Comparison of Data Envelopment Analysis Methods Used in Airport Benchmarking

David Schaar and Lance Sherry, Ph.D.
Dept. of Systems Engineering and Operations Research
George Mason University
Fairfax, Virginia, USA

Abstract—**Airport efficiency has been shown to contribute to the overall performance efficiency of the air transportation network. Airport performance efficiency is benchmarked annually and widely published. These benchmarks use several techniques, including several Data Envelopment Analysis (DEA) methods.**

This paper examines the differences in results using three DEA methods (Cooper-Charnes-Rhodes (CCR), Banker-Charnes-Cooper (BCC), and Slacks-Based Measure of efficiency (SBM)) on data from 45 airports from 1996 to 2000. The results from the three DEA methods yielded wide variation in results. For example, the CCR analysis showed that efficiencies degraded from small to medium to large airports. The BCC analysis showed no significant difference in efficiency among the three classes of airports. The SBM analysis yielded degraded efficiency from large to medium to small airports. The implications of these results on the use of DEA in benchmarking and the need for guidelines for selection of DEA models and the interpretation of DEA results is discussed.

Keywords-Data Envelopment Analysis; CCR; BCC; SBM; airport efficiency

I. Introduction

Airport congestion is growing and traveler satisfaction is dropping [1], yet demand is expected to grow at major hub airports [2].

In this environment, the operational efficiency of airports is one of the important determinants of the system's future success. An important component to achieving operational efficiency is to use performance measurement to understand which airports are performing well and which are underperforming and also to understand what the drivers are behind good and bad performance, all with the ultimate goal of improving performance in the areas that are deficient.

As we will show in this paper, much has been written in the academic literature about measuring airport efficiency but several opportunities for improvement exist in this work.

One such opportunity relates to the theoretical foundations for evaluating performance: A number of studies have been conducted where airports are ranked against each other. However, in these studies there has been limited focus on the selection of the evaluation models used to identify good and bad performers. Several choices exist in this area and a study of each model is necessary to understand which is the most appropriate in a given situation.

This paper will highlight this issue through the use of an example. We will re-run an earlier analysis of airport performance and highlight the impact of the choices in the areas listed above.

II. Background

A. General Background

Performance benchmarking as a management tool is frequently cited as having been pioneered by Robert C. Camp at the Xerox Corporation in the late 1970's [7],[8] where he used this tool to identify shortcomings of that company's photocopier production and distribution. He used benchmarking to identify several key practices from the photocopier industry as well as from other industries, all serving to ultimately improve the company's photocopier business. The performance measurement and benchmarking techniques have all evolved significantly since that time, but the ultimate goals for any benchmark remain the same: To identify performance gaps and to find practices that will help close that gap.

Benchmarking has since been applied both in industry and academia in numerous studies. Dattakumar and Jagadeesh [6] conducted an extensive literature review and found more than 350 publications on the topic as of June 2002.

Airport benchmarking studies began appearing in the literature in the early 1990's with one of the first such studies being that of Tolofari, Ashford, and Caves [9] in which airports operated by the British Airport Authority were compared against each other. A series of benchmarking studies have since appeared with a variety of geographic foci and covering a diverse set of performance parameters. A few examples of these studies are listed in Table I. TABLE I.

