Received 6 November 2019
Revised 23 January 2020
10 February 2020
28 February 2020
Accepted 4 March 2020

The efficiency analysis of world top container ports using two-stage uncertainty DEA model and FCM

Thi Quynh Mai Pham
Department of International Transportation System, Mokpo National Maritime University, Mokpo, Republic of Korea and Department of Economics, Vietnam Maritime University, Haiphong, Viet Nam, and

Gyei Kark Park and Kyoung-Hoon Choi
Mokpo National Maritime University, Mokpo, Republic of Korea

Abstract

Purpose – The purpose of this paper is to present an integrated model to measure the operational efficiency of the top 40 container ports in the world for a five-year continuous period using a two-stage uncertainty data envelopment analysis (UDEA) combined with fuzzy C-means clustering method (FCM).

Design/methodology/approach – UDEA model is adopted for measuring the efficiency of container ports to overcome the limitation of the basic model, which is unable to handle uncertain data that are easy to meet in practice. FCM algorithm is implemented to find similar distribution efficiency scores of two stages and the cluster similar efficiency scores of container ports into various groups.

Findings – The combination of the two-stage UDEA model and the FCM algorithm provided a more comprehensive view when evaluating the performance of container ports. The UDEA results show that most of the container ports have reduced their profitability level in the second stage and most of the efficient container ports have turned into inefficient ones because of their small scale.

Originality/value – This paper proposes using the two-stage UDEA model to evaluate port efficiency based on two main aspects of productivity and profitability. Moreover, it combines DEA and FCM algorithms to offer a more comprehensive view when measuring the performance of container ports.

Keywords Container ports, Operational efficiency, Two-stage uncertainty data envelopment analysis, Fuzzy C-means clustering method

Paper type Case study

1. Introduction

Container shipping is a leapfrog development in the transport history, contributing considerably to open a new and more economical transport method. After the introduction of container shipping in the 1960s, it has generated an enormous change to the world transport industry. Recently, sea container transport has accounted for more than 52 per cent of the total maritime cargo and become an extremely significant trend in the shipping industry.

This research is a part of the project titled "SMART-Navigation project," funded by the Ministry of Oceans and Fisheries.

Moreover, the alliances become even bigger and stronger in an era of reconstruction of new alliances. To reduce the transport cost and make use of economies of scale as well, shipping companies increase the vessel size that just can get into some of the standard world ports. Therefore, the ports all over the world, especially container ports are now planning to invest in the long-term infrastructure to be able to cope suitably with the shipping alliances' strategies. Thus, strengthening operational efficiency becomes more and more important and be a key issue for the port operators to find a way to attract more customers and cargoes.

In the past, the survey method (Clark *et al.*, 2004) and parameter analysis-Stochastic frontier analysis method (Estache *et al.*, 2002) were used primarily to measure the efficiency of the container ports. However, these methods have some significant limitations. The survey method relies on the subjectivity of the participants and existing surveys of the efficiency of the container port have only been managed at a specific point in time. The parametric method requires the functional form, the production technology specification and the separation of noise and inefficiency rely on the strong assumptions on the distribution of the error term. Especially, these methods are unsuitable in the multi-input and multi-output port production systems. The other one is a non-parametric method called data envelopment analysis (DEA), which is used by several studies. The DEA is a program-based method for mathematical evaluation of the efficiency of multiple input-output and decision-making unit set (DMUs). The DEA technique is frequently used to measure the efficiency of container ports because the computation is non-parametric, which uses linear programming to determine the efficiency frontier and can handle multiple input-output.

However, the traditional DEA model has some limitations that must be considered. Because it is focused on the frontier, minor changes in data can change efficient frontiers significantly. In other words, when a data change impacting the efficient frontier and could change the state of that unit; for example, the inefficient DMU becomes efficient DMU or vice versa. Consequently, to successfully apply the basic DEA model, we must have an accurate measurement of both the inputs and outputs. Nevertheless, the observed values of the input and output data in port performance are sometimes collected over a very long period and have no exact value; for example, ship waiting time. The development of containerized cargoes is causing issues for ports as they reach the capacity limits of numerous resources, increasingly leading to port congestion. Ship waiting time reflects the qualification of the service of the container port. Its value varies by each month of each year, and it is difficult to find out the average value. On the other hand, even with the average value, which is different according to each calculation may not accurately reflect the quality of port services. Therefore, it increases the need for using a new upgrade DEA model to measure port efficiency in a way that can handle the uncertain value in the model such as Wen *et al.* (2014) and Pham *et al.* (2016).

In addition, almost all previous research studies just focused on one aspect of port operation without the multi-dimensional assessment, which was necessary to have a more comprehensive view of port performance. Therefore, this article proposes using the two-stage uncertainty DEA model to evaluate port efficiency based on two main aspects of productivity and profitability.

Moreover, the study incorporates the fuzzy C-means clustering method (FCM) to cluster the container ports into groups based on the operational efficiency results obtained by the two-stage uncertainty DEA model, which, in turn, can further assess the ability of operation and competition of container ports. Besides, it helps to detect the potential container port group whose operational efficiencies are able to be improved. After that, it is possible to investigate the reason of inefficiency by using slack-based DEA model and propose some suggestions to improve the operational efficiency of some potential container ports.

The article is organized as follows. Section 2 reviews the previous studies. Section 3 discusses an illustration of used methodologies. Section 4 describes the used data and the model results, as well as their implications. Section 5 summarizes the empirical findings with some concluding remarks.

2. Literature review

2.1 Research on port efficiency using data envelopment analysis model

Several studies have been conducted using the DEA model in relation to efficiency analysis in the existing literature. Tongzon (2001) used the DEA-CCR model for the technical efficiency analysis of 12 container ports in Australia. Turner *et al.* (2004) used the DEA model to measure infrastructure productivity and Tobit regression to examine the determinants of infrastructure productivity in American container ports. Ryoo (2005) evaluated the efficiency of the ports in Busan and Gwangyang. Wang and Cullinane (2006) also tried to measure the efficiency of 104 container terminals in Europe. Rios and Maçada (2006) did try to analyze the relative efficiency of the operations in the container terminals of Mercosur in the years of 2002, 2003 and 2004 using DEA-BCC model. Kwon (2007) measured the efficiency of 22 North-East Asian port. Park (2011) analyzed the efficiency of 11 container terminals of the Busan port and Gwangyang port. Kim and Hwang (2012) analyzed the efficiency of the major container ports in Korea and China by comparing the results of the transportation process before and after the financial crisis. Schoyen and Odeck (2013) evaluated the technical efficiency of the Norwegian container ports relative to a frontier composed of the best among the Norwegian ports and comparable that ports in Nordic and the UK ports by using DEA model. Wiśnicki *et al.* (2017) conducted nine European terminals used different handling technology to measure efficiency applying DEA model while Wanke *et al.* (2018) assessed the efficiency of six major Nigeria ports from 2003 to 2007 by applying a two-stage fuzzy-based methodology. In this way, previous studies were just based on certain physical input variables such as berth length, handling equipment, the total area and the most commonly used output variable, which is container throughput. While there are many other variables considered as affecting the port's operational efficiency such as the port handling capacity, ship call, berth productivity and the liner shipping connectivity index (LSCI) were not considered properly. Furthermore, most of the previous research studies did not mention assessing the port efficiency in different aspects and considering some common output variables to become potential input ones for another aspect.

