

17th Meeting of the EURO Working Group on Transportation, EWGT2014, 2-4 July 2014,  
Sevilla, Spain

## Features selection based on fuzzy entropy for Data Envelopment Analysis applied to transport systems

Sara Bray<sup>a\*</sup>, Leonardo Caggiani<sup>a</sup>, Mauro Dell’Orco<sup>a</sup> and Michele Ottomanelli<sup>a</sup>

<sup>a</sup>DICATECh – Politecnico di Bari, via E. Orabona 4, Bari 70125, Italy

---

### Abstract

Recently, great attention has been paid to the data envelopment analysis (DEA) for the analysis of efficiency of transportation systems. In real world applications, the data of production processes cannot be precisely measured or can be affected by ambiguity. This is particularly worrying when assessing efficiency with frontier-type models, such as Data Envelopment Analysis (DEA) models, since they are very sensitive to possible data errors. Many research works have faced the problem of using DEA models when the inputs and outputs are uncertain. Fuzzy Theory based methods are one of the approaches that have been recently proposed even without a determined (or unique) framework. In this work we have defined a fuzzy version of the classical DEA models, and, in particular, a feature selection analysis has been developed to investigate the effects of uncertainty on the efficiency of the considered transport services. The feature selection method developed in this paper is based on fuzzy entropy measures and it can be applied to DMUs (Decision Making Units) on the entire frontier. Having identified the efficient and inefficient DMUs in fuzzy DEA analysis, the focus is on the stability of classification of DMUs into efficient and inefficient performers. A numerical example is then presented, considering as DMUs a set of international container ports with given number of inputs and outputs properly modified.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

Selection and peer-review under responsibility of the Scientific Committee of EWGT2014

*Keywords:* Data Envelopment Analysis; Efficiency; Fuzzy theory; Uncertainty; Transportation Services; Fuzzy Entropy

---

---

\* Corresponding author. *E-mail address:* [sara.bray@poliba.it](mailto:sara.bray@poliba.it)

## 1. Introduction

A distinctive feature of the contemporary container port industry is that competition has become fiercer than ever (Tongzon and Heng, 2005). Consequently, improving productivity sufficiently to accommodate a large portion of the forecasted increase in container traffic presents a particular challenge to terminal operators and port authorities.

As the demand for international trade and global logistic services continues to increase, to remain competitive, container terminals have to invest heavily in sophisticated equipment or in dredging channels to accommodate the most advanced and largest container ships (Le-Griffin and Murphy, 2006; Cullinane et al., 2006). It is worth noting that limits in available land constrain a merely physical expansion, especially for urban-centric ports, and increase environmental concerns (Park, Kim and Lu, 2008). In addition, also excessive and inappropriate investments can induce inefficiency and wasting of resources. In this context, expanding port capacity by improving the productivity of terminal facilities, and exploring the critical factors affecting the productivity appears to be a viable solution (Park and Lu, 2010). For a container terminal, productivity performance makes a significant contribution to the terminal's survival prospects and competitive advantage (Park and Lu, 2010). Traditionally, the performance of a container terminal has been evaluated with numerous attempts at calculating and seeking to improve or optimize the operational productivity of cargo handling at the berth and container yard (Lu, Huo and Park, 2012). Under this circumstance, to take the top positions as major container ports, and to keep the competitive advantages, it is necessary that these ports should not only extend their facilities, but also maximize the efficiency of their own operations.

In the literature the Data Envelopment Analysis (DEA) methodology has been applied to the evaluation of container terminal performance. In traditional DEA models, it is assumed that all inputs and outputs are exactly known. But in real world, this assumption may not always be true. On the other hand, in more general cases, the data for evaluation are sometimes imprecise and vague. In order to consider such uncertain input data we have proposed (Bray et al, 2014) to use fuzzy set theory in DEA analysis to evaluate port efficiency.

To this end the present study was supported by the South-East Europe (SEE) Transnational Cooperation Project GIFT – Green Intermodal Freight Transport in which a research task is aimed at defining models and methods to evaluate the efficiency of freight transport corridors using uncertain data.

However, if container terminal managers can gain a proper appreciation of their various productivity factors, they may be able to identify which factors have a more positive influence on productivity.

The core of feature selection analysis for evaluating terminal productivity is to remove input variables one by one, then re-estimate the correlation between productivity and investment. From this perspective, feature selection methods provides a more appropriate benchmark for identifying which factors are more critical for productivity improvement. Thus, in this paper we propose a feature selection method applied to Fuzzy Data Envelopment Analysis.