TABLE I. REPRESENTATIVE AIRPORT BENCHMARKING STUDIES

Study Title	Inputs (resources)	Outputs	Analytical Technique
Measuring airport quality from the airlines' viewpoint [10]	Airport charges; minimum connecting times; number of passenger terminals; number of runways; distance to nearest city center	Level of satisfaction from the airline users of each airport	DEA
Size versus efficiency: a case study of US commercial airports [11]	Operational costs; non-operational expenses; number of runways; number of gates	Passenger throughput; aircraft movements; aeronautical revenue; non-aeronautical revenue; percentage of on-time operations	DEA
Relative efficiency of european airports [12]	Air transport movement study: Airport surface area; total length of runways; number of aircraft parking positions at terminals; number of remote aircraft parking positions	Aircraft movements	DEA / Parametric TFP
	Passenger movement study: Terminal size; number of aircraft parking positions at terminals; number of check-in desks; number of baggage claim belts	Passenger throughput	DEA / Parametric TFP
Developing measures of airport productivity and performance [13]	Terminal efficiency study: Number of runways; number of gates; terminal area; number of employees; number of baggage collection belts; number of public parking spots	Passenger throughput; cargo throughput	DEA
	Aircraft movement study: Airport surface area; number of runways; runway area; number of employees	Aircraft movements	DEA
Performance Based Clustering for Benchmarking of US Airports [14]	Operational costs; number of employees; number of gates; number of runways	Operational revenue; passenger throughput; aircraft movements; cargo throughput	DEA
Measuring airports' operating efficiency: a summary of the 2003 ATRS global airport benchmarking report [15]	Number of employees; number of runways; number of gates; terminal area; purchased goods, materials, and services (outsourcing)	Passenger throughput; cargo throughput; aircraft movements; non-aeronautical revenue	Parametric TFP
Measuring total factor productivity of airports - an index number approach [16]	Capital expenditure	Aeronautical revenue; non-aeronatical revenue	Index number TFP

As is clear from Table I, besides being used in the analysis that will be examined here, Data Envelopment Analysis (DEA) is an analytical technique that is frequently used in benchmarking studies. DEA is a non-parametric methodology used to assess the efficiency of a Decision-Making Unit (DMU – e.g. an airport) in converting a set of inputs into outputs.

A number of different versions of the basic DEA model have been developed to address a series of potential shortcomings of the original DEA model. We will examine three variations of the DEA methodology and study the differences in assumptions of the three as well as compare the different outcomes of each. The purpose of this analysis is to understand whether the choice of methodology may have an impact on the outcome of an analysis.

B. *CCR model*

In all variations of the DEA models, the DMU(s) with the best inherent efficiency in converting inputs $X_1, X_2, \ldots, X_n$ into outputs $Y_1, Y_2, \ldots, Y_m$ is identified, and then all other DMUs are ranked relative to that most efficient DMU.

For DMU 0, the basic DEA model (so-called CCR after Charnes, Cooper, and Rhodes [5]) is calculated as follows:

$$\max h_0 = \frac{\sum_r u_r y_{rj_0}}{\sum_i v_i x_{ij_0}}$$

$$\text{subject to } \frac{\sum_r u_r y_{rj}}{\sum_i v_i x_{ij}} \leq 1 \quad \text{for each unit j} \qquad (1)$$

$$u_r, v_i \geq 0$$

The interpretation of u_r and v_i is that they are weights applied to outputs y_{rj} and inputs x_{ij} and the are chosen to maximize the efficiency score h_0 for DMU 0. The constraint forces the efficiency score to be no greater than 1 for any DMU. An efficiency frontier is calculated, enveloping all datapoints in a convex hull. The DMU(s) located on the frontier represent an efficiency level of 1.0, and those located inside the frontier are operating at a less than full efficiency level, i.e. less than 1.0.

The above fractional program is executed once for each participating DMU, resulting in the optimal weights being determined for each DMU.

Before solving the problem, the denominator in the objective function is removed and instead an additional constraint is added. Also, the original constraint is manipulated in order to convert the fractional program to a linear program. These two steps result in the following linear program:

$$\max h_0 = \sum_r u_r y_{rj_0}$$

$$\text{subject to} \quad \begin{aligned} & \sum_r u_r y_{rj} - \sum_i v_i x_{ij} \le 0 \\ & \sum_i v_i x_{ij_0} = 1 \\ & u_r, v_i \ge 0 \end{aligned} \tag{2}$$

In simpler notation, this is written as:

$$\max(\mathrm{v,u}) = u y_0$$

$$\text{subject to} \quad \begin{aligned} & -vX + uY \le 0 \\ & v x_0 = 1 \\ & v \ge 0, u \ge 0 \end{aligned} \tag{3}$$