On the other hand, in the previous studies, the input and output variables are all obvious factors. However, there are some cases when analyzing port performance, ambiguous variables also exist. To address this, some researchers have proposed several models of handling uncertain data. For example, Wen *et al.* (2014), Pham *et al.* (2016) and Lio and Liu (2017) all introduced the uncertain DEA models based on the basic DEA model and the uncertainty theory but applying them into different basic DEA model. However, the studies of Wen *et al.* (2014) and Lio and Liu (2017) only stopped with a hypothetical sample set meanwhile the research of Pham *et al.* (2016) had been tested and compared with the results got from the basic DEA software, as well as being applied directly to evaluate the efficiency of the world major container ports in

2016. Therefore, this article will continue to use the uncertainty DEA model in Pham *et al.* (2016) and expand in an extensive multi-dimensional study to analyze and evaluate many aspects of the operation of major container ports over the past five years, accompanying by an analysis of the excesses in the usage of inputs or the shortages in the production of outputs to provide some suggestions that might increase the efficiency of the container ports.

2.2 Research on classification using fuzzy C-means clustering method

FCM is widely used in various fields as a tool of classifying data, discovering groups and identifying the key properties in the underlying data. In terms of economics, Zhou (2011) attempted to analyze the influence factors of the financial market in all lines. The study indicated that effective compartmentalizing clustering measured the standard of good and bad clustering. Yin (2013) studied the clustering of supply chain units, transportation modes and work orders into different unit-transportation-work order families. This research can prove that FCM is an efficient tool to cluster data, especially with many-dimensional data sets. However, with the basic FCM method, the users must decide the number of clusters in advance, which in a certain way, the results are not regarded as objective. Therefore, to deal with the important problem in the classification process, Park *et al.* (2009) developed a new fuzzy clustering algorithm that can get the optimal number of clusters suitable to the data set. He proposed a new algorithm used by modifying the increase and the re-initialization algorithm. In this article, to consider the optimal number of clusters, the new fuzzy algorithm of Park will be applied.

3. Methodology

3.1 Uncertainty data envelopment analysis model

3.1.1 Basic data envelopment analysis model. DEA was developed by Charnes *et al.* (1978). The two most widely used DEA models are the DEA-CCR (Charnes *et al.*, 1978) and DEA-BCC (Banker *et al.*, 1984). The key difference between CCR and BCC model is that the CCR model assumes a constant return to scale (CRS) while the BCC model assumes a variable return to scale (VRS). CRS implies that a change in the amount of the input will lead to a similar change in the amount of outputs and all observed production combinations can be scaled up or down proportionally. BCC model, on the other hand, allows for VRS and is graphically represented by a piecewise linear convex frontier.

3.1.1.1 Data envelopment analysis-CCR model. Let's assume that there are n DMUs to be evaluated. Each DMU consumes varying amounts of m different inputs to generates different outputs. Specifically, $\mathrm{DMU_j}$ consumes amounts $\mathrm{X_j} = [x_{ij}]$ of inputs (i = 1, . . ., m) and produces amounts $\mathrm{Y_j} = [y_{rj}]$ of outputs (r = 1, . . ., s). The $s \times n$ matrix of output measures is denoted by Y, and the $m \times n$ matrix of input measures is denoted by X. Also, let's assume that $x_{ij} > 0$ and $y_{rj} > 0$. Consider the problem of evaluating the relative efficiency for anyone of the n DMUs, which will be identified as $\mathrm{DMU_0}$. The relative efficiency for $\mathrm{DMU_0}$ is calculated by maximizing the weighted sum of the target output. The weighted sum of the target inputs is equal to unity and the differences between the weighted sum of the outputs and the weighted sum of the inputs are smaller than zero and expressed as the equation (1):

$$\begin{cases} Max\theta = \sum_{r=1}^{s} u_r y_{rj0} \\ s.t. \sum_{i=1}^{m} v_i x_{io} = 1, \\ \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{\mathrm{ij}} \leq 0, j = 1, \ldots, n, \\ u_r \geq 0, r = 1, \ldots, s, \\ v_i \geq 0, i = 1, \ldots, m. \end{cases} \tag{1}$$

where u_r and v_i are assigned to output r and input i, respectively.

Definition 1 (efficiency). DMU_0 is CCR efficient if $\theta* = 1$ and there exists at least one optimal with $v_r* > 0$ and $u_i* > 0$ are optimal solution of equation (1). Otherwise, DMU_0 is inefficient.

3.1.1.2 Data envelopment analysis.

BCC model BCC model is expressed as equation (2):

$$\begin{cases} Max\theta = \sum_{r=1}^{s} u_r y_{rj0} - u_0 \\ s.t. \sum_{i=1}^{m} v_i x_{io} = 1, \\ \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x - u_0 \leq 0, j = 1, \ldots, n, \\ u_r \geq 0, r = 1, \ldots, s, \\ v_i \geq 0, i = 1, \ldots, m. \end{cases} \tag{2}$$

Definition 2. If DMU has CCR efficient then it also has BCC efficient.

Running the above model for each DMU, the BCC-efficiency scores are obtained. These scores are called "pure technical efficiency scores." For each DMU the CCR-efficiency score will not exceed the BCC-efficiency score. Except for u_O, which may be positive, negative or zero, all the variables of the function in equation (2) are constrained to be non-negative.

3.1.2 Uncertainty theory. Uncertainty theory was developed by Professor Baoding Liu of Tsinghua University in 2007. In this part, the basic concept of uncertain variables is as follows.

Definition 3 (Liu, 2013). An uncertain variable is a function ξ measured in the uncertainty space (Γ, L and M) for a set of real numbers and $\{\xi \in \mathrm{B}\}$ is an event for a Borel set.

Definition 4 (Liu, 2013). The uncertainty distribution of the uncertain variable ξ for any real number X is defined as $\Phi(\mathrm{x}) = \mathrm{M}\{\xi \leq \mathrm{x}\}$. For example, the uncertainty distribution of the linear uncertainty variable $\xi \sim \mathrm{L}(\mathrm{a}, \mathrm{b})$ is given by equation (3):

$$\phi(\mathrm{x}) = \begin{cases} 0 if x \leq a \\ \dfrac{x-a}{b-a} if a \leq x \leq b \\ 1 if x \geq b \end{cases} \tag{3}$$

where a and b are real numbers ($a < b$) and denoted by $\mathrm{L}(a, b)$.

Definition 5 (Liu, 2013). Regular uncertainty distributions Φ (x) and ξ can be uncertain variables. The inverse function $\Phi - 1$ (α) is then the inverse uncertainty distribution of ξ. For example, the inverse uncertainty distribution of the linear uncertain variable L (a, b) is given by equation (4).

$$\Phi^{-1}_{(\alpha)}(1-\alpha)a+\alpha b \tag{4}$$

Definition 6 (Liu, 2013). The objective function $f(x, \xi_1, \xi_2, \ldots, \xi_n)$ increases strongly for ξ_1, $\xi_2, \ldots, \xi_m$, and reduces strongly for $\xi_{m+1}, \xi_{m+2}, \ldots, \xi_n$. If the uncertainty distributions $\Phi_1, \Phi_2, \ldots, \Phi_n$ and $\xi_1, \xi_2, \ldots, \xi_n$ are independent uncertain variables, the expected objective function $E[f(x, \xi_1, \xi_2, \ldots, \xi_n)$ will be like equation (5):

$$\int_{\mathrm{df}}^{1}\left(\mathrm{x}, \phi_1^{-1}(\alpha), \cdots \phi_{\mathrm{m}}^{-1}(\alpha), \phi_{\mathrm{m+1}}^{-1}(1-\alpha), \cdots \phi_{\mathrm{n}}^{-1}(1-\alpha)\right)\mathrm{d}\,\alpha \tag{5}$$

Definition 7 (Liu, 2013). The objective function $g(x, \xi_1, \xi_2, \ldots, \xi_n)$ increases strongly for ξ_1, $\xi_2, \ldots, \xi_k$ and decreases strongly for $\xi_{k+1}, \xi_{k+2}, \ldots, \xi_n$. The uncertainty distributions $\Phi_1, \Phi_2, \ldots, \Phi_n$ and $\xi_1, \xi_2, \ldots, \xi_n$ are independent uncertain variables, and the constraint is shown in equation (6):

$$\mathrm{M}\{g(x, \xi_1, \xi_2, \ldots, \xi_n) \leq 0\} \geq \alpha \tag{6}$$

It also maintains necessary and sufficient conditions.

$$\mathrm{g}\left(x, \phi_1^{-1}(\alpha), \ldots, \phi_k^{-1}(\alpha), \phi_{k+1}^{-1}(1-\alpha), \ldots, \phi_n^{-1}(1-\alpha)\right) \leq 0 \tag{7}$$

Definition 8 (Liu, 2013). j = 1, 2, …, p and the objective function $f(x, \xi_1, \xi_2, \ldots, \xi_n)$ increase strongly for $\xi_1, \xi_2, \ldots, \xi_m$ and decrease strongly for $\xi_{m+1}, \xi_{m+2}, \ldots, \xi_n$. Besides, the objective function $g(x, \xi_1, \xi_2, \ldots, \xi_n)$ increases strongly for $\xi_1, \xi_2, \ldots, \xi_k$ and decreases strongly for $\xi_{k+1}, \xi_{k+2}, \ldots, \xi_n$. If the uncertainty distributions $\Phi_1, \Phi_2, \ldots, \Phi_n$ and $\xi_1, \xi_2, \ldots, \xi_n$ are independent uncertain variables, then the uncertain program in equation (7):

$$\begin{gathered}\min E\left[f(x, \xi_1, \xi_2, \ldots, \xi_n)\right]\\ \mathrm{s.t.M}\{g(x, \xi_1, \xi_2, \ldots, \xi_n) \leq 0\} \geq \alpha_j, j=1,2,\ldots,p\end{gathered} \tag{8}$$

Is equivalent to the crisp mathematical programming:

$$\begin{gathered}min\int_{\mathrm{df}}^{1} f\left(\mathrm{x}, \Psi_1^{-1}(\alpha), \cdots \Psi_{\mathrm{m}}^{-1}(\alpha), \Psi_{\mathrm{m+1}}^{-1}(1-\alpha), \cdots \Psi_{\mathrm{n}}^{-1}(1-\alpha)\right)\mathrm{d}\,\alpha\\ \mathrm{s.t.M}\{g(x, \xi_1, \xi_2, \ldots, \xi_n) \leq 0\} \geq \alpha_j, j=1,2,\ldots,p\end{gathered} \tag{9}$$

3.1.3 Uncertainty data envelopment analysis. Like traditional DEA, the objective of the uncertain DEA-CCR model is to maximize the efficiency level of DMU j_O subject to the constraints. The uncertain DEA model can be seen as follows:

$$
\begin{cases}
Max\,\theta = \sum_{r=1}^{s} u_r \tilde{y_{rj0}} \\
s.t. M\left(\sum_{i=1}^{m} v_i \tilde{x_0} \leq 1\right) \geq \alpha, \\
M\left\{\sum_{r=1}^{s} u_r \tilde{y_{rj}} - \sum_{i=1}^{m} v_i \tilde{x} \leq 0\right\} \geq \alpha, j = 1, \ldots n \\
u_r \geq 0, r = 1, \ldots, s, \\
v_i \geq 0, i = 1, \ldots, m.
\end{cases} \tag{10}
$$

As the uncertain measure is involved, this definition is different from the traditional efficiency definition. A natural idea is to provide a confidence level α, $\alpha \epsilon$ (0.5, 1) at which it is desired that the uncertain constraints hold. In other words, the event may not happen within $1 - \alpha$.

The model in equation (10) is an uncertain programming model, which is too complex to compute directly. The next part will give its equivalent crisp model to simplify the computation process:

$$
\text{Max} \int_0^1 \sum_{r=1}^{s} u_r \psi_{rj0}^{-1} d\alpha
$$

$$
\begin{cases}
s.t. \sum_{i=1}^{m} v_i \phi_{ij0}^{-1}(\alpha) \leq 1 \\
\sum_{r=1}^{s} u_r \psi_{rj}^{-1}(\alpha) - \sum_{i=1}^{m} v_i \phi_{ij}^{-1}(1-\alpha) \leq 0 \\
\sum_{r=1}^{s} u_r \geq 0, r = 1, \ldots, s j = 1, \ldots n \\
\sum_{i=1}^{m} v_i \geq 0, i = 1, \ldots, m
\end{cases} \tag{11}
$$

Similarly, uncertain DEA-BCC programming:

$$
Max\ \theta = \sum_{r=1}^{s} u_r \tilde{y_{rj0}} - u_0
$$

$$
\begin{cases}
s.t.\ M\left(\sum_{i=1}^{m} v_i \tilde{x_{n0}} \leq 1\right) \geq \alpha, \\
M\left\{\sum_{r=1}^{s} u_r \tilde{y_{rj}} - \sum_{i=1}^{m} v_i \tilde{x_n} - u_0 \leq 0\right\} \geq \alpha,\ j = 1, \ldots n \\
u_r \geq 0,\ r = 1, \ldots, s, \\
v_i \geq 0,\ i = 1, \ldots, m. \\
u_0\ \textit{free in sign}
\end{cases} \tag{12}
$$

Is equivalent to the crisp mathematical programming:

$$
\begin{gathered}
\text{Max}\int_0^1 \sum_{r=1}^{s}\left(u_r\psi_{rj0}^{-1}-u_0\right)d\alpha \\
\begin{cases}
s.t.\sum_{i=1}^{m} v_i\phi_{ij0}^{-1}(\alpha)\leq 1 \\
\sum_{r=1}^{s} u_r\psi_{rj}^{-1}(\alpha)-\sum_{i=1}^{m} v_i\phi_{ij}^{-1}(1-\alpha)-u_0\leq 0 \\
\sum_{r=1}^{s} u_r\geq 0, r=1,\ldots,s j=1,..n \\
\sum_{i=1}^{m} v_i\geq 0, i=1,\ldots,m \\
u_0\,\mathit{free\,in\,sign}
\end{cases}
\end{gathered}
\tag{13}
$$

3.1.4 Investigation of inefficiency causes. A useful and interesting part of the efficiency analysis is to investigate the causes of inefficiency. The following model combines CCR and BCC models into one model (Cooper *et al.*, 2011). The constraint set λ_j ⊠ 0 corresponds to the CRS (CCR) model and the constraint set $\sum_{j=1}^{n}\lambda_j=1$ corresponds to the VRS (BCC) model:

$$
\begin{gathered}
min\,\theta-\epsilon\left(\sum_{i=1}^{m} s_i^-+\sum_{r=1}^{s} s_r^+\right) \\
\begin{cases}
s.t.\sum_{j=1}^{n} x_{ij}\lambda_j+s_i^- =\theta x_{i0} i=1,2,\ldots m; \\
\sum_{i=1}^{m} y_{rj}\lambda_j-s_r^+ =y_{r0}\; r=1,2,\ldots s; \\
\lambda_j\geq 0\, j=1,2,\ldots n(CRS) \\
\sum_{j=1}^{n}\lambda_j=1, i=1,2,\ldots m(VRS)
\end{cases}
\end{gathered}
\tag{14}
$$

The ε in the objective function is called the non-archimedean, which is defined as infinitely small or less than any real positive number. The presence of ε allows a minimization over efficiency score (θ) to preempt the optimization of slacks, s_i^- s_r^+ Model 14 first obtains optimal efficiency scores (θ*) from Model 1 and calculates them, and then obtains slack values and optimizes them to achieve the efficiency frontier.

3.2 Fuzzy C-means clustering method

FCM is the most widely used fuzzy clustering algorithm. FCM was developed by Dunn in 1973 and improved by Bezdek in 1981 to handle the problem of overlapping clusters, which cannot be solved in the classical models. FCM method can use fuzzy theory to assign data to a plurality of clusters using the membership degree between 0 and 1 without belonging to a specific cluster. The principle of fuzzy clustering is to maximize similarity within a cluster, which means minimizing the sum of the distance between data and fuzzy cluster center and maximize dissimilarity between clusters, which means maximizing the distance between centers of the clusters.

Let X = $\{x_1, x_2, \ldots, x_n\}$ be the set of data in a p dimensional space where n is the number of data and p is the number of data properties.

When classifying the data into clusters, we express the non-inference of the center of each cluster and the data as the Euclidean distance as shown in equation (15):

$$d_{ik} = \|X_k - V_i\| \tag{15}$$

At that time, the center of the cluster is expressed by equation (16):

$$v_i = \frac{\sum_{k=1}^{n} (U_{ik})^m x_k}{\sum_{k=1}^{n} (U_{ik})^m} \tag{16}$$

Calculating membership of data belonging to a cluster:

$$U_{ik} = \left(\sum_{j=1}^{c} \left(\frac{x_k - v_j}{x_k - x_j} \right)^{2/(m-1)} \right)^{-1} \forall i, \forall k \tag{17}$$

where c is the number of clusters, m is a weight that determines the fuzzy level, and the larger m is, the fuzzier the partition is.

The optimal number of clusters can be determined by the number of clusters minimized in the equation (18), and the increase of clusters when the difference of values is below the threshold value, that is, the number of clusters is increased one by one:

$$\mathrm{S(c)} = \sum_{k=1}^{n} \sum_{i=1}^{c} (U_{ik})^m \left(\|x_k - v_i\|^2 - \|v_i - \bar{x}\|^2 \right) \tag{18}$$

where $\bar{x}$ is the average data. Step-by-step flowchart of optimal cluster number.

Optimal fuzzy cluster number:
Step 1: Select c ($2 \leq c < n$), m ($1 < m < \infty$) and the convergence criterion ε
Step 2: Set the initial values of the c partitioning matrixes $U^{(l)}$ as appropriate $U^{(l)}_{l=0}$
Step 3: (Repeat)
Calculate the center v of each cluster,
Update the $U^{(l+1)}_{lk}$
If $\|U^{(l+1)} - U^{(l)}\| \leq \varepsilon$, the process ends. Otherwise, the process returns to step 2
That time we can get the optimal clustering result of cluster number two (c = 2)
Step 4: Calculate the objective function
Step 5: (Repeat 2)
Increase the number of cluster c = 3, 4, 5, . . .
Repeat from step 1 to step 4 until the condition $|S(c+1) - S(c) \leq M|$ is achieved
Where M is a threshold number.

4. The efficiency analysis results and discussion

This part will evaluate the efficiency of the top container ports in the world. These leading container ports have a significant influence on container trading and maritime trade as well.

The data set for analysis includes 40 container ports that come from the list of top 100 container ports in the world in terms of annual container throughput. The analysis period adopted is in five continuous years from 2013 to 2017, and the port data were assembled in the relevant websites for the port authorities, United Nation Conference on Trade and Development, Journal of Commercial, Cargo Smart, etc.

The scientific definitions of input and output variables are important to the application of DEA model. The input and output variables should reflect the actual objectives and the process of container ports' productivity and profitability as accurately as possible. The following steps are followed in the two-stage uncertainty DEA method to measure the multi-stage efficiency of container ports.

4.1 Stage 1. Productivity

The input variables include the necessary physical facilities of container ports, which can directly influence the container handling operation such as berth length ($x1$), container cranes ($x2$), total area ($x3$), port handling capacity ($x4$) and ship waiting time ($x5$). While container berth length, total area and port handling capacity represent infrastructure and handling capacity providing an overview of port assets, the number of container crane directly influences the increase in port capacity that more containers bring increased efficiency and flexibility allowing a port to operate with more vessels simultaneously. Besides, ship waiting time, which is an indicator of how well working time is being used. It is the time from the time that a ship arrives at anchorage to the time that it arrives at the berth to load/unload cargoes or receive services and vice versa. These waiting times could be affected by the labor disputes, work practices such as equipment breakdown, port congestion, ship problems or bad weather. Therefore, the ship waiting time will be different according to each vessel or each type of vessel and cannot be collected with a certain value. The data of ship waiting time was collected indirectly through various expert's review and be regarded as an uncertain variable, which will be handled using the uncertainty DEA model in the first stage.

On the output side, the port output can be multi-dimensional depending on the objective that the port wants to achieve. In terms of productivity, the chosen output variables are container throughput ($y1$), ship call ($y2$), berth productivity ($y3$) and LSCI ($y4$). While container throughput is unquestionably the most important and widely used variable of container port that it is directly related to the need for cargo-related facilities and services, ship call represents the attractiveness of the container port and berth productivity represents the speed of handling operation at a berth, which partly shows the level and quality of port service. The last variable, LSCI captures a country's level of integration into the existing liner shipping network that means a measure of connectivity to maritime shipping and a measure of trade facilitation.

The correlation was estimated between input and output variables of the first stage to check the relationships between them. Table I shows that there is no significant correlations within the input variables except $x4$ (port handling capacity) and $x1$ (berth length) and $x2$ (container crane) but the correlations are lower than 0.8; this is evidences that these variables have positive relationships when the berth length and container crane increase, the port handling capacity also increases and vice versa. Especially, Table I also indicates the significant relationship between the input and output variables.

The efficiency values obtained by the proposed method for all the 40 container ports are reported in Table II. A value of one represents ideal efficiency. The results of the first stage show that during five continuous years, there are five Chinese ports, one Omanis port, one

Brazilian port and one Korean port continuously considered as efficient container ports in terms of productivity.

Meanwhile, from 2014 Shanghai port became an efficient container port, Singapore port seems to have lost its status of an efficient port in the past three years. The reason is that Singapore port focuses on shipping and support services with nearly 50 per cent of value-added got from this part. Thus, the container throughput recently has not increased significantly, and the container handling rate or berth productivity is not high and experience ship congestion as well because of many ship calls to the port.

The results show that European container ports had lower efficiency level than the container ports from Asia; especially Rotterdam port, which is an entry port of the whole Europe, fell into serious congestion in 2014 and efficiency score seemed to continue a low level because of the concentrated plan to develop a new big terminal. On the other side, we see an improvement in the efficiency score of the ports in Hochiminh (Vietnam), Manila (Philippines), Seattle/Tacoma (the USA) and Vancouver (Canada), comparing to the top efficient container ports.

4.2 Stage 2. Profitability

Currently, the role of some variables is transformed from output to input variables. In the second stage, the chosen input variables are container throughput ($x1$), ship call ($x2$), berth productivity ($x3$) and container handling charge ($x4$) while the output variables are yearly revenue of the ports (Y). In terms of measuring profitability, container throughput is the factor related to production efficiency, ship call is related to value-added services of the ports, berth productivity is related to the level of port services and container handling charge is one of the components related to the port's income. In this stage, there is no uncertain variable; thus, the basic DEA model will be applied.

The correlation was again estimated between input and output variables of the second stage to check the relationships between them. Table III shows that there are no significant correlations within the input variables and shows a negative correlation between $x4$ and other input variables. This indicates that an increase in container handling charges could affect port revenue-related factors.

The results of the second stage show that during five continuous years, only Shanghai ports, Santos port and New York/New Jersey port are continuously considered as efficient container ports in terms of profitability. Table IV shows that when the efficiency score of Singapore and Tianjin ports increasingly lower than other leading container ports, the other container ports in Xiamen, Hamburg and Hochiminh proved better prospects. Many other container ports have a very low-efficiency score, comparing to the first stage; especially is the situation of some container ports such as Incheon port, Port Said port and Barcelona

	x_1	x_2	x_3	x_4	x_5	y_1	y_2	y_3	y_4
x_1	1.000								
x_2	0.633	1.000							
x_3	0.139	0.090	1.000						
x_4	0.790	0.778	0.104	1.000					
x_5	0.105	0.046	0.027	0.087	1.000				
y_1	0.786	0.882	0.056	0.904	0.078	1.000			
y_2	0.287	0.341	0.024	0.403	0.070	0.436	1.000		
y_3	0.573	0.611	0.199	0.670	0.001	0.615	0.324	1.000	
y_4	0.273	0.313	0.012	0.422	0.026	0.548	0.444	0.354	1.000

Table I. Correlation matrix for inputs and outputs in the first stage

Table II. Efficiency results of world major container ports in the first stage (productivity) during 2013-2017

No.	DMU	Country	2013 CCR	2013 BCC	2014 CCR	2014 BCC	2015 CCR	2015 BCC	2016 CCR	2016 BCC	2017 CCR	2017 BCC
1	Shanghai	China	0.995	1.000	1.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	Singapore	Singapore	1.000	1.000	1.00	1.000	0.828	0.869	0.873	1.000	0.872	1.000
3	Shenzhen	China	0.993	1.000	0.98	1.000	0.972	1.000	0.956	1.000	0.972	1.000
4	Ningbo-Zhoushan	China	1.000	1.000	1.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	Busan	Korea	0.924	0.949	0.88	0.900	0.886	0.892	0.884	0.900	0.982	0.982
6	Guangzhou	China	0.982	1.000	1.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000
7	Qingdao	China	1.000	1.000	1.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000
8	Tianjin	China	0.835	0.983	0.84	1.000	0.851	1.000	0.867	1.000	0.832	1.000
9	Rotterdam	The Netherland	0.873	0.880	0.62	0.626	0.536	0.539	0.514	0.523	0.533	0.545
10	Port klang	Malaysia	0.883	0.886	0.90	0.912	0.603	0.636	0.617	0.646	0.820	0.833
11	Xiamen	China	0.803	1.000	0.91	0.946	0.842	0.879	0.842	0.842	0.850	0.852
12	Kaohsiung	Taiwan	0.768	0.818	0.82	0.823	0.786	0.805	0.834	0.839	0.788	0.790
13	Los Angeles	USA	0.629	0.676	0.67	0.703	0.669	0.680	0.749	0.762	0.719	0.724
14	Hamburg	Germany	0.595	0.598	0.48	0.479	0.489	0.492	0.467	0.482	0.494	0.511
15	Tanjung Pelepas	Malaysia	0.969	0.970	0.87	0.895	0.928	0.945	0.894	0.904	0.972	0.975
16	Laem Chabang	Thailand	0.908	0.911	0.98	0.985	0.898	0.930	0.922	0.934	0.938	0.944
17	Long Beach	USA	0.809	1.000	0.81	1.000	1.000	1.000	1.000	1.000	1.000	1.000
18	New York/New Jersey	USA	0.760	0.795	0.85	0.881	0.895	0.914	0.951	0.955	0.992	1.000
19	Colombo	Sri Lanka	0.459	0.513	0.59	0.612	0.555	0.557	0.681	0.717	0.593	0.594
20	Hochiminh	Vietnam	0.437	0.458	0.37	0.428	0.468	0.485	0.661	0.707	0.629	0.643
21	Tangjung Priok	Indonesia	0.850	0.862	0.77	0.779	0.701	0.738	0.716	0.738	0.647	0.704
22	Bremen	Germany	0.554	0.581	0.73	0.780	0.633	0.648	0.687	0.757	0.619	0.688
23	Jawaharlal Nehru	India	0.931	1.000	1.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000
24	Valencia	Spain	0.439	0.442	0.51	0.529	0.568	0.571	0.668	0.703	0.664	0.680
25	Manila	Philippine	0.443	0.457	0.50	0.549	0.644	0.677	0.991	1.000	0.967	1.000
26	Lianyungang	China	1.000	1.000	1.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000
27	Taicang	China	1.000	1.000	1.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000
28	Algeciras	Spain	1.000	1.000	1.00	1.000	1.000	1.000	0.889	0.906	0.841	0.841
29	Piraeus	Greece	0.518	0.602	0.51	0.526	0.507	0.509	0.583	0.600	0.599	0.608
30	Savannah	USA	0.563	0.582	0.66	0.700	0.688	0.715	0.748	0.766	0.763	0.775
31	Salalah	Oman	1.000	1.000	1.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000
32	Santos	Brazil	1.000	1.000	1.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000

(*continued*)

			2013		2014		2015		2016		2017	
No.	DMU	Country	CCR	BCC	CCR	BCC	CCR	BCC	CCR	BCC	CCR	BCC
33	Felixstowe	England	0.635	0.767	0.82	0.986	0.720	0.729	0.703	0.808	0.671	0.777
34	Seattle/Tacoma	USA	0.596	0.669	0.49	0.497	0.731	0.741	0.763	0.960	0.750	0.960
35	Vancouver	Canada	0.791	0.838	0.77	0.798	0.884	0.919	1.000	1.000	1.000	1.000
36	Nanjing	China	1.000	1.000	1.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000
37	Marsaxlokk	Malta	0.957	0.991	0.93	0.982	0.941	1.000	1.000	1.000	0.954	1.000
38	Incheon	Korea	1.000	1.000	1.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000
39	Port Said	Egypt	0.656	0.734	0.65	0.754	0.585	0.694	0.620	0.766	0.587	0.766
40	Barcelona	Canada	0.878	0.896	0.84	0.862	0.820	0.869	0.885	0.919	0.892	0.921

Table II.

port. The scores of CCR model of these container ports are significantly low, but the scores of BCC model equal to 1 that means these container ports are regarded as to scale inefficiency. The scale is too small to compare and compete with other major container ports like Shanghai port.

Then, FCM is used for a more comprehensive analysis of the results of the two stage-uncertainty DEA model. The chosen component variables obtain the CCR scores of the two-stage uncertainty DEA model, as there are too many container ports considered as efficient when using BCC model. After the calculation process, it is indicated that except for the year of 2016, the optimal cluster number is nearly four (c = 4) for all the years of observation.

Figure 1 presents the classifying results of FCM method. The best cluster is in blue color including the port in Shanghai, New York/New Jersey, Ningbo Zhoushan and Santos. The second cluster in orange color represents a high-efficiency score in terms of productivity but a low score in the profitability aspect, following with grey and yellow color, respectively. Additionally, it does not change considerably in cluster members from 2013 to 2015, and the difference between the best clusters with the remaining cluster was remarkable, but largely disappeared since 2016.

Based on the results above, the port cluster should be improved, firstly, is the second cluster. Because this cluster contains the container ports as efficient in terms of productivity but inefficient in terms of profitability, needs to focus on finding ways to improve the efficiency of the second aspect. As the model used in this paper remain an input-oriented model, it is significant to consider the excesses of input factors to gain higher efficiency. In fact, the factors such as container throughput, ship call and berth productivity are meaningful for container port operations, so the decline in these factors is unreasonable. Therefore, it can only affect the container handling charge of the port.

Table V show several container ports in Cluster 2, which need to reduce the container handling charge to achieve a higher level of efficiency. For example, if the container handling charges of the Jawaharlal Nehru port are reduced by 14.88 units, the efficiency score can be increased to 59.66. Remarkably, the charges of Salalah and Barcelona ports have to be reduced greatly to increase efficiency.

The remaining container ports in Cluster 2 are unable to change port efficiency by reducing container handling charges. In other words, the current charges are considered to be the optimal charges. For example, Incheon port of Korea is in this situation. Therefore, it seems that the unique way to improve efficiency is to increase the port revenue. According to recent studies, more attention should be paid to the high value-added services of container ports to increase the profitability of the ports. Container ports, such as the container ports in Korea, which concentrate only on the basic operation of the ports, are less profitable and less competitive than other container ports, which focus on developing value-added services. This study suggests that improving the profitability of the container ports should be the priority of the port authorities.

Table III. Correlation matrix for inputs and outputs in the second stage

	x_1	x_2	x_3	x_4	y
x_1	1.000				
x_2	0.436	1.000			
x_3	0.615	0.324	1.000		
x_4	−0.143	−0.094	−0.138	1.000	
y	0.508	0.188	0.356	0.259	1.000

Table IV. Efficiency results of world major container ports in the second stage (profitability) during 2013-2017

No.	DMU	Country	2013 CCR	2013 BCC	2014 CCR	2014 BCC	2015 CCR	2015 BCC	2016 CCR	2016 BCC	2017 CCR	2017 BCC
1	Shanghai	China	1.000	1.000	1.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	Singapore	Singapore	0.777	0.947	0.78	0.944	0.527	0.763	0.528	0.780	0.538	0.793
3	Shenzhen	China	0.056	0.494	0.05	0.444	0.070	0.533	0.088	0.590	0.047	0.517
4	Ningbo-Zhoushan	China	0.482	0.678	0.47	0.718	0.723	0.819	0.648	0.780	0.640	0.823
5	Busan	Korea	0.064	0.522	0.08	0.729	0.102	0.729	0.124	0.830	0.094	0.915
6	Guangzhou	China	0.008	0.610	0.00	0.432	0.006	0.472	0.017	0.520	0.015	0.513
7	Qingdao	China	0.131	0.640	0.08	0.465	0.124	0.465	0.145	0.620	0.183	0.709
8	Tianjin	China	1.000	1.000	1.00	1.000	1.000	1.000	0.785	0.900	0.651	0.825
9	Rotterdam	The Netherland	0.057	0.585	0.06	0.572	0.132	0.616	0.179	0.730	0.195	0.748
10	Port klang	Malaysia	0.148	0.760	0.20	0.807	0.364	0.838	0.203	0.740	0.189	0.753
11	Xiamen	China	0.315	0.692	0.45	0.780	0.594	0.770	0.724	0.930	0.904	1.000
12	Kaohsiung	Taiwan	0.145	0.582	0.18	0.656	0.244	0.660	0.238	0.650	0.193	0.653
13	Los Angeles	USA	0.086	0.562	0.10	0.625	0.137	0.658	0.208	0.700	0.122	0.633
14	Hamburg	Germany	0.306	0.676	0.28	0.619	0.339	0.621	0.423	0.840	0.416	0.791
15	Tanjung Pelepas	Malaysia	0.463	0.875	0.52	0.925	0.430	0.793	0.381	0.800	0.328	0.797
16	Laem Chabang	Thailand	0.053	0.664	0.06	0.720	0.091	0.711	0.086	0.770	0.095	0.810
17	Long Beach	USA	0.073	0.568	0.09	0.676	0.114	0.676	0.198	0.800	0.185	0.771
18	New York/New Jersey	USA	1.000	1.000	1.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000
19	Colombo	Sri Lanka	0.076	0.701	0.08	0.605	0.071	0.598	0.090	0.640	0.083	0.670
20	Hochiminh	Vietnam	0.389	1.000	0.36	1.000	0.673	1.000	0.736	1.000	0.577	1.000
21	Tangjung Priok	Indonesia	0.124	0.747	0.01	0.718	0.082	0.730	0.090	0.730	0.095	0.737
22	Bremen	Germany	0.137	0.598	0.13	0.560	0.177	0.596	0.185	0.590	0.168	0.596
23	Jawaharlal Nehru	India	0.094	0.918	0.10	0.926	0.146	0.895	0.160	0.930	0.135	0.897
24	Valencia	Spain	0.037	0.714	0.04	0.714	0.051	0.769	0.049	0.740	0.045	0.714
25	Manila	Philippine	0.032	0.833	0.04	0.835	0.035	0.833	0.035	0.820	0.034	0.834
26	Lianyungang	China	0.102	0.793	0.10	0.821	0.109	0.785	0.100	0.810	0.090	0.818
27	Taicang	China	0.057	1.000	0.06	1.000	0.156	0.901	0.043	0.840	0.036	0.838
28	Algeciras	Spain	0.024	0.658	0.02	0.628	0.027	0.655	0.023	0.640	0.028	0.676
29	Piraeus	Greece	0.039	0.737	0.03	0.733	0.045	0.808	0.041	0.770	0.039	0.716
30	Savannah	USA	0.159	0.905	0.16	0.874	0.161	0.815	0.217	0.890	0.213	0.903
31	Salalah	Oman	0.077	0.882	0.07	0.933	0.081	1.000	0.086	0.900	0.077	0.786
32	Santos	Brazil	1.000	1.000	1.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000

(*continued*)

No.	DMU	Country	2013		2014		2015		2016		2017	
			CCR	BCC	CCR	BCC	CCR	BCC	CCR	BCC	CCR	BCC
33	Felixstowe	England	0.180	0.758	0.17	0.800	0.184	0.800	0.231	1.000	0.191	0.970
34	Seattle/Tacoma	USA	0.257	1.000	0.25	1.000	0.353	1.000	0.334	1.000	0.300	1.000
35	Vancouver	Canada	0.101	0.932	0.10	0.956	0.117	0.950	0.137	1.000	0.121	0.923
36	Nanjing	China	0.012	0.946	0.02	1.000	0.015	1.000	0.019	0.950	0.051	0.936
37	Marsaxlokk	Malta	0.010	1.000	0.05	1.000	0.009	1.000	0.010	1.000	0.009	1.000
38	Incheon	Korea	0.051	1.000	0.05	1.000	0.085	1.000	0.087	1.000	0.081	1.000
39	Port Said	Egypt	0.011	1.000	0.02	1.000	0.032	1.000	0.049	1.000	0.081	1.000
40	Barcelona	Canada	0.094	1.000	0.09	1.000	0.120	1.000	0.096	1.000	0.080	1.000

Table IV.

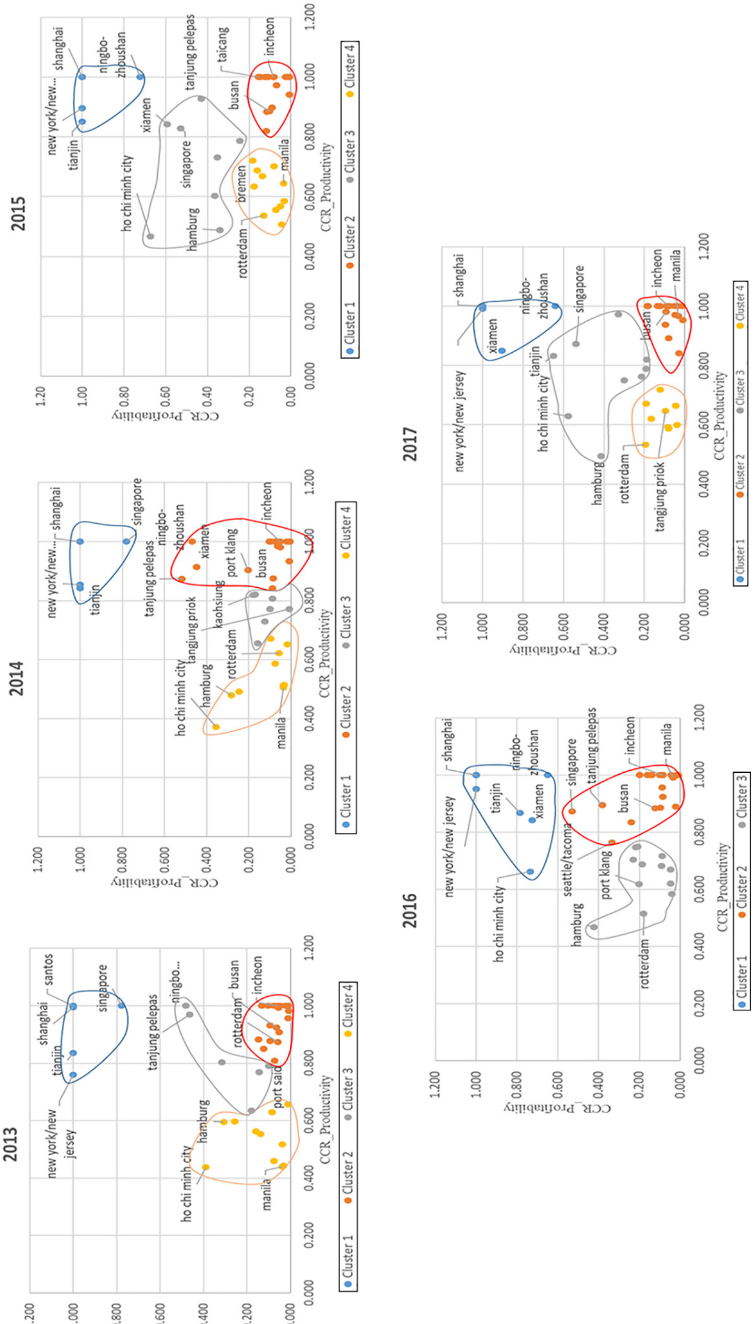

Figure 1.
The results of fuzzy C-means clustering

Port	Year	Container handling charge	Slack	Target
Jawaharlal Nehru	2013	74.53	−14.88	59.66
	2014	75.47	−11.17	64.30
	2015	85.41	0.00	85.41
	2016	68.04	−2.72	65.32
	2017	84.42	−2.56	81.86
Algeciras	2013	258.80	−80.69	178.10
	2014	262.05	−103.17	158.88
	2015	296.56	−104.87	191.68
	2016	236.26	−72.15	164.11
	2017	293.13	−132.85	160.28
Salalah	2013	93.17	−22.98	70.19
	2014	229.56	−162.26	67.30
	2015	276.16	−175.46	100.69
	2016	220.00	−147.28	72.72
	2017	272.97	−186.30	86.67
Marsaxlokk	2013	86.96	−41.00	45.96
	2014	88.05	−36.74	51.31
	2015	99.64	−12.99	86.65
	2016	79.38	−25.18	54.20
	2017	98.49	−36.03	62.47
Barcelona	2013	258.80	−197.76	61.03
	2014	262.05	−197.43	64.62
	2015	296.56	−213.15	83.41
	2016	236.26	−159.18	77.07
	2017	293.13	−184.75	108.39

Table V. Targets of container handling charge for several container ports in Cluster 2

5. Conclusion

This paper conducted research studies about the operational efficiency of the top container ports in the world using two stage-uncertainty DEA model and FCM. In the efficiency measurement section, it has properly explored the two-stage uncertainty DEA model as an alternative approach to the basic DEA model. Previous studies on DEA model have usually measured operational efficiency by using specific input and output variables. In sharp contrast, this paper has divided the operational efficiency into two stages by positively transforming the inputs in each stage, which shows the efficiencies according to each process and stage-wise role of inputs and outputs. Besides, this model considered the uncertain factor, which was unchosen in previous research studies, as one of the important factors influences the port efficiency. The results show that there are many container ports regarded as efficient container ports in the first stage but then turned into inefficiency in the second stage. The applicable results pointed out that most of the container ports have reduced their profitability level in the second stage and most of the efficient container ports turned into inefficient ones are due to their small scale.

Subsequently using the FCM method can properly classify the container ports into specific clusters. The clustering results sufficiently revealed that for the chosen period, the ports in Shanghai, Santos, New York/New Jersey are to be the best productive and profitable ports in the world and placed in the best cluster. The characteristics of the second cluster are that they are almost the best productive but unprofitable ports in the list and the two Korean container ports including Busan and Incheon are consistently clustered in this group during the five-year period.

This paper, then, carefully selects the potential port cluster for the ability analysis to change the efficiency. As most of the variables such as container throughput, ship call and berth productivity represent a significant role, it is focused typically on analyzing the possibility of changing the container handling charge. The result shows that while some port can reduce the charge to increase efficiency, some others need to seek the inevitable alternative improvement in increasing the revenue. According to recent studies, to increase sufficiently the generated revenue of the ports, the port authorities need to focus more on the high value-added services in the port. Many large container ports in the world now no longer focus on improving only container throughput but also invest much in the field of shipping supported services, auxiliary services for ships such as the ports in Singapore and Shanghai or focusing on the port-related industries in the port's hinterland like Rotterdam port. Through the above research results, it is recommended that improvement of the profitability of the container ports should be the priority of the port authorities. The proper strategic decision is awaiting to improve the efficiency of the container ports in the near future.

References

Banker, R.D., Charnes, A. and Cooper, W.W. (1984), "Some models for estimating technical and scale inefficiencies in data envelopment analysis", *Management Science*, Vol. 30 No. 9, pp. 1078-1092.

Charnes, A., Cooper, W.W. and Rhodes, E. (1978), "Measuring the efficiency of decision-making units", *European Journal of Operational Research*, Vol. 2 No. 6, pp. 429-444.

Clark, X., Dollar, D. and Micco, A. (2004), "Port efficiency, maritime transport costs, and bilateral trade", *Journal of Development Economics*, Vol. 75 No. 2, pp. 417-450.

Cooper, W.W., Seiford, L.M. and Zhu, J. (2011), *Handbook on Data Envelopment Analysis*, Springer Science + Business Media, New York, NY.

Estache, A., Gonzalez, M. and Trujillo, L. (2002), "Efficiency gains from port reform and the potential for yardstick competition: lessons from Mexico", *World Development*, Vol. 30 No. 4, pp. 545-560.

Kim, M.S. and Hwang, C.S. (2012), "A comparative study on the competitiveness of major shipping ports in Korea and China after the global financial crisis", *China and China Sinology*, Vol. 16, pp. 1-34.

Kwon, S.H. (2007), "Comparative study on the efficiency analysis of northeast Asia ports: focused on DEA model", working papers, Pukyong National University Graduate School.

Lio, W. and Liu, B. (2017), "Uncertain data envelopment analysis with imprecisely observed inputs and outputs", *Fuzzy Optimization and Decision Making*, Vol. 17 No. 3, pp. 357-373.

Liu, B. (2013), *Uncertainty Theory*, 4th Springer-Verlag, Berlin.

Ryoo, D.K. (2005), "Efficiency measurement of container terminals in Korea using DEA", *Marine Logistics Research*, Vol. 47, pp. 21-38.

Park, H.K. (2011), "The data envelopment analysis of container terminals to transshipment cargo", *Journal of Korea Port Economic Association*, Vol. 26 No. 1, pp. 1-19.

Park, G.K., Han, X., Kim, Y.K. and Oh, S.W. (2009), "Intelligent DB retrieval system for marine accidents using FCM", *Journal of Korean Institute of Intelligent Systems*, Vol. 19 No. 4, pp. 568-573.

Pham, T.Q.M., Kim, H.Y. and Lee, C.H. (2016), "불확실성 DEA모델을 이용한 컨테이너 항만의 효율성 분석 연구 [a study on the analysis of container port's efficiency using uncertainty DEA model]", 한국항만경제학회 *[Journal of Korea Port Economic Association]*, Vol. 32 No. 1, pp. 165-178.

Rios, L.R. and Maçada, A.C.G. (2006), "Analysing the relative efficiency of container terminals of Mercosul using DEA", *Maritime Economics and Logistics*, Vol. 8 No. 4, pp. 331-346.

Schoyen, H. and Odeck, J. (2013), "The technical efficiency of Norwegian container ports: a comparison to some Nordic and UK container ports using data envelopment analysis (DEA)", *Maritime Economics and Logistics*, Vol. 15 No. 2, pp. 197-221.

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. 107-122.

Turner, H., Windle, R. and Dresner, M. (2004), "North American container port productivity: 1984-1997", *Transportation Research Part E: Logistics and Transportation Review*, Vol. 40 No. 4, pp. 339-356.

Wanke, P., Nwaogbe, O.R. and Chen, Z. (2018), "Efficiency in Nigerian ports: handling imprecise data with a two-stage fuzzy approach", *Maritime Policy and Management*, Vol. 45 No. 5, pp. 699-715.

Wang, T.F. and Cullinane, K. (2006), "Chapter 23 data envelopment analysis (DEA) and improving container port efficiency", *Research in Transportation Economics*, Vol. 17, pp. 527-566.

Wen, M., Guo, L.H., Kang, R. and Yang, Y. (2014), "Data envelopment analysis with uncertain inputs and outputs", *Journal of Applied Mathematics*, Vol. 2014, pp. 1-7.

Wiśnicki, B., Chybowski, L. and Czarnecki, M. (2017), "Analysis of the efficiency of port container terminals with the use of the data envelopment analysis method of relative productivity evaluation", *Management Systems in Production Engineering*, Vol. 25 No. 1, pp. 9-15.

Yin, X.F. (2013), "A fuzzy C-means based hybrid evolutionary approach to the clustering of supply chain", *Journal of Computers and Industrial Engineering*, Vol. 66 No. 4, pp. 768-780.

Zhou, Y. (2011), "Research finance market based on fuzzy C-means clustering", paper presented at the International Conference on Computer Science and