit Rai
Department of Electrical Engineering
Galgotias College of Engineering & Technology
Greater Noida, India
amit.rai@galgotiacollege.edu

Abstract –**The Data Envelopment Analysis (DEA) is a non-parametric mathematical tool for evaluating the relative performance of homogenous (i.e. similar inputs and outputs) decision-making units (DMUs). Since the introduction in 1978, countless DEA models have been suggested in view of its popularity worldwide. The common issues encountered by researchers in DEA models are: (1) traditional models discriminate poorly among DMUs; (2) variable nature of inputs/outputs weight; (3) subjectivity in selection of models; (4) possibility of alternative optima in cross-evaluation model. Many methodologies have been proposed to face such issues. In this paper, some of DEA models which utilize the concept of Shannon's entropy are presented, those can resolve aforesaid problems. In the end, an example illustrates the validity and effectiveness of the presented methodologies.**

Keywords – Data Envelopment Analysis, Shannon's entropy, relative efficiency.

I. INTRODUCTION

Data Envelopment Analysis (DEA) is a powerful mathematical tool for assessing the performance of a set of organizational entities termed as decision making units (DMUs). The DEA is a technique based upon the linear programming (LP) formulation that utilize the multiple inputs and multiple outputs (similar for each DMU under evaluation) for assessment [1]. The application of DEA tool have been widely used in organization related to power [2], education [3], banking [4], healthcare [5], agribusiness [6], supply chain management [7] and many more and to provide them the best ranking framework, countless DEA models have been suggested.

In today's competitive environment, it is essential to assess the performances of peer entities (*i.e.* DMUs) to rank them in terms of efficiency. In doing so, the researchers face issues like the classical DEA model often measures more than one DMU as efficient DMU and unable to discriminate further to compare or rank [8]. Furthermore, the flexible nature of multiplier weights of inputs/outputs allows DMU to maximize its efficiency score, results poor ranking evaluation as the inefficient unit might place on efficiency frontier [9]. Thereafter, each DEA model has its pros and cons, so issues arise on the subjectivity in DEA model selection [10]. The major setback in cross evaluation model of DEA is the non-uniqueness of the multiplier weights leads to alternate optimal solution for ranking [11].

In this paper, four combined Shannon's entropy based DEA models are presented to evaluate the relative efficiency of DMUs in order to find the best practices. These methodologies of DEA are able to handle abovementioned problems to a great extent. The rest of this paper is organized as follow: section 2 discusses the literature review related to work. Section 3 presents the mathematical modeling of combined Shannon's entropy and DEA models. Illustrative example and discussion have been reported in section 4 followed by the conclusion and remarks in section 5.

II. LITERATURE REVIEW

First DEA model (CCR) was proposed by Charnes et al. [1] in 1978 and is based on the concept of productivity and efficiency [12]. Banker et al. [13] introduced BCC model which assumes the variable return to scale (VRS). Both CCR and BCC are radial models means; either input oriented or output oriented. A non-radial model like Additive model [14] reduces input and increases output simultaneously. Additive model also named as Pareto-Koopmans model. The major setbacks in traditional models are poor discrimination in efficient units because of the self-evaluation scheme. The extension tool for classical DEA models was proposed by Sexton et al. [15], known as cross efficiency model. The main advantage of cross efficiency technique is peer-evaluation. Li and Reeves [16] introduced the Multi-criteria DEA (MCDEA) model in order to remove discriminatory capability of DEA by combining the multiple objectives in a single DEA model.

It is a well-known fact that classical DEA models are unable to compare or rank DMUs because of flexible nature of multiplier weights and poor discrimination power. To the best of our knowledge, common weight technique was firstly introduced by Roll et al. [17]. Later on, many methodologies have been proposed to calculate common set of weights, refer [8,18,19]. To accumulate the different viewpoints and models of DEA, Soleimani-damaneh et al. [10] used the key concept of Shannon's entropy and formulated efficiency scores. Thereafter, many researchers utilized the application of Shannon's entropy in the field of DEA. For example, refer to Wu et al. [11,20], Qi et al. [19], Xie et al. [21], Adhikari et al. [22] and Ghosh et al. [23,24].