The feature selection could be very useful in DEA models since it allow to simplify the model, reducing the inputs, relevant measurements and computational burdens. This make the model more oriented to practical use (Luukka, 2011). In addition, insignificant features from the can be removed and the consequent model is more easy to understand. Feature selection has been applied in classical DEA using bootstrap techniques for example in the analysis of airports system efficiency (Button and Neiva, 2014).

In this work for the feature selection process, we have proposed to extend the Fuzzy -DEA model presented in Bray et al. (2014) with a new technique based on fuzzy entropy measures (Luukka, 2011) that allow to identify the more significant inputs to be processed in the Fuzzy-DEA model. We have applied the method to a set of six input data for 16 major international container ports (DMUs). Feature selection method was implemented in MATLAB™ computing environment.

The results of this study can provide a useful reference to port managers for developing improvement strategies. The paper is structured as follows. Section 2 gives a brief literature review on measurement of transport systems efficiency and feature selection analysis. Section 3 gives the descriptions of the method built for evaluating port efficiency considering uncertain data and describes feature selection method based on fuzzy entropy . Section 4 presents feature selection method based on fuzzy entropy through a case of study.

Finally, conclusions are drawn in Section 5.

## 2. Efficiency and feature selection analysis: literature review

In this chapter the literature of efficiency and feature selection analysis is reviewed. First of all DEA studies applied to container terminal/port areas are examined.

The Data Envelopment Analysis methodology has been applied to the evaluation of container terminal performance in the literature. For example, in Roll and Hayuth (1993) the first work to advocate the application of the DEA technique to the terminals' context is presented; it remains a purely theoretical exposition, rather than a genuine application. DEA window analysis using panel data relating to the eight container ports in Japan is presented in Itoh (2002). In Tongzon and Heng (2005), DEA-CCR (Charnes Cooper and Rhodes model) and DEA-Additive models are used to analyze the efficiency of four Australian and 12 other international container ports. Applying DEA to estimate the relative efficiency of a sample of Portuguese and Greek seaports is given in Barros and Athanassiou (2004). In Cullinane and Wang (2007), the relevance of DEA was analyzed to estimate the productive efficiency of the container port industry. Available DEA panel data approaches were applied to a sample of 25 leading container ports and evaluated in Cullinane and Wang (2010). In Lin and Tseng (2007), five models of DEA were applied to identify trends in port efficiency of major container ports in the Asia-Pacific region. The impact of different groups on the efficiency of 28 container ports from 12 countries and regions in Asia was studied in Wu, Yan and Liu (2009). Also air transport systems have been analyzed in Button and Neiva (2014) where a bootstrap approach is proposed to select the main system features.

Nevertheless the port competitiveness measurement, and consequently port classification, are very complex because of the uncertainty due to the lack of available ports data, or to imprecision, and/or vagueness of information; so that traditional mathematical techniques and models could not be the proper approach. In these cases it could be useful to face the problem using soft computing techniques based on fuzzy logic, which have been proved to be more suitable when facing uncertainty. In literature there are few works that explicitly consider the uncertainty lying in freight transportation analysis, especially in measurement of container port efficiency, and even less in Container Terminals (CT) classification (Chou, 2007 and 2010; Huang et al., 2003; Caggiani et al., 2012; Iannucci et al., 2011). However it is relevant to notice that Chou (2007 and 2010) and Huang et al. (2003) apply Multi-Criteria Decision Making (MCDM) methods together with fuzzy feature of indicators. In the port classification it may be deemed appropriate to focusing upon fuzzy approach.

The objectives of all these studies are reflected by the variable specification in models, and now we are going to review the literature on the specification of inputs and outputs, in particular the specification of input variables. In Wu, Yan and Liu (2010), the DEA method has been proven to be a suitable tool for evaluating performance with multiple inputs and outputs in respect to 77 global container ports. It was found that the number of berths and the capital deployed are the most sensitive measures impacting the performance of most container ports. The specification of inputs in the literature is not as unified as that of outputs. We can recognize two groups of input specification that are not mutually exclusive. One group of studies considers as input variables: labour and capital (Liu, 1995; Coto-Millan et al., 2000; Estache et al., 2002; Cullinane and Song, 2003). Another group of studies specifies inputs based on the infrastructure and equipment information, that is, terminal quay length, terminal area, number of cargo handling equipment and storage capacity (Tongzon and Heng, 2005; Cullinane, et al., 2002; Cullinane and Song, 2006; Sun et al., 2006).