Finally, before solving, the linear program is converted to its dual for computational efficiency reasons:

$$\min(\theta,\lambda) = \theta$$

$$\text{subject to} \quad \begin{aligned} & \theta x_0 - X\lambda \ge 0 \\ & Y\lambda \ge y_0 \\ & \lambda \ge 0 \end{aligned} \tag{4}$$

With the addition of slack variables, the dual problem becomes:

$$\min(\theta,\lambda) = \theta$$

$$\text{subject to} \quad \begin{aligned} & \theta x_0 - X\lambda = s^- \\ & Y\lambda = y_0 + s^+ \\ & \lambda \ge 0, s^+ \ge 0, s^- \ge 0 \end{aligned} \tag{5}$$

The slack variables can be interpreted as the output shortfall and the input overconsumption compared to the efficient frontier.

C. BCC model

The CCR model is designed with the assumption of constant returns to scale. This means that there is no assumption that any positive or negative economies of scale exist. It is assumed is that a small airport should be able to operate as efficiently as a large one – that is, constant returns to scale. In order to address this, Banker, Charnes, and Cooper developed the BCC model [3].

The BCC model is closely related to the standard CCR model as is evident in the dual of the BCC model:

$$\min(\theta,\lambda) = \theta$$

$$\text{subject to} \quad \begin{aligned} & \theta x_0 - X\lambda = s^- \\ & Y\lambda = y_0 + s^+ \\ & e\lambda = 1 \\ & \lambda \ge 0, s^+ \ge 0, s^- \ge 0 \end{aligned} \tag{6}$$

The difference compared to the CCR model is the introduction of the convexity condition $e\lambda = 1$. This additional constraint gives the frontiers piecewise linear and concave characteristics.

D. SBM model

Finally, the second adjustment to the basic CCR model is the Slacks-Based Measure of efficiency (SBM), proposed by Tone [18]. The motivation for the development of this model is the observation that while both the CCR and the BCC models calculate efficiency scores, neither is able to take into account the resulting amount of slack for inputs and outputs. Consequently, the purpose of this model is to minimize the input and output slacks, resulting in this fractional program, which is converted to a linear program before solving:

$$\min(\lambda, s^+, s^-)\ \rho = \frac{1 - \frac{1}{m}\sum_{i=1}^{m} s_i^- / x_{i0}}{1 + \frac{1}{s}\sum_{r=1}^{s} s_r^+ / y_{r0}}$$

$$
\begin{aligned}
& x_0 - X\lambda = s^- \\
\text{subject to} \quad & Y\lambda = y_0 + s^+ \qquad (7) \\
& \lambda \geq 0, s^+ \geq 0, s^- \geq 0
\end{aligned}
$$

III. Method of Analysis

A. Introduction

The analysis presented here is based on a re-run and extension of the analysis performed by Bazargan and Vasigh [11].

B. Inputs and Outputs

Input and output data was collected from a series of sources in order to assemble the same inputs and outputs as in the Bazargan and Vasigh study for the same airports and years (1996 – 2000). Table II lists the airports covered in the analysis and Table III lists the input and output variables and the source from which they were collected.

TABLE II. AIRPORTS INCLUDED IN STUDY

Large	Medium	Small
ATL	BNA	ALB
DEN	CLE	BHM
DFW	DAL	BOI
DTW	IND	COS
EWR	MCI	DAY
IAH	MDW	ELP
JFK	MEM	GEG
LAS	MSY	GSO
LAX	OAK	GUM
MIA	PDX	LIT
MSP	RDU	OKC
ORD	SJC	ORF
PHX	SJU	RIC
SFO	SMF	ROC
STL	SNA	TUL

TABLE III. AIRPORT DATA SOURCES

Inputs	Source
Operating expenses	FAA Form 127
Non-operating expenses	FAA Form 127
Number of runways	NFDC 5010 database
Number of gates	Airport websites
Outputs	**Source**
Aeronautical revenue	FAA Form 127
Non-aeronautical revenue	FAA Form 127
Portion of on-time operations	BTS Form 41
Total enplanements	TAF
Number of air carrier operations	TAF
Number of other operations	TAF

C. Model Selections

In the Bazargan and Vasigh study, the model selected was the CCR model. No discussion was provided regarding why this model was selected ahead of others; presumably it was selected on the premise of being the "default" DEA model.