III. METHODOLOGY

A. Shannon's entropy

Information entropy (called Shannon) was introduced by Claude E. Shannon [25] and originally used in the field of information theory to quantify the degree of uncertainty in a random process. Later, its application has been wide used in other field such as economics, management, engineering and so on. First model of DEA that combined the concept of Shannon entropy proposed by Soleimani-damaneh *et al.* in his paper [10]. Thereafter, many studies utilize the theory in formulation of different DEA methods, refer [11,19,21],. To compute efficiency scores of the DMUs, entropy is evaluated by following steps given below:

Step 1. Normalize the data as

$$\propto_{ij}=\frac{\beta_{ij}}{\sum_{i=1}^{n}\beta_{ij}} \quad ,j=1,\dots,J$$

Step 2. Calculate Shannon's entropy e_j as

$$e_j=-\frac{1}{\ln n}\sum_{i=1}^{n}\propto_{ij}.\ln\propto_{ij} \quad ,j=1,\dots,J$$

Step 3. Compute the degree of importance

$$\omega_j=\frac{1-e_j}{J-\sum_{j=1}^{J}e_j}$$

In step 3, the degree of importance is computed to describe the important of the criterion. If, criterion have smaller the value of e_j viz. larger will be the value of ω_j, means it will provide more information and will be the more important in efficiency evaluation.

B. Shannon's entropy based DEA models

In this paper, four methodologies are presented those combine the concept of Shannon's entropy and DEA to deliver a compromise solution for decision makers (DM).

a) Combined efficiency score

This method is proposed by [10], where Shannon's entropy is used to combine the efficiency score obtained through different conventional models of DEA. The idea of the methodology is to remove the subjectivity on the choice of model and to provide more realistic results.

Algorithm 1: Combined efficiency score
1. Select input and output data for $DMU_i(i=1,\dots,I)$ as $X_m(m=1,\dots,M)$ and $Y_r(r=1,\dots,R)$
2. Compute relative efficiency score through different DEA models $\{l_1,l_2,\dots,l_j\}$
3. Construct an efficiency matrix $[\beta_{ij}]$ (ex. - [E_{CCR}, E_{BCC}, $E_{Additive}$, …])
4. Calculate the degree of importance of each model (refer section 3.1)
5. Evaluate combined efficiency score as $\theta_i=\sum_{j=1}^{J}\omega_j\beta_{ij}\ ,i=1,\dots,I$

Algorithm 1 summarizes the procedure for computation of combined efficiency score. Selection of DEA models to construct efficiency matrix can vary (it depends on DM).

b) Ultimate cross efficiency

Cross efficiency model is basically an extension tool of DEA. It has several advantages such as remove unrealistic weight scheme; provide unique order of DMUs and many more. But, the main disadvantage is the presence of alternate optimal solution because of the non-uniqueness of the input/output weights obtained from original DEA model. Wu *et al.* [11] utilized the Shannon's entropy to define the multiplier weights in order to obtain ultimate cross efficiency score. The methodology is summed up in algorithm 2.

Algorithm 2: Ultimate cross efficiency
1. Select input and output data for $DMU_i(i=1,\dots,I)$ as $X_m(m=1,\dots,M)$ and $Y_r(r=1,\dots,R)$
2. Obtain multiplier input/output weights (μ_{mi} *and* π_{ri})
3. Evaluate cross efficiency matrix $[E]_{ii}$, refer [26]
4. Find transpose of cross efficiency matrix $[\mathbb{E}]=[E]^T$
5. Determine the closeness as $d_{ii}=\frac{\mathbb{E}_{ii}}{\max\{\mathbb{E}_{i1},\mathbb{E}_{i2},\dots,\mathbb{E}_{iI}\}}$, where $d_{ii}\in[0,1]$
6. Consider $[\beta_{ij}]=[d_{ii}]$ to compute Shannon entropy
7. Calculate the degree of importance (here weights) for each criterion (refer section 3.1)
8. Evaluate combined efficiency score as $\theta_i=\sum_{j=1}^{J}\omega_j\beta_{ij}\ ,i=1,\dots,I$