In the studies considering labour and capital information as inputs, the configuration of container ports/terminals is neglected, because all factors are aggregated into a single capital variable. In the second group, the studies do not consider labour information, but the specification reflects a more accurate design of the port, with the underlying assumption that requests for labour in the production pertain to the equipment, according to a certain ratio. In this context it is necessary to be cautious because this assumption is not always proper and different equipment requires different numbers of workers and different skill levels.

However, the existing literature reveals a lack of empirical evidence in relation to the comparative effectiveness of feature selection analysis in an application to the port industry.

This paper aims to fill this gap by applying a new approach to analyze container terminal productivity.

### 3. Feature selection method based on fuzzy entropy measure

In this section we will first introduce briefly the fuzzy DEA model (fuzzy CCR input-oriented model) proposed by the authors in Bray et al. (2014) to model uncertain data as a fuzzy sets (Zadeh, 1965). Afterwards, the method used to identify which factors (inputs) have more influence on productivity has been discussed. The concept of fuzzy set theory can incorporate the traditional DEA models so, in this way, we represent input or output data as fuzzy symmetrical triangular numbers.

In fuzzy set theory, the more the membership value approaches one, the closer the corresponding abscissa value is to the respective linguistic variable (fuzzy set). If the membership functions are triangular, then all the considered fuzzy constraints can be expressed as inequalities and depend on the satisfaction  $h$  (Zimmermann, 1996). More formally, assume that there are  $n$  DMUs to be evaluated. Each DMU consumes varying amounts of  $m$  different inputs to produce  $s$  different outputs. Specifically, DMU <sub>$j$</sub>  consumes amounts  $X_j = [x_{ij}]$  of inputs ( $i = 1; \dots; m$ ) and produces amounts  $Y_j = [y_{rj}]$  of outputs ( $r = 1; \dots; s$ ). Moreover, assume that  $x_{ij} > 0$  and  $y_{rj} > 0$ . Let then consider the problem of evaluating the relative efficiency for anyone of the  $n$  DMUs, which will be identified as DMU<sub>0</sub>.

We have figured out that sensitivity analysis would be the next big step of this research to analyze all the inputs used in the Fuzzy-DEA model so, for the sake of completeness, the model is shown below (Bray et al., 2014):

Fuzzy CCR input-oriented model

Max  $h$

Subject to:

$$\sum_r u_r y_{r0} \geq h \quad (1)$$

$$\sum_i v_i x_{i0} = 1$$

$$\sum_r u_r y_{rj} - \sum_i v_i x_{ij} \leq 0$$

$$u_r, v_i \geq 0$$

for  $j = 1, \dots, n$ ;  $r = 1, \dots, s$  and  $i = 1, \dots, m$

$$\begin{cases} \Delta \leq a + [(q_+ - b)(1 - h)] \\ \Delta \geq a - [(a - q_-)(1 - h)] \end{cases}$$

Where  $u_r$  and  $v_i$  are weights assigned to output  $r$  and input  $i$ , respectively and  $\Delta$  represents the fuzzy input ( $x_{ij}$ ) or the fuzzy output ( $y_{rj}$ ) that we are considering. Inequalities representing the fuzzy constraints are reported at the end of the model. Full details on the model, on the symbolism adopted and on the case of study are available in Bray et al. (2014).

In many cases, it is of interest to have suitable measures of impreciseness and vagueness, so called fuzziness measures. The specificity of fuzzy sets is to capture the idea of partial membership (Zadeh, 1965). Taking into consideration the concept of fuzzy sets, De Luca and Termini (1971) suggested that corresponding to Shannon probabilistic entropy (1948), the measure of fuzzy entropy should be:

$$H_1(A) = -\sum_{j=1}^n (\mu_A(x_j) \log \mu_A(x_j) + (1 - \mu_A(x_j)) \log (1 - \mu_A(x_j))) \quad (2)$$

where  $\mu_A(x_j)$  are the membership values.