One of the purposes of the present study was to examine the impact of different model selections. Hence, efficiency scores were calculated using three different, basic DEA models: CCR, BCC, and SBM. These are considered to be among the standard DEA models by Cooper, Seiford, and Tone [19]. These models were all executed in output-oriented mode, meaning that efficiency scores were calculated by analyzing which levels of outputs should be "possible" to generate if an airport were operating at full efficiency given the current levels of inputs. These scores were calculated using the "Learning Version" of the DEA-Solver software [4].

D. Model Configuration

One of the particular choices made in the Bazargan and Vasigh study, driven by practical considerations, was to create an artificial, "super-efficient" airport by assigning it the minimum input values present in the dataset and the maximum output values. The reason for this was that when the CCR model was initially executed with only the basic list of airports included, a very large portion of the airports were ranked as fully efficient (1.0). The cause of that outcome was the relatively small number of airports (45) as compared to the number of parameters considered (10). In general, this will frequently be a phenomenon in analysis having a low ratio of DMUs to parameters.

Bazarghan and Vasigh also describe the theory of adding a condition requiring all weights u and v to be >= ε (an infinitesimal value) in order to avoid setting all but one input and one output's weights to a non-zero value. However, it is not indicated if this was used in the study and if it was, there is no specification of the value used for this infinitesimal value is provided. Consequently, the model will in this case be run without this additional constraint.

IV. Results and Discussion

A. Comparison of CCR Model Runs

The results of the CCR model runs are shown in Table IV, Table V, and Table VI. These values are in general lower than the values calculated in the Bazargan and Vasigh study and the reason for this can not be ascertained since the raw data of the previous study were not available. It can be postulated that the difference is due to differences in the inputs; in particular if the values assigned to the artificial, "super-efficient" airports were different in the two studies, since that would cause a general offset between the two.

What is noteworthy, however, is that the general trend from the Bazargan and Vasigh study is confirmed: Large airports are in general exhibiting lower efficiency scores than medium-sized airports, and medium-sized airports are exhibiting lower scores than small airports. These results are displayed in Figure 1. and Figure 2. The Kruskal-Wallis test was performed to determine whether the differences in efficiency ranks among the groups were significant, and that was found to be the case. The results from the Kruskal-Wallis test are presented in Table VII.

TABLE IV. RESULTS OF CCR RUNS FOR LARGE AIRPORTS

Airport	1996	1997	1998	1999	2000
ATL	0.1910	0.1990	0.2000	0.2000	0.2000
DEN	0.1515	0.1667	0.1667	0.1435	0.1281
DFW	0.1216	0.1231	0.1380	0.1199	0.1251
DTW	0.1405	0.1310	0.1256	0.1400	0.1432
EWR	0.2335	0.2595	0.2614	0.2511	0.2884
IAH	0.2202	0.2141	0.2149	0.2145	0.2238
JFK	0.2500	0.2491	0.2485	0.2500	0.2500
LAS	0.2034	0.2088	0.1977	0.2005	0.1946
LAX	0.2953	0.2995	0.2950	0.2698	0.2719
MIA	0.2500	0.2500	0.2500	0.2500	0.2500
MSP	0.2064	0.2015	0.1932	0.2155	0.2192
ORD	0.1667	0.1667	0.1667	0.1667	0.1667
PHX	0.2723	0.2765	0.2587	0.2674	0.2607
SFO	0.1784	0.1962	0.1783	0.1946	0.1804
STL	0.1937	0.2083	0.2032	0.2117	0.2134
Average	0.2050	0.2100	0.2065	0.2064	0.2077