Ultimate cross efficiency method actually eliminates the assumption of averaging the efficiency scores attained from peer-evaluation and self-evaluation of a DMU.

c) Common weights DEA model using Shannon's entropy

In conventional DEA models, flexibility in multiplier weights of input/output allows DMU to attain its maximum efficiency. Because of it, typically more than one DMU is evaluated as efficient which cannot be further classified for ranking purpose. The common weight DEA techniques are effective in order to solve such problems. In this paper, Shannon's entropy is used to generate common set of weights for assessment. This method is proposed by [19] and steps given in algorithm 3 describe the methodology.

Algorithm 3: Common weights DEA model using Shannon's entropy
1. Select input and output data for $DMU_i(i=1,\dots,I)$ as $X_m(m=1,\dots,M)$ and $Y_r(r=1,\dots,R)$
2. Normalize the input and output data as $X_{mi_0}=\frac{X_{mi_0}}{\max X_{mi}}$ *and* $Y_{ri_0}=\frac{Y_{ri_0}}{\max Y_{ri}}$, $i=1,\dots,I$
3. Compute allowable maximum weight restriction ε_0 (refer eq. 7 in [19])
4. Using ε_0 in CCR model determine non-zero optimal weights $[V]and[U]$
5. Consider $[\beta_{ij}]=[V]$ *and* $[\beta_{ij}]=[U]$ to calculate Shannon's entropy (refer section 3.1)
6. Evaluate degree of importance by considering input and output together (refer eq. 17 in [19])
7. Determine the common set of weights for all DMU_iusing

$\theta_i = \sum_{j=1}^{J} \omega_j \beta_{ij}\ , i = 1, \dots, I$

8. Calculate the efficiency score as $eff. = \text{sum of weighted outputs} / \text{sum of weighted inputs}$

$[V]$ *and* $[U]$ are the non-zero optimal weights for input and output respectively.

d) Comprehensive efficiency score

Xie *et al.* proposed a methodology in [21] to enhance the discriminatory power of conventional DEA model. For this, DEA efficiencies are first computed for all different combination of variable (X_m *and* Y_r) subsets and then Shannon's entropy is applied to quantify the importance of each subset in efficiency assessment. In last, comprehensive efficiency score is obtained which improvises the outcomes compare to the conventional model.

Algorithm 4: Comprehensive efficiency score

---

1. Select input and output data for $DMU_i (i = 1, \dots, I)$ as $X_m (m = 1, \dots, M)$ and $Y_r (r = 1, \dots, R)$
2. Generate all different combination of variable subsets (having at least one input and one output)
3. Compute DEA efficiencies for all combinations using any one conventional model (ex. - CRR, BCC…)
4. Construct efficiency matrix as$[\mathbb{E}] = \{E_{c1}, E_{c2}, \dots, E_{ck}\}$
5. Consider $[\beta_{ij}] = [\mathbb{E}]$ to calculate Shannon's entropy
6. Calculate the degree of importance of each variable subset (refer section 3.1)
7. Evaluate comprehensive efficiency score as $\theta_i = \sum_{j=1}^{J} \omega_j \beta_{ij}\ , i = 1, \dots, I$

In algorithm 4, the total count for all probable subsets is $k = (2^m - 1) \times (2^r - 1)$ and E_{ck} refers to efficiency of k^{th} combination.

IV. ILLUSTRATION

In this section, a numerical example (shown in the Table I) is considered to illustrate the above presented methodology. There are 12 DMUs, where each DMU has 2 inputs X_1 and X_2 and 4 outputs Y_1, Y_2, Y_3 and Y_4. The key characteristics of data are given in last five row of Table I. The efficiency scores obtained for all four models have been reported in Table II. For convenience, efficiency scores are defined as Eff._{M1}, Eff._{M2}, Eff._{M3} and Eff._{M4} for combined efficiency score, ultimate cross efficiency, common weight model and comprehensive efficiency score, respectively. In the later stages, the superiority of each model is discussed in detail.