Newer fuzzy entropy measures were introduced by Parkash et al. (2008), defined as:

$$H_2(A; w) = \sum_{j=1}^n w_j \left( \sin \frac{\pi \mu_A(x_j)}{2} + \sin \frac{\pi (1 - \mu_A(x_j))}{2} - 1 \right) \quad (3)$$

and

$$H_3(A; w) = \sum_{j=1}^n w_j \left( \cos \frac{\pi \mu_A(x_j)}{2} + \cos \frac{\pi(1-\mu_A(x_j))}{2} - 1 \right) \quad (4)$$

These fuzzy entropy measures were used in feature selection process (Luukka, 2011).

In particular, by applying equation (3) and (4), we get entropy values for the features that we are considering. If the uncertainty is high, we expect to get high entropy values.

Based on this assumption, the decision of neglecting a feature is made according to the highest entropy value, since we assume that the contribution of features getting highest entropy values is not relevant.

In other words, applying this method based on fuzzy entropy, we simply find the feature with largest fuzzy entropy value and remove that feature from the data set so that it is not considered in Fuzzy-DEA model. After removing this feature, the procedure is repeated and other features can be removed using this approach (Luukka, 2011). In earlier studies this method was tested only with medical data (Luukka, 2006). In this paper we have adapted the more recent findings in Luukka (2011) to transport planning problems as first step of a Fuzzy DEA model. The next section will provide a case of study, where the feature selection method, previously described, is applied to container ports data set. Specifically, the results obtained by the application of this method are analysed.

#### 4. Case of study

If no significant improvements in classification accuracy are achieved, reducing number of features still has many advantages, like reduction of the data set dimension and therefore simplification of the classification task.

Reducing the number of features to be measured for model implementation makes the efficiency analysis faster, more convenient and less expensive. Simpler models with fewer inputs lead to a more transparent and comprehensible model, providing better explanation of expected results. Fewer model inputs result in simpler models that train and execute faster, and allow training on smaller data sets without the risk of over fitting. Discarding irrelevant and redundant features reduces noise and spurious correlations with the outputs and avoids problems of collinearity between inputs. In this section we report the application of the feature selection model based on fuzzy entropy to four Australian and twelve other international container ports (Tongzon, 2001). Ideally, all activities and resources present in the port should be taken into account to calculate efficiency. However, the decision upon which variables to include in the efficiency evaluation function, largely depends on the availability and quality of the data. For instance, the definition of port outputs depends on the activities undertaken by the port, and therefore it can include the number of passengers arriving/departing/transferring in/from the port; the number of vehicles, or the volume of different handled goods.

For all these reasons, here we apply and extend the methodology shown in the previous section, which is commonly used for selecting features in medical field. In particular, we will determine a classification in order of growing importance of six inputs used in a previous study, to measure efficiency with a fuzzy DEA model (Bray et al, 2014). For sake of completeness, we will give the entire database of outputs and inputs used. The outputs are four: TEUs handled (the number of twenty foot container equivalent units handled), shipcalls (number of ship visits), shiprate (ship working rate, which measures the number of containers moved per working hour), crane prod. (crane productivity, which measures the number of containers moved per crane per working hour); the input are six: nocranes (number of cranes), noberths (number of container berths), notugs (number of tugs), termiarea (terminal area), delaytime (delay time) and labor (proxied by the number of port authority employees).

An important input influencing port outputs is the amount of delay time, which is the difference between total berth time plus time waiting to berth, and the time between the start and finish of ship working, and is an indicator of how well working time is being used. These delays could be due to labor disputes, work practices such as meal breaks, equipment breakdown, port congestion, perceived ship problems or bad weather. In Table 1 and 2 are reported the entire dataset used for making a ranking of ports efficiency using fuzzy CCR input-oriented model discussed in the previous section. In particular, in Table 1 we have considered a unique class for all ports, while in the Table 2 we have divided ports in two classes according to the TEUs value. In fact, in table 2 we can see that, if the TEU's value is less than 1.000.000 TEUs (smaller ports) the port belongs to the first class, otherwise it belongs to the second class (larger ports).