TABLE V. RESULTS OF CCR RUNS FOR MEDIUM SIZED AIRPORTS

Airport	1996	1997	1998	1999	2000
BNA	0.2140	0.2140	0.2179	0.2143	0.2186
CLE	0.2054	0.2132	0.2142	0.2085	0.2183
DAL	0.5603	0.5087	0.5257	0.5319	0.5262
IND	0.2811	0.3043	0.3044	0.3026	0.3067
MCI	0.2674	0.2816	0.2812	0.2836	0.2843
MDW	0.2361	0.2298	0.2296	0.2260	0.2296
MEM	0.2206	0.2070	0.2065	0.2256	0.2289
MSY	0.2765	0.2799	0.2869	0.2845	0.2916
OAK	0.4802	0.4704	0.5455	0.5147	0.5455
PDX	0.2518	0.2764	0.2778	0.2855	0.2767
RDU	0.4480	0.4152	0.4027	0.3749	0.3807
SJC	0.3022	0.3156	0.3051	0.3057	0.2951
SJU	0.3374	0.3967	0.4215	0.4155	0.4378
SMF	0.4121	0.4252	0.4171	0.4205	0.4019
SNA	1.0000	1.0000	0.9914	1.0000	0.9882
Average	0.3662	0.3692	0.3752	0.3729	0.3753

TABLE VI. RESULTS OF CCR RUNS FOR SMALL AIRPORTS

Airport	1996	1997	1998	1999	2000
ALB	0.4772	0.5101	0.5047	0.4895	0.5099
BHM	0.5345	0.5400	0.5482	0.5461	0.5698
BOI	0.8089	0.5714	0.7842	0.5392	0.8086
COS	0.5191	0.5561	0.5672	0.5725	0.5625
DAY	0.3472	0.3708	0.3674	0.3620	0.3652
ELP	0.6856	0.6810	0.6863	0.6876	0.6757
GEG	0.4460	0.4662	0.4726	0.5016	0.4695
GSO	0.8082	0.8553	0.8350	0.8115	0.8459
GUM	0.5714	0.5714	0.5714	0.5714	0.5714
LIT	0.8031	0.8385	0.8493	0.8590	0.8710
OKC	0.5854	0.6006	0.6053	0.6155	0.5961
ORF	0.3994	0.4555	0.4217	0.4122	0.4306
RIC	0.4214	0.4470	0.4403	0.4326	0.4439
ROC	0.4341	0.4572	0.4570	0.4321	0.4374
TUL	0.5098	0.5332	0.5373	0.5349	0.5472
Average	0.5568	0.5636	0.5765	0.5578	0.5803

Figure 1. Results of Bazargan and Vasigh study

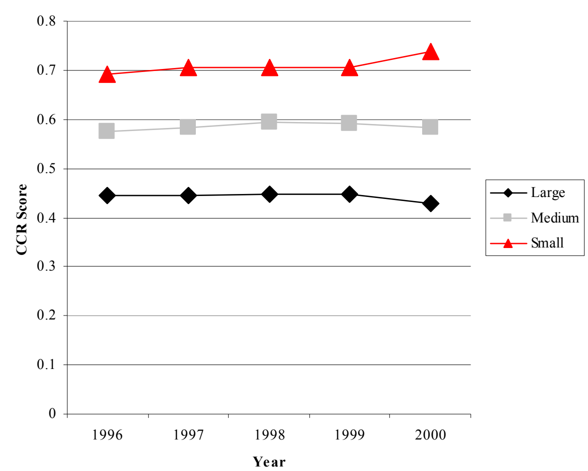

Figure 2. Results of CCR re-run of Bazargan and Vasigh study

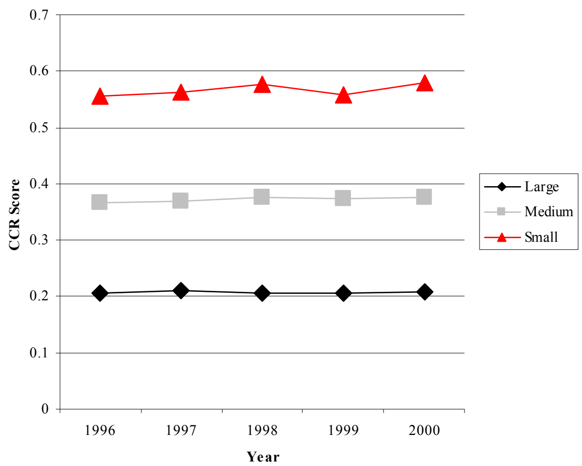

TABLE VII. RESULTS OF KRUSKAL-WALLIS TEST ON CCR RANKS

	Mean efficiency rank				
Year	**Large airports**	**Medium airports**	**Small airports**	**Chi square**	**Asymptotic significance**
1996	36.13	22.40	10.47	28.691	0.000
1997	36.07	23.07	9.87	29.848	0.000
1998	36.33	22.60	10.07	30.200	0.000
1999	36.33	22.53	10.13	29.876	0.000
2000	36.27	22.60	10.13	29.716	0.000

B. *Comparison of CCR, BCC, and SBM Results*

In applying the three DEA models (CCR, BCC, and SBM) to the same dataset, some striking results emerged: While the CCR analysis showed that small airports were more efficient than both medium sized and large airports, and that medium sized airports were more efficient that large airports, in the BCC model no significant difference among the three groups was observed, and in the SBM dataset the difference was in fact reversed. These findings are displayed in the following figures and tables, where the average efficiency scores are displayed and the results of the Kruskal-Wallis tests on the efficiency ranks.

Figure 3. Results of BCC analysis

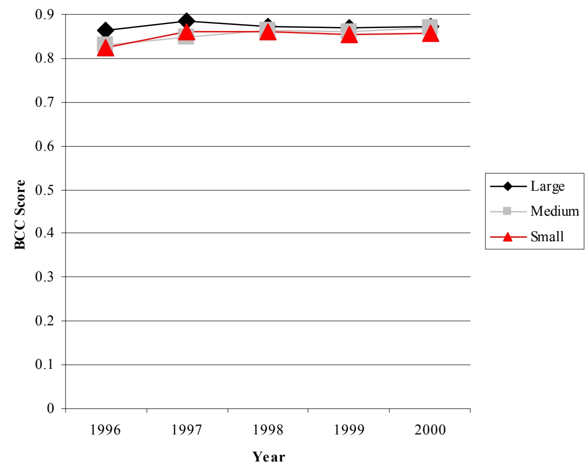

Figure 4. Results of SBM analysis

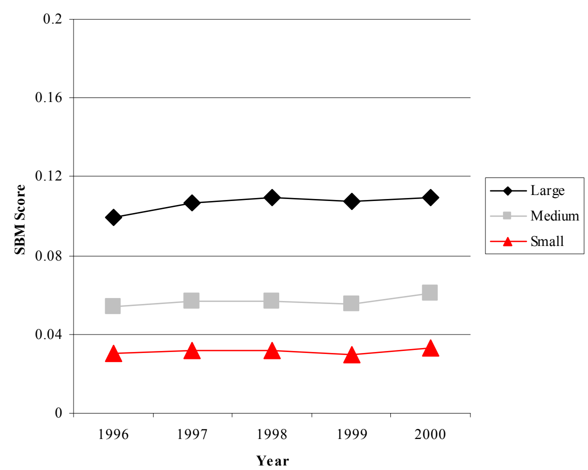

TABLE VIII. RESULTS OF KRUSKAL-WALLIS TEST ON BCC RANKS

	Mean efficiency rank				
Year	**Large airports**	**Medium airports**	**Small airports**	**Chi square**	**Asymptotic significance**
1996	18.53	23.47	27.00	3.149	0.207
1997	22.33	25.67	21.00	1.006	0.605
1998	23.33	24.97	20.70	0.808	0.668
1999	23.40	23.37	22.23	0.077	0.962
2000	21.40	22.63	24.97	0.572	0.751

TABLE IX. RESULTS OF KRUSKAL-WALLIS TEST ON SBM RANKS

Year	Mean efficiency rank			Chi square	Asymptotic significance
	Large airports	Medium airports	Small airports		
1996	9.67	23.60	35.73	29.589	0.000
1997	9.67	23.67	35.67	29.449	0.000
1998	9.47	23.40	36.13	30.939	0.000
1999	9.33	23.33	36.33	31.710	0.000
2000	9.47	22.93	36.60	32.010	0.000

These results point very clearly to the fact that the choice of DEA model has a strong impact on the outcome of this study and that selecting the appropriate model is paramount. While the selection of the appropriate model is key, and previous studies have pointed this out, the literature does not provide many examples of or guidance for selecting the most appropriate model.

In comparing which inputs and outputs are assigned significance in the three different models, the following two observations can be made:

- The same weights were assigned to outputs in the BCC as in the CCR model, and these weights were all focused on a single output for each airport. However, input weights differed between the two models
- In comparing the CCR and the SBM models, the same input variables are assigned non-zero weights, although these weights are different. On the output side, the SBM model assigns non-zero weights to a much broader set of variables than does the CCR model, which usually assigns non-zero values to only a single variable. It can be argued that this leads to the SBM model covering a more comprehensive and thereby accurate scope, rather than only considering a single output variable as in the CCR case.

Notably, running the BCC model, which considers variable returns to scale, resulted in either constant or *decreasing* returns to scale. While impossible to depict graphically in a multi-dimensional model, the concept of decreasing returns to scale is reflected in the shape of the convex hull spanning the DMU(s). In practical terms, this means that rather than observing any economies of scale in terms of efficiency, the model resulted in either no difference or actually a negative effect on airport efficiency as scale grew.

C. Observations on the Weakness of Using Artificial, "Super-Efficient" DMUs

In running the CCR analysis using the artificial, "super-efficient" DMU, one serious problem that arises is the fact that for nearly every non-efficient DMU, the model selects just one single input and a single output to have a non-zero weight while all others are assigned a zero weight. That means that the resulting efficiency score is highly skewed and does not give an accurate representation of a DMU's efficiency.

In the analysis, Bazargan and Vasigh make reference to the use of an infinitesimal value in a constraint to avoid setting any of the weights to 0, but make no reference as to whether this was used and if so, what values were used. As was observed in this analysis, using such a technique is important in order to achieve an accurate assessment of all DMUs across all input and output parameters rather than focusing the analysis on just a subset of parameters.

An alternative approach to handle the issue of setting weights to 0 is the so-called cross-efficiency score, as originally proposed by Doyle and Green [17]. This method calculates the optimal weights for each DMU in the typically manner for CCR, BCC, or SBM, but then applies each DMU's weights to all other DMUs, and then computes an average score for each DMU. While this does aid in mitigating the effects of one DMU's score being calculated on the basis of only a small set of inputs and outputs, it also can be argued to run counter to one of the fundamental principles of DEA; the assumption that a management tradeoff is made between performing well across the given inputs and outputs for an individual DMU, and therefore the model should permit each DMU to appear in the "best possible light" by selecting its optimal weightings.

V. CONCLUSIONS AND FUTURE WORK

The analysis in this paper has shown that depending on the DEA model chosen, radically different results may be obtained. Consequently, any study of airport efficiency needs to begin with a thorough examination of the models available and a motivation for why a particular model was selected. Without an upfront analysis of this kind, a study's final results may be called into question.

In order to simplify such analysis, a thorough study of the implications of each of the most commonly used models is necessary. Such a study needs to examine the implications and interpretations of each model in an airport context since the characteristics of each model may take on different meanings depending on the application area.

Finally, a further, broader area worthy of additional analysis relates to the selection of inputs and outputs to airport efficiency benchmarks: Before proceeding to calculating the airports' efficiency scores, any study needs to present a discussion of what the true goals of an airport are – for example maximizing passenger throughput, maximizing aircraft movements, minimizing delay, maximizing profits, minimizing costs, etc. Only after determination of the airports' goals can the appropriate inputs and outputs be selected. Many of the analyses to-date have largely omitted this discussion and appear to have selected inputs and outputs merely on the availability of data.

REFERENCES

[1] Bureau of Transportation Statistics, "Flight Delays, Mishandled Bags, Consumer Complaints increase in 2007",

http://www.bts.gov/press_releases/2008/dot017_08/html/dot017_08.html, accessed 10 February 2008.

[2] Federal Aviation Administration, "Terminal Area Forecast Summary 2007 - 2025," HQ-08371, http://www.faa.gov/data_statistics/aviation/taf_reports/media/TAF2007-2025Summary.pdf, accessed 10 February 2008.

[3] R.D. Banker, A. Charnes, and W.W. Cooper, "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis," Management Science vol. 30, iss. 9, pp. 1078-1092, 1984.

[4] K. Tone, "DEA-Solver-LV" software, version 3.0. http://www.saitech-inc.com

[5] A. Charnes, W.W. Cooper and E. Rhodes, "Measuring the Efficiency of Decision Making Units," European Journal of Operational Research 2, pp. 429-444, 1978.

[6] R. Dattakumar and R. Jagadeesh, "A review of literature on benchmarking" Benchmarking: An International Journal, vol. 10, iss. 3, pp. 176-209, 2003.

[7] J. Spendolini, The Benchmarking Book. New York, NY: Amacom, 1992.

[8] M. Yasin, "The theory and practice of benchmarking: then and now," Benchmarking: An International Journal, vol. 9, iss. 3, pp. 217-243, 2002.

[9] S. Tolofari, N. Ashford, and R. Caves, "The cost of air services fragmentation," Loughborough University of Technology, Department of Transport Technology Working Paper TT9010.

[10] N. Adler, J. Berechman, "Measuring airport quality from the airlines' viewpoint: an application of data envelopment analysis," Transport Policy, vol. 8, iss. 3, pp. 171-181, 2001.

[11] M. Bazargan, B. Vasigh, "Size versus efficiency: a case study of US commercial airports," Journal of Air Transport Management, vol. 9, iss. 3, pp. 187-193, 2003.

[12] E. Pels, P. Nijkamp, P. Rietveld, "Relative effiency of European airports," Transport Policy, vol. 8, iss. 3, pp. 183-192, 2001.

[13] D. Gillen, A. Lall, "Developing measures of airport productivity and performance: an application of data envelope analysis," Transportation Research Part E: Logistics and Transportation Review, vol. 33, iss. 4, pp. 261-273, 1997.

[14] J. Sarkis, S. Talluri, "Performance based clustering for benchmarking of US airports," Transportation Research Part A: Policy and Practice, vol. 38, iss. 5, pp. 329-346, 2004.

[15] T. H. Oum, C. Yu, "Measuring airports' operating efficiency: a summary of the 2003 ATRS global airport benchmarking report," Transportation Research Part E: Logistics and Transportation Review, vol. 40, iss. 6, pp. 515-532, 2004.

[16] P. G. Hooper, D. A. Hensher, "Measuring total factor productivity of airports – an index number approach," Transportation Research Part E: Logistics and Transportation Review, vol. 33, iss. 4, pp. 249-259, 1997.

[17] J. Doyle and R. Green, "Efficiency and Cross-Efficiency in DEA: Derivations, Meanings, and Uses," The Journal of the Operational Research Society, vol. 45, iss. 5, pp. 567-578, 1994.

[18] K. Tone, "A slacks-based measure of efficiency in data envelopment analysis," European Journal of Operational Research, vol. 130, iss. 3, pp. 498-509, 2001.

[19] W. W. Cooper, L. M. Seiford, K. Tone. Data Envelopment Analysis. New York, NY: Springer, 2007.