In order to obtain Eff._{M1}, the efficiency scores are evaluated from different DEA models; here CCR, BCC, Additive, Multi-criteria DEA (MCDEA) models are considered to build the decision making matrix ($[\beta_{ij}]$). The radial graph in Fig 1 represents the results simulated from these models. The column 2 of Table II reports the Eff._{M1}, where DMU 1 and 5 are found as the frontier for remaining DMUs. However, DMU 9 is measured as relatively efficient from first three models {CCR, BCC and Additive}, but Eff._{M1} model computed DMU 9 as least efficient DMU. This is because of the both models of MCDEA found DMU 9 as poorest performer in all. The importance degree of MCDEA (D_j and Min M) models are 0.2648 and 0.5845 respectively.

Furthermore, column 3 of Table II shows the Eff._{M2} which is calculated from the application of Shannon's entropy, applied on cross efficiency matrix. It is observed that, DMU 1 is ranked 1st among all units. On the other side, DMU 9 is in the last of ranking chart. However, average cross efficiency model predicts DMU 2 as most inefficient unit, with score 0.6074. In opposite, Eff._{M2} model find DMU2 as second most efficient unit among all, with score 0.9604. The Eff._{M2} model provides better simulations for ranking purpose than average cross efficiency score model, as seen in graph represented in Fig 2.

TABLE I. INPUT AND OUTPUT DATA FOR 12 DECISION MAKING UNITS FROM [27]

DMUs	Inputs		Outputs			
	X_1	X_2	Y_1	Y_2	Y_3	Y_4
1	17.02	5	42	45.3	14.2	30.1
2	16.46	4.5	39	40.1	13	29.8
3	11.76	6	26	39.6	13.8	24.5
4	10.52	4	22	36	11.3	25
5	9.5	3.8	21	34.2	12	20.4
6	4.79	5.4	10	20.1	5	16.5
7	6.21	6.2	14	26.5	7	19.7
8	11.12	6	25	35.9	9	24.7
9	3.67	8	4	17.4	0.1	18.1
10	8.93	7	16	34.3	6.5	20.6
11	17.74	7.1	43	45.6	14	31.1
12	14.85	6.2	27	38.7	13.8	25.4
Mean	11.05	5.77	24.08	34.48	9.98	23.83
Median	10.82	6.00	23.50	35.95	11.65	24.60
Standard deviation	4.76	1.28	12.42	8.96	4.50	4.85
Minimum	3.67	3.8	4	17.4	0.1	16.5
Maximum	17.74	8	43	45.6	14.2	31.1