Table 1. Ports data. Ports aggregated in a unique class

International ports	TEUs	Ship calls	Ship rate	Crane prod	N. of cranes	N. of cont. berths	N. of tugs	Term. area (m <sup>2</sup> × 10 <sup>3</sup> )	Labor (Units)	Delay time (h)	Port Classes
1. Melbourne	904618	823	20,8	14,8	16	12	6	1184.1	829	8	1
2. Hong Kong	13460343	12880	45	45	64	18	24	2198,3	800	5	1
3. Hamburg	3054320	4178	37,2	19,6	52	14	25	3030	1168	0,2	1
4. Rotterdam	4935616	5544	32	16	66	18	15	4158	981	1,7	1
5. Felixstowe	2042423	2677	56,4	23,5	29	13	3	1432	1824	0,6	1
6. Yokohama	3911927	11908	47	47	41	20	34	1823,2	472	6	1
7. Singapore	12943900	24015	40	39,3	95	17	12	2979,21	978	2,3	1
8. Keelung	2320397	3144	24	24	23	14	9	339	690	13	1
9. Sydney	695312	759	22,8	13,4	14	11	3	1124.5	635	9,5	1
10. Fremantle	202680	692	13,3	12,9	5	7	5	273	498	9	1
11. Brisbane	249439	556	21	12,5	6	3	5	474	200	5,5	1
12. Tilbury	394772	347	32,8	18,2	11	4	2	519	750	4,5	1
13. Zeebrugge	553175	1608	36,7	26,2	16	9	5	2311.1	21	1	1
14. La Spezia	871100	1045	23,9	17,1	8	7	8	270	177	3,7	1
15. TanjungPriok	1421693	3239	18	18	10	6	11	310	1513	50	1
16. Osaka	987948	2375	32	32	24	13	10	1154	1070	4	1

Table 2. Ports data. Ports divided in two classes according to value of TEUs

International ports	TEUs	Ship calls	Ship rate	Crane prod	N. of cranes	N. of cont. berths	N. of tugs	Term. area (m <sup>2</sup> )	Labor (Units)	Delay time (h)	Port Classes
10. Fremantle	202680	692	13,3	12,9	5	7	5	273000	498	9	1
11. Brisbane	249439	556	21	12,5	6	3	5	474000	200	5,5	1
12. Tilbury	394772	347	32,8	18,2	11	4	2	519000	750	4,5	1
13. Zeebrugge	553175	1608	36,7	26,2	16	9	5	2311100	21	1	1
9. Sydney	695312	759	22,8	13,4	14	11	3	1124500	635	9,5	1
14. La Spezia	871100	1045	23,9	17,1	8	7	8	270000	177	3,7	1
1. Melbourne	904618	823	20,8	14,8	16	12	6	1184100	829	8	1
16. Osaka	987948	2375	32	32	24	13	10	1154000	1070	4	1
15. TanjungPriok	1421693	3239	18	18	10	6	11	310000	1513	50	2
5. Felixstowe	2042423	2677	56,4	23,5	29	13	3	1432000	1824	0,6	2
8. Keelung	2320397	3144	24	24	23	14	9	339000	690	13	2
3. Hamburg	3054320	4178	37,2	19,6	52	14	25	3030000	1168	0,2	2
6. Yokohama	3911927	11908	47	47	41	20	34	1823250	472	6	2
4. Rotterdam	4935616	5544	32	16	66	18	15	4158000	981	1,7	2
7. Singapore	12943900	24015	40	39,3	95	17	12	2979211	978	2,3	2
2. Hong Kong	13460343	12880	45	45	64	18	24	2198300	800	5	2

The results of this study are reported in Table 3 and 4. As we can see in both cases (class one and class two of ports), we have a classification in order of growing importance where delay time is always the most important input for every future possible analysis (e.g. fuzzy DEA analysis), while the terminal area is the input less important and

then less relevant for future possible applications.

So, in this way, we can reduce significantly the number of features (inputs) used, for example, for measuring ports efficiency. Furthermore, it is worth underlining the strong influence of delay time as selected input in all kind of container ports analysis. Indeed, this is one of the goals we wanted to achieve.

Tables 3 - 4. Feature selection method results. Unique class of ports -Two classes of ports

Inputs classified in order of growing importance (ports of Table 1)	Inputs classified in order of growing importance (ports of Table 2)
Terminal area (m <sup>2</sup> )	Terminal area (m <sup>2</sup> )
Number of container berths	Number of cranes
Labor (UNITS)	Number of container berths
Number of cranes	Labor (UNITS)
Number of tugs	Number of tugs
Delay time (h)	Delay time (h)

## 5. Conclusions

In this paper the feature selection in classification based problems in transportation planning is highlighted. The proposed method simplifies the application of other models reducing the number of measurements to be taken. A numerical example in section 4 shows that feature selection method using fuzzy entropy measures is giving good results in freight transport field (container ports).

This method is applied to the same data set, and feature selection is conducted to underline which inputs used in previous studies (fuzzy DEA analysis) have more influence and importance in container ports inputs dataset.

