ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)
Vol 3, Issue 3, March 2018

Deterministic Scenario vs. Stochastic Scenario in Data Envelopment Analysis (DEA): A Case on Iranian Forest Management Units

[1] Majid Zadmirzaei Soleimandarabi, [2] Soleiman Mohammadi Limaei, [3] Alireza Amirteimoori
[1] [2] Department of Forestry, Faculty of natural resources, University of Guilan, Iran, [3] Department of Applied Mathematics, Faculty of Science, Islamic Azad University, Rasht, Iran

Abstract: **-- The aim of this paper is to estimate the relative efficiency of Iranian forest management units with both deterministic and stochastic Data Envelopment Analysis (DEA). Towards this end, the required data were collected from 14 Iranian forest management units included 2 inputs (growing stock, plantation costs) and 2 outputs (harvesting revenues, forest protection tasks). In deterministic scenario, the most frequently applied DEA model-BCC is used in which results showed that most of the forest management units are operating at low efficiency levels (just 35.71% efficient units). However, in forestry, some external uncertain factors such as socio-economic and climate factors influence the magnitude of harvesting revenues of forest management units. Thus, in uncertain scenario, an output- oriented Chance Constrained Data Envelopment Analysis (CCDEA) is used. Results were significantly different according to different risk criterion levels (α). Nevertheless, using the Kendall's tau correlation test showed by increasing the probability levels (1- α) of CCDEA their result are closed to deterministic DEA, it means that there were 5 efficient units (2, 3, 8, 9, and 10) in the all scenarios. Therefore, according to the output-oriented nature of DEA models, the managers of the inefficient units should increase their output while their inputs proportions remain unchanged; otherwise they will not be able to promote their overall productivity.**

Keywords— Deterministic vs. stochastic DEA models, Forest management units, Kendall's tau correlation test, measuring the performance.

I. INTRODUCTION

The foundation of efficiency evaluation is to identify the corresponding production possibility set for making a good decision as the challengeable issue to improve the total productivity of any kind of organization. Owing to the multiple benefits and advantages (economic, ecological and social functions) offered by the forest and also the non-market nature of part of these outputs, measuring the efficiency in forestry is highly demanding as well [1]. In doing so, the well documented method in the operations research society is Data Envelopment Analysis (DEA). DEA, which is a mathematical optimization model, has used to establish a best practice group from a set of observed units in which then provide the basis for rating units. DEA also quantifies the level of inefficiency and selects exemplary units for inefficient units to mimic. However, there is a weakness in conventional DEA models; in fact, deterministic DEA models do not allow stochastic variations in input and output such as data entry errors. As a result, DEA efficiency measurement may be sensitive to such variations. A Decision Making Units (DMU) which is measured as efficient relative to other DMUs may turn inefficient if such random variations are considered. Thus, to close this substantial gap, Stochastic DEA (SDEA) methods have therefore been designed to deal with the problems which are introduced by uncertainty [2]. The Chance Constrained DEA (CCDEA) is an important applied SDEA method for managing risk arising from random variations in natural. It is one way to manipulate uncertain data in DEA via probability distributions. Seminal work by Sengupta showed how stochastic variables could be included in the non-parametric framework [3]. Afterwards, the other researchers made major breakthroughs in this aspect [4-6]. In forestry, some external uncertain factors such as socio-economic and climate factors can influence over marketable or quantitative forest outputs such as harvesting revenue, growing stock, etc, so they should be considered as random variable, however, some other outputs (forest protection tasks, non-wood forest products…) are unmarketable or qualitative that never be affected by external uncertain factors; thus they should be considered as deterministic outputs. With this intention, it is better to use a hybrid output- oriented CCDEA model with both random and deterministic output variables. Hence, in this study, the DEA approaches are applied to compare two different scenarios on forest management units. Such an application may provide an illustrative linkage between the deterministic and stochastic DEA in real decisional problems on forest management background.

ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)
Vol 3, Issue 3, March 2018

II. METHODOLOGY

DEA models

Mathematically, a variety of DEA models are available to measure the relative efficiency. However, when the conventional DEA models are applied in practical problems, several issues are raised. Thus, the DEA models often have to be modified for overcoming these problems. Here we treat only BCC model in both deterministic and stochastic approaches although the results can also hold for other DEA models with their associated production possibility sets.

B. Deterministic DEA approach

Suppose we have a set of n peer DMUs, {DMUj: j = 1, 2, …, n}, which produce multiple outputs y_{rj}, (r = 1, 2, ..., s) ∈ R_{+}^{S}, by utilizing multiple inputs x_{ij}, (i = 1, 2, ..., m) ∈ R_{+}^{M}. Here, due to the output-oriented nature of these models, the objective function tries to increase output amounts (1a, 1b) by fixing input levels. In fact, θ is a real decision variable and λ is a non- negative vector of decision variable. So, it is clear that $\theta \geq 1$, but when a DMU is fully efficient if and only if the optimal solution problem $\theta^{*} = 1$ **Error! Reference source not found.**.

$$e^{BCC}(y,x)\max = \theta \qquad (1a)$$

$$st:$$

$$\sum_{j=1}^{n} x_{ij}\lambda_j \leq x_{i0} \qquad (1b)$$

$$\sum_{j=1}^{n} y_{rj}\lambda_j \geq \theta y_{r0} \qquad (1c)$$

$$\sum \lambda_j = 1 \qquad (1d)$$

$$\lambda_j \geq 0 \qquad (1e)$$

C. Stochastic DEA approach

As mentioned before, in the real world, we face up to insecure data. The one advantage of Chance Constrained DEA (CCDEA) model is that, these models deal with insecure data. According to our real problem, we consider one stochastic and one deterministic output that denoted by $\tilde{y} = (\tilde{y}_1, \dots, \tilde{y}_S) \in \mathbb{R}_{+}^{S}$ and$y = (y_1, \dots, y_S) \in \mathbb{R}_{+}^{S}$, respectively. So, obviously, the deference between model (1) and (2) is just the first constraint (2b). Here, OT and OK are sets of random and deterministic output index, respectively, P means “Probability’’ and α is considered as a risk criterion representing utility of a manager that is a predetermined number between 0 and 1, so (1-α) shows the probability levels.

$$Max = \theta \qquad (2a)$$

$$st:$$

$$P\left(\sum_{j=1}^{n}\lambda_j \tilde{y}_{rj} \geq \theta \tilde{y}_{ro}\right) \geq 1-\alpha \qquad r \in O_T \qquad (2b)$$

$$\sum_{j=1}^{n}\lambda_j y_{rj} \geq \theta y_{ro} \qquad r \in O_K \qquad (2c)$$

$$\sum_{j=1}^{n}\lambda_j x_{ij} \leq x_{i0} \qquad (2d)$$

$$\sum_{j=1}^{n}\lambda_j = 1 \qquad (2e)$$

$$\lambda_j \geq 0 \qquad (2f)$$

It can be easily thought that the above CCDEA model needs to be reformulated to obtain its feasibility. Indeed, CCDEA is a kind of robust LP with random constraints in which they can been easily expressed as the Second-Order Cone Programming (SOCP), interested reader is refer to section 4.4.2: Second-Order Cone Programming on page 156 of Boyed and Vandenberghe for more detailed discussions regarding this issue [8]. In this study, the first constraint (2b) should be reformulated as the deterministic equivalent one by CCP proposed by Charnes and Cooper **Error! Reference source not found.**. Now suppose $\zeta_r \geq 0$ is the “external slack” for the rth output. By “external slack” we refer to slack outside the braces. We can choose the value of this external slack, so it satisfies

$$P\left(\sum_{j=1}^{n}\lambda_j \tilde{y}_{rj} - \theta \tilde{y}_{ro} \geq 0\right) = 1-\alpha+\zeta_r$$

There must then exist a positive number $S_r^{+} \geq 0$ such that

$$P\left(\sum_{j=1}^{n}\lambda_j \tilde{y}_{rj} - \theta \tilde{y}_{ro} \geq s_r^{+}\right) = 1-\alpha$$

This positive value of $S_r^{+} \geq 0$ allows a still further increase in $\tilde{y}_{r0}$ for any set of sample observations without worsening any other input or output. It is easy to see that $\zeta_r \geq 0$ if and only if $S_r^{+} = 0$.

We suppose our output is a random variable with a multivariate normal distribution and known parameters. We

ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)
Vol 3, Issue 3, March 2018

also restrict our attention to the class of zero-order decision rules, now we have

$$\alpha = \mathrm{P}\left(\sum_{j=1}^{n}\lambda_j \tilde{y}_{rj} - \theta \tilde{y}_{ro} \le s_r^+\right) = \mathrm{P}\left(\tilde{z}_i \le \frac{s_r^+ - (\sum_{j=1}^{n}\lambda_j y_{rj} - \theta y_{ro})}{\sigma_r^o(\lambda,\theta)}\right),$$

Where $\tilde{Z}$ is the standard normal random variable (with zero mean and unit variance), and

$$(\sigma_r^o(\lambda,\theta))^2 = \sum_{j\neq 0}^{n}\sum_{k\neq 0}^{n}\lambda_j\lambda_k \operatorname{cov}\left(\tilde{y}_{rj}, \tilde{y}_{rk}\right) +$$

$$2(\lambda_0 - \theta)\sum_{j\neq 0}^{n}\lambda_j \operatorname{cov}\left(\tilde{y}_{rj}, \tilde{y}_{r0}\right) + (\lambda_0-\theta)^2 \operatorname{var}\left(\tilde{y}_{r0}\right)$$

is the variance of $\sum_{j=1}^{n}\lambda_j \tilde{y}_{rj} - \theta \tilde{y}_{ro}$. We note that

$$\Phi\left[\frac{s_r^+ - \left(\sum_{j=1}^{n}\lambda_j y_{rj} - \theta y_{ro}\right)}{\sigma_r^o(\lambda,\theta)}\right] = \alpha$$

Where Φ is the standard normal distribution with its limit of integration shown in the square brackets. We note that this expression is free of random elements. Because Φ is normal we know it has an inverse. Thus we can also write this as:

$$\frac{s_r^+ - \sum_{j=1}^{n}\lambda_j y_{rj} - \theta y_{ro}}{\sigma_r^o(\lambda,\theta)} = \Phi^{-1}(\alpha),$$

i.e.,

$$\sum_{j=1}^{n}\lambda_j y_{rj} - s_r^+ + \Phi^{-1}(\alpha)\sigma_r^o(\lambda,\theta) = \theta y_{r0}$$

Where Φ^{-1} is the so-called "fractile function".

It should be noted that the functional form $(\sigma_r^o(\lambda,\theta))^2$ causes a nonlinear programming, so we Should replace u_r^2 with $(\sigma_r^o(\lambda,\theta))^2$, until it can be transformed to a quadratic problem then Put this all together we find that we have the following problem to solve by algorithms available for this class of problems (such as interior points or barrier methods).

$$Max = \theta \qquad (3a)$$

$$\sum_{j=1}^{n}\lambda_j y_{rj} - s_r^+ + \Phi^{-1}(\alpha)u_r = \theta y_{ro} \qquad r\in O_T \qquad (3b)$$

$$\sum_{j=1}^{n}\lambda_j y_{rj} \ge \theta y_{ro} \qquad r\in O_K \qquad (3c)$$

$$\sum_{j=1}^{n}\lambda_j x_{ij} \le x_{i0} \qquad (3d)$$

$$u_r^2 = \sum_{j\neq 0}^{n}\sum_{k\neq 0}^{n}\lambda_j\lambda_k \operatorname{cov}\left(\tilde{y}_{rj}, \tilde{y}_{rk}\right) + \qquad (3e)$$

$$2(\lambda_0-\theta)\sum_{j\neq 0}^{n}\lambda_j \operatorname{cov}\left(\tilde{y}_{rj}, \tilde{y}_{r0}\right) + (\lambda_0-\theta)^2 \operatorname{var}\left(\tilde{y}_{r0}\right)$$

$$\sum_{j=1}^{n}\lambda_j = 1 \qquad (3f)$$

$$s_r^+ \ge 0, \lambda_j \ge 0, u_r \ge 0 \qquad (3g)$$

This problem, which is free of random elements, is the desired deterministic equivalent for (2) - a term which is justified because an optimal choice of the variables in (3) will also be optimal for (2) and, vice versa, an optimal solution of (2) will also be optimal for (3).

D. Kendall's tau correlation test

To compare the results of two different scenarios the Kendall's tau correlation test is used in order to know how much correlation there is between these scenarios. The Kendall's tau is a nonparametric correlation and it should be used rather than Spearman's coefficient you have a small data set with a large number of tied ranks. This means that if you rank all of the scores and many scores have the same rank, the Kendall's tau should be used. Although when Spearman's statistic is more popular of the two coefficients, there is much to suggest that Kendall's statistic is actually a better estimate of the correlation in the population **Error! Reference source not found.**.

III. DATA COLLOCATION AND CASE STUDY

Iranian Caspian forests, which are located at Guilan province in the north of Iran, are considered as a real case study. Industrial harvesting occurs only in these forests because the severe climatic conditions and forest degradation, forests in other regions are not exploited for industrial wood production.
Based on the aims of this study and forestry experts' ideas,

ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)
Vol 3, Issue 3, March 2018

two deterministic inputs: growing stock and plantation costs were considered, moreover, one stochastic output (harvesting revenues) and one deterministic output (forest protection tasks) were included in DEA analyses. It should be noted that the length of the planning horizon includes 10 years. Hence, this case, the average data of a ten-year period were considered. Moreover, the monetary data were adjusted by the Consumer Price Index (CPI) of Iran in the base year 2011 (Table I).

Table I Mean Values of Input and Output Variables of Forest Management Units Used In Dea Models

		Input		Output	
DMUs	Forest management units	Growing stock (m^3/ha)	Plantation costs (Iranian million Rials/ha)	Harvesting revenues (Iranian million Rials/m^3/ha)	Forest* protection tasks
1	Shafaroud 17	210.52	53.07177	14173.86	1.53
2	Shafaroud 14	170.38	39.80383	45585.07	4.33
3	Shafaroud 11	293.74	100.8364	89788.29	4.4
4	Shafaroud 9	297.81	437.8421	66850.6	1.46
5	Shafaroud 4	349.189	788.1158	25351.57	2.46
6	Shafaroud 2	197.246	68.9933	42826.83	1.46
7	Nav 14	255.4	31.84306	36958.66	4.46
8	Nav 12	155.43	87.56842	30357.46	1.46
9	Nav 3	273.81	132.6794	142413.3	4.33
10	Nav 2	219.62	254.2138	104653.9	2.46
11	Nav 1	246.23	63.68613	71330.72	2.33
12	Lomer 8	231.77	236.1694	22571.13	1.53
13	Lomer 2	196.71	119.4115	30773.3	1.53
14	Lomer 1	432	443.1493	134588.7	2.46

* using questionnaires with five-level Likert item

IV. RESULTS AND DISCUSSION

According to results of BCC model, most of the DMUs were inefficient. In other words, there were just five (35.71%) efficient forest management units. The results of CCDEA model, were completely different with respect to different risk criteria levels (α), however, the Kendall's tau correlation test shows that there is a good correlation between two models by increasing the probability levels (1-α). The results of DEA models are shown in Table II and Table III. Moreover, the results of this study are in line with those obtained by other researchers. For instance, in a recent study the researchers figured out that by increasing the probability levels the results of Fuzzy DEA and CCDEA are really closed to basic DEA [11].

Table Ii Deterministic and Stochastic Efficiency Scores Of Forest Management Units

DMUs	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Percent efficient units (%)
DEA	0.36	1	1	0.47	0.55	0.37	0.72	1	1	1	0.52	0.34	0.44	0.55	35.71
DEA (α=0.1)	1	1	1	0.46	0.51	0.6	1	1	1	1	1	1	0.41	0.9	64.28
DEA (α=0.2)	1	1	1	0.46	0.51	0.6	1	1	1	1	1	0.35	0.41	0.9	57.14
DEA (α=0.3)	0.36	1	1	0.46	0.51	0.6	1	1	1	1	1	0.35	0.41	0.9	50
DEA (α=0.4)	0.36	1	1	0.46	0.53	0.44	1	1	1	1	0.6	0.35	0.41	0.66	42.85

Table III Matrix of Kendall's Tau Correlation Test

Correlation	DEA	DEA (α=0.1)	DEA (α=0.2)	DEA (α=0.3)	DEA (α=0.4)
DEA	1				
DEA (α=0.1)	0.598**	1			
DEA (α=0.2)	0.653**	0.693**	1		
DEA (α=0.3)	0.615**	0.644**	0.851**	1	
DEA (α=0.4)	0.690**	0.708**	0.857**	0.890**	1

V. CONCLUSION

To sum up, in this study, two different deterministic and stochastic DEA scenarios were compared in forest management units, as a real problem. In this regard, a hybrid stochastic output-oriented CCDEA model was considered against deterministic one, because in forest management section, harvesting revenue as a marketable and or/ quantitative output can be affected by external uncertain factors but forest protection task as a unmarketable or qualitative one never be affected by them. Apparently the results of these two different scenarios showed some differences although the Kendall's tau correlation test shows that there is a good direct correlation between them. Thus, the modified CCDEA model can suitably discriminate the efficiency and/or inefficiency of each unit and can be applied to hedge against risk and uncertainty in the forest management. Finally, according to these practical consequences, there were just 5 (35.71%) efficient units (2, 3, 8, 9, and 10) in the all scenarios. Therefore, base upon the output- oriented nature of DEA models, the managers of the inefficient units should increase their output while their inputs proportions remain unchanged; otherwise they will not be able to promote their overall productivity.

REFERENCES

[1] M. Zadmirzaei, S. Mohammadi limaei, L. Olsson, and A. Amirteimoori, “Assessing the impact of the external non-discretionary factor on the performance of forest management units using DEA approach,” Journal of Forest Research., vol. 22, no. 3, pp. 144-152, 2017.

[2] M. Khodabakhshi, “An output oriented super-efficiency measure in stochastic data envelopment analysis: Considering Iranian electricity distribution companies,” Computers & Industrial Engineering., vol. 58, no. 4, pp. 663–671, 2010.

[3] J. K. Sengupta, “Efficiency measurement in stochastic input-output systems,” International Journal of Systems Science., vol. 13, pp. 273–287, 1982.

[4] K. Land, C. A. K. Lovell, and S. Thore, “Chance-constrained data envelopment Analysis,” Managerial and Decision Economics., vol. 14, pp. 541–554, 1994.

[5] O. B. Olesen, and N. C. Petersen, “Chance Constrained Efficiency Evaluation.” Management Science., vol. 41, pp. 442-457, 1995.

[6] W. W. Cooper, H. Z. Deng, M. Huang, and S. X. Li, “Chance constrained programming approaches to congestion in stochastic data envelopment analysis,” European Journal of Operations Research., vol. 155, no. 2, pp. 487–501, 2004.

[7] W. W. Cooper, L. M. Seiford, and J. Zhu, “Handbook on Data Envelopment Analysis,” 2nd Edition, Springer Science+Business Media, LLC, pp. 512, 2011.

[8] S. Boyed, and L. Vandenberghe, “Convex Optimization”, Cambridge University Press. United Kingdom, pp. 716, 2004.

[9] A. Charnes, and W.W. Cooper, “Deterministic Equivalents for Optimizing and Satisfying under Chance Constraints,” Operations Research., vol. 11, no. 1, pp. 18–39, 1963.

[10] A.P. Field, “Discovering statistics using SPSS for windows: advanced techniques for the beginner,” SAGE publication, Social Science. United Kingdom, pp. 496, 2000.

[11] A. Azadeh, S.M. Alem, “A flexible deterministic, stochastic and fuzzy Data Envelopment Analysis approach for supply chain risk and vendor selection problem: Simulation analysis,” Expert Systems with Applications, vol. 37, pp. 7438–7448, 2010.