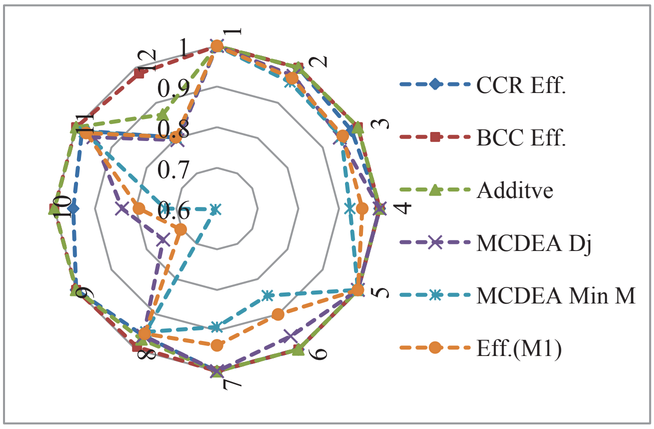

Fig. 1. Efficiency scores evaluated through different DEA model and Eff.M1

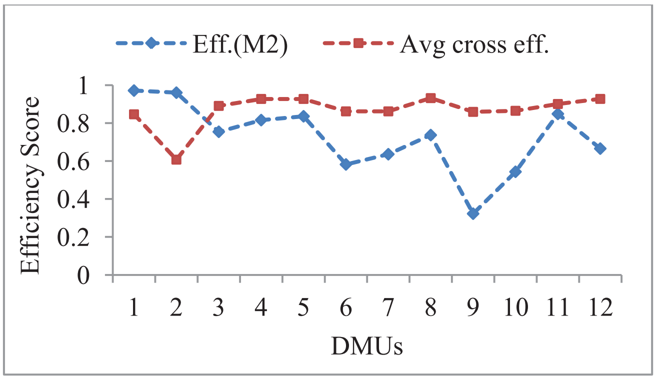

Fig. 2. Eff.M2 vs. average cross efficiency

TABLE II. EFFICIENCY SCORES FROM ALL FOUR MODELS

DMUs	$Eff._{M1}$	$Eff._{M2}$	$Eff._{M3}$	$Eff._{M4}$
1	1	0.9714	0.854495	0.9099
2	0.97	0.9604	0.832459	0.9171
3	0.9571	0.7543	0.875217	0.7346
4	0.9572	0.8159	0.943592	0.8882
5	1	0.8358	0.959064	0.9208
6	0.9007	0.5821	0.83208	0.5124
7	0.9365	0.6359	0.873315	0.5535
8	0.9554	0.7368	0.827073	0.6196
9	0.7031	0.3226	0.592303	0.2913
10	0.7919	0.5435	0.741181	0.4931
11	0.9716	0.8486	0.79118	0.6848
12	0.8025	0.6656	0.733868	0.6725

Moreover, the $Eff._{M3}$ is evaluated using common weights which are derived from non-zero optimal weights. The common weights are $V_1 = 11.02461, V_2 = 3.395809, U_1 = 3.327106, U_2 = 3.228551, U_3 = 1.681605$ and $U_4 = 3.633787$. The $Eff._{M3}$ results are provided in column 4 of Table II. The Shannon's entropy value of non-zero input/output weights, combined (input and output) degree of importance and $Eff._{M3}$ are graphically represented in Fig 3. DMU 9 is depicted as the least efficient unit and DMU 5 is scored maximum efficiency (0.959064) comparative to others. DMU 1 which is ranked 1st in first two model (M1 and M2) achieve score 0.854495 in M3. Surprisingly, DMU4 become the runner-up in leader board with a score of 0.943592 in M3.

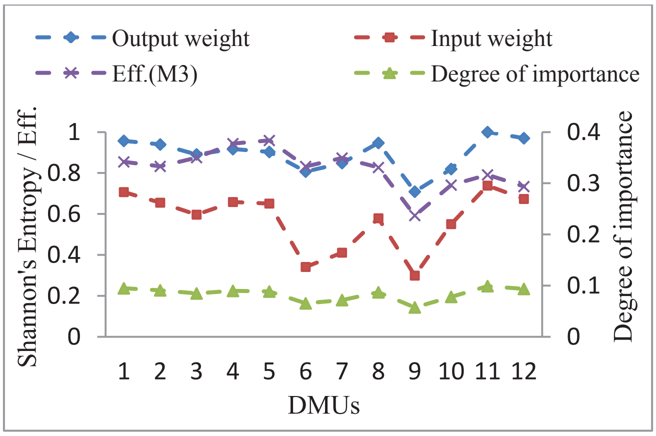

Fig. 3. Shannon's Entropy value for input/output weights, Eff.M3 and degree of importance