The productivity of a container ports is influenced by a range of factors (inputs), which are removed one by one according to the amount of the input relevant information. Feature selection (FS) has been shown to be a powerful approach to deal with large number of data, by selecting relevant features from data set and, at the same time, removing irrelevant and/or redundant (highly correlated with others) features that harm the quality of the results. A good feature selection techniques should be able to identify and model the noisy and misleading features from the domain problem and help to get a minimal feature subsets, still keeping the important information present in the original data (Jensen and Sheng, 2008).

However, the acquisition of data is quite difficult, and even the combinations of independent/input variables and dependent/output variables utilized in this study have to be adjusted. In further research, this study will enlarge the number of ports and variables. The individual ports simulation model will represent the direction for future investigations.

In addition a comparative analysis with a step-wise feature selection method will be carried out.

## Acknowledgements

This study was supported by the South-East Europe (SEE) Transnational Cooperation Project GIFT – Green Intermodal Freight Transport – Project No. SEE/C/0003/3.3/X (<http://www.gift-project.eu>)

## References

- Bandemer, H. and Nather, W., 1992. Fuzzy data analysis. Kluwer Academic Publisher.
- Barros, C.P., Athanassiou, M., 2004. Efficiency in European seaports with DEA: Evidence from Greece and Portugal. *Maritime Economics and Logistics*, vol. 6, no. 2, p. 122-140.
- Bray S., Caggiani L., Dell’Orco, M., Ottomanelli M., 2014. Measuring transport systems efficiency under uncertainty by fuzzy sets theory based Data Envelopment Analysis. *Procedia-Social and Behavioral Sciences*, Vol. 111, pp. 770-779.

- Button, K., and Neiva, R., 2014. Economic Efficiency of European Air Traffic Control Systems. *Journal of Transport Economics and Policy (JTEP)*, 48(1), 65-80
- Caggiani L., Iannucci G., Ottomanelli M., Tangari L., Sassanelli, D., 2012. A fuzzy data meta training system for ranking hub container terminals. *European Transport*, vol. 51
- Chou C.C., 2010. Application of FMCDM model to selecting the hub location in the marine transportation: A case study in south-eastern Asia. *Mathematical and Computer Modelling*, Vol. 51, pp. 791-801.
- Chou C.C., 2007. A fuzzy MCDM method for solving marine transshipment container port selection problems. *Applied Mathematics and Computation*, N.186, pp. 435-444.
- Coto-Millan, P., Banzo-Pino, J., and Rodriguez-Alvarez, A., 2000. Economic Efficiency in Spanish Ports: Some Empirical Evidence. *Maritime Policy and Management*, 27 (2), pp. 169-174.
- Cullinane, K. P. B. & Song, D. W., 2006. Estimating the Relative Efficiency of European Container Ports: A Stochastic Frontier Analysis. In: Cullinane, K., and Talley, W. K. Eds. *Research in Transportation Economics, Vol 16 Port Economics*, Amsterdam, Netherlands.
- Cullinane, K. P. B. and Song, D. W., 2003. A Stochastic Frontier Model of the Productive Efficiency of Korean Container Terminals. *Applied Economics*, 35 (3)
- Cullinane, K. P. B., Song, D. W. and Gray, R., 2002. A Stochastic Frontier Model of the Efficiency of Major Container Terminals in Asia: Accessing the Influence of Administrative and Ownership Structures. *Transportation Research Part A*, 36, pp. 743-762.
- Cullinane, K., Wang T.F., 2010. The efficiency analysis of container port production using DEA panel data approaches. *OR Spectrum*, 32, pp. 717- 738.
- Cullinane, K., Wang, T.F., Song, D.W., 2006. The technical efficiency of container ports: Comparing data envelopment analysis and stochastic frontier analysis. *Journal of Transportation Research Part A*, vol. 40, p. 354-374.
- Cullinane, K.P.B., Wang, T.F., 2007. Data envelopment analysis (DEA) and improving container port efficiency. In: Brooks, M., Cullinane, K.P.B. (eds.) *Devolution, port governance and port performance*, Research in Transportation Economics, vol. 17, p. 517-566, Elsevier, Amsterdam.
- De Luca, A., & Termini, S., 1971. A definition of non-probabilistic entropy in setting of fuzzy set theory. *Information Control*, 20, 301-312.
- Estache, A., Gonzalez, M., and Trujillo, L., 2002. Efficiency Gains from Port Reform 185 and the Potential for Yardstick Competition: Lessons from Mexico. *World Development*, 30 (4), pp. 545-560.
- Huang, W.C., Teng, J.Y., Huang, M.J. and Kou, M.S., 2003. Port competitiveness evaluation by fuzzy multicriteria grade classification model. *Journal of Marine Science and Technology*, Vol. 11, No. 1, pp. 53-60.
- Itoh, H., 2002. Efficiency changes at major container ports in Japan: A window application of data envelopment analysis. *Review of Urban and Regional Development Studies*, vol. 14, no. 2, p. 133-152
- Jensen, R., Qiang, S., 2008. *Computational intelligence and feature selection: Rough and Fuzzy Approaches*. IEEE Press
- Le-Griffin, H.D., Murphy, M., 2006. Container port productivity: experiences at the ports of Los Angeles and Long Beach. *METRANS 2006 Conference Proceedings*, p. 1-21.
- Lin, L.C., Tseng, C.C., 2007. Operational performance evaluation of major container ports in the Asia-Pacific region. *Journal of Maritime Policy and Management*, vol. 34, no. 6, p. 535-551.
- Liu, Z., 1995. The Comparative Performance of Public and Private Enterprises: The Case of British Ports. *Journal of Transport Economics and Policy*, 29 (3), pp. 263-274.
- Lu, B., Huo, Y.F., Park, N.K., 2012. A study on logistics requirements potential model based on spatial economics. *Proceedings of XX. International Conference on Material Handling, Constructions and Logistics*, p. 351-356.
- Lu, B., Park, N.K., 2010. Operational performance evaluation of Korean major container terminals. *Journal of Korean Navigation and Port Research*, vol. 34, no. 9, p. 719-726
- Luukka P. 2011. Feature selection using fuzzy entropy measures with similarity classifier. *Expert Systems with Applications*, 38, 4600-4607.
- Luukka, P., and Leppalampi, T., 2006. Similarity classifier with generalized mean applied to medical data. *Computers in Biology and Medicine*, 36, 1026–1040.
- Park, N.K., Kim, J.Y., Lu, B., 2008. Critical factors for container port productivity. *Proceedings of International Association of Maritime Economists*, p.262-272.
- Park, N.K., Lu, B., 2010. A study on productivity factors of Chinese container ports. *Journal of Korean Navigation and Port Research*, vol. 34, no. 7. p. 559- 566.
- Parkash, O. M., Sharma, P. K., & Mahajan, R., 2008. New measures of weighted fuzzy entropy and their applications for the study of maximum weighted fuzzy entropy principle. *Information Sciences*, 178(11), 2389–2395.
- Roll, Y., Hayuth, Y., 1993. Port performance comparison applying data envelopment analysis (DEA). *Maritime Policy and Management*, vol. 20, no. 2, p. 153-161,
- Shannon, C. E., 1948. A mathematical theory of communication. *Bell System Technical Journal*, 379–423, 623–659.
- Sun, X., Yan, Y., & Liu, J., 2006. Econometric Analysis of Technical Efficiency of Global Container Operators, *Proceedings of the 11<sup>th</sup> International Conference of Hong Kong Society for Transportation Studies: Sustainable Transportation*, pp. 667-676.
- Tongzon, J., 2001. Efficiency Measurement of Selected Australian and Other International Ports Using Data Envelopment Analysis. *Transportation Research Part A: Policy and Practice*, Vol. 35, No. 2, (pp. 113-128).
- Tongzon, J., Heng, W., 2005. Port privatization, efficiency and competitiveness: some empirical evidence from container ports. *Transportation Research A: Policy and Practice*, vol. 39, no. 5, p. 405-424.
- Wu, J., Yan, H., Liu, J., 2009. Groups in DEA based cross-evaluation: An application to Asian container ports. *Maritime Policy & Management*, vol. 36, no. 6, p. 545-558.

- Wu, J., Yan, H., Liu, J., 2010. DEA models for identifying sensitive performance measures in container port evaluation. *Maritime Economics & Logistics*, vol. 12, no. 3, p. 215-236.
- Zadeh, L., 1965. Fuzzy sets. *Information and Control*, 8, 338–353.
- Zimmermann, H.J., 1996. *Fuzzy Set Theory and Its Applications*. Kluwer Academic Publisher, Dordrecht.