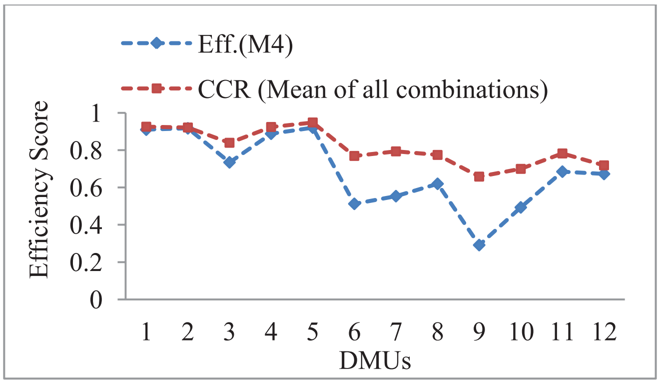

Fig. 4. EffM4 and mean CCR value for all combinations

Fig 4 portrayed the graphical description of $Eff._{M4}$ and mean of efficiencies ($CCR_{Eff.}$) of DMUs for all possible subsets. For the given data, the total number of different combinations is $k = (2^2 - 1) \times (2^4 - 1) = 45$. The scores of $Eff._{M4}$ are reported in column 5 of Table II. It is observed that, CCR model produces 7 out of 12 DMUs as the relatively efficient (i.e. score =1), but Eff_{M4} model enhances the discrimination power of CCR model and simulates DMU 5 as the only most efficient unit. DMU 2 and 1 occupy the rank 2nd and 3rd respectively with the efficiency score of 0.9171 and 0.9099 respectively. DMU 9 achieve score 0.2913 which is lowest in table.

The simulation results for above all four models have been carried out on Window 7 (64-bit) PC having Core i5, 2.66 GHz Processor, 4 GB RAM and 500 GB Hard disk Drive.

V. CONCLUSION

This paper presents four different combined Shannon's entropy and DEA approaches which helps in improving the efficacy of DEA. The major remarks that can be pointed out are as given below:

- Combined efficiency score model (M1) evaluate the degree of importance for each model considered in order to generate decision making matrix. This model removes the subjectivity in DEA model selection.
- Ultimate cross efficiency model (M2) eradicates the assumption of average and use Shannon's entropy to determine the $Eff._{M2}$.
- In common weight technique (M3), common weights for all participating DMUs are evaluated using Shannon's entropy. The methodology promises to be more effective in discriminating DMUs.
- The efficiencies computed from different specification of variable subsets are aggregated by utilizing the principle of Shannon's entropy to obtain comprehensive score ($Eff._{M4}$).

Overall, the presented models {M1, M2, M3 and M4} are able to handle the multifaceted problem to a great extent.

REFERENCES

[1] A. Charnes, W. W. Cooper and and E. Rhodes, "Measuring the efficiency of decision making units," *Eur. J. Oper. Res.*, vol. 2, pp. 429–444, 1978.

[2] V. Kumar, N. Kumar, S. Ghosh, and K. Singh, "Indian thermal power plant challenges and remedies via application of modified

data envelopment analysis," *Int. Trans. Oper. Res.*, vol. 00, pp. 1–23, 2014.
[3] J. Nazarko and J. Šaparauskas, "Application of DEA method in efficiency evaluation of public higher education institutions," *Technol. Econ. Dev. Econ.*, vol. 20, no. 1, pp. 25–44, 2014.
[4] A. E. Laplante and J. C. Paradi, "Evaluation of Bank Branch Growth Potential Using Data Envelopment Analysis," *Omega, http://dx.doi.org/10.1016/j.omega.2014.10.009*, 2014.
[5] S. Kohl, J. Schoenfelder, A. Fügener, and J. O. Brunner, "The use of Data Envelopment Analysis (DEA) in healthcare with a focus on hospitals," *Health Care Manag. Sci.*, 2018.
[6] E. Toma, C. Dobre, I. Dona, and E. Cofas, "DEA applicability in assessment of agriculture efficiency on areas with similar geographically patterns," *Agric. Agric. Sci. Procedia*, vol. 6, pp. 704–711, 2015.
[7] S. Soheilirad, K. Govindan, A. Mardani, E. K. Zavadskas, M. Nilashi, and N. Zakuan, "Application of data envelopment analysis models in supply chain management : a systematic review and meta-analysis," *Ann. Oper. Res.*, 2017.
[8] Y. Wang, Y. Luo, and L. Liang, "Ranking decision making units by imposing a minimum weight restriction in the data envelopment analysis," *J. Comput. Appl. Math.*, vol. 223, no. 1, pp. 469–484, 2009.
[9] K. Khalili-damghani and M. Fadaei, "A comprehensive common weights data envelopment analysis model : Ideal and anti-ideal virtual decision making units approach," *J. Ind. Syst. Eng.*, vol. 11, no. 3, pp. 281–306, 2018.
[10] M. Soleimani-damaneh and M. Zarepisheh, "Shannon's entropy for combining the efficiency results of different DEA models : Method and application," *Expert Syst. Appl.*, vol. 36, no. 3, pp. 5146–5150, 2009.
[11] J. Wu, J. Sun, L. Liang, and Y. Zha, "Determination of weights for ultimate cross efficiency using Shannon entropy," *Expert Syst. Appl.*, vol. 38, pp. 5162–5165, 2011.
[12] M. J. Farrell, "The Measurement of Productive Efficiency," *J. R. Stat. Soc.*, vol. 120, no. 3, pp. 253–290, 1957.
[13] R. D. Banker, A. Charnes, and W. W. Cooper, "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis," *Manage. Sci.*, vol. 30, no. 9, pp. 1078–1092, 1984.
[14] A. Charnes, W. W. Cooper, B. Golany, L. Seiford, and J. Stutz, "Foundations of data envelopment analysis for Pareto-Koopmans efficient empirical production functions," *J. Econom.*, vol. 30, no. 1–2, pp. 91–107, 1985.
[15] T. R. Sexton, R. H. Silkman, and A. J. Hogan, "Data Envelopment Analysis : Critique and Extensions," in *R. H. Silkman (ed.). Measuring Efficiency: An Assesrment of Data Envelopment Analysis. New Directions for Program Evaluation, San Francisco*, no. 32, 1986, pp. 73–105.
[16] X. B. Li and G. R. Reeves, "Multiple criteria approach to data envelopment analysis," *Eur. J. Oper. Res.*, vol. 115, no. 3, pp. 507–517, 1999.
[17] Y. Roll, W. D. Cook, and B. Golany, "Controlling Factor Weights in Data Envelopment Analysis," *IIE Trans.*, vol. 23, no. 1, pp. 2–9, 1991.
[18] C. Kao and H. Hung, "Data envelopment analysis with common weights : the compromise solution approach," *J. Oper. Res. Soc.*, vol. 56, pp. 1196–1203, 2005.
[19] X.-G. Qi and B. Guo, "Determining Common Weights in Data Envelopment Analysis with Shannon's Entropy," *Entropy*, vol. 16, pp. 6394–6414, 2014.
[20] J. Wu, J. Sun, and L. Liang, "DEA cross-efficiency aggregation method based upon Shannon entropy," *Int. J. Prod. Res.*, vol. 50, no. 23, pp. 6726–6736, 2012.
[21] Q. Xie, Q. Dai, Y. Li, and A. Jiang, "Increasing the Discriminatory Power of DEA Using Shannon's Entropy," *Entropy*, vol. 16, pp. 1571–1585, 2014.
[22] A. Adhikari, S. Basu, I. Biswas, and A. Banerjee, "A route efficiency analysis using Shannon entropy-based modified DEA method and route characteristics investigation for urban bus transport in India," *Inf. Syst. Oper. Res.*, pp. 1–28, 2017.
[23] S. Ghosh, V. K. Yadav, and V. Mukherjee, "Evaluation of cumulative impact of partial shading and aerosols on different PV array topologies through combined Shannon's entropy and DEA," *Energy*, pp. 765–775, 2017.
[24] S. Ghosh, V. K. Yadav, V. Mukherjee, and P. Yadav, "Evaluation of relative impact of aerosols on photovoltaic cells through combined Shannon's entropy and Data Envelopment Analysis (DEA)," *Renew. Energy*, vol. 105, pp. 344–353, 2017.
[25] C. E. Shannon, "A Mathematical Theory of Communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 1948.
[26] J. Doyle and R. Green, "Efficiency and Cross-efficiency in DEA : Derivations , Meanings and Uses," *J. Oper. Res. Soc.*, vol. 45, no. 5, pp. 567–578, 1994.
[27] S. Jen and S. Toshiyuki, "A unified framework for the selection of a Flexible Manufacturing System," *Eur. J. Oper. Res.*, vol. 85, pp. 297–315, 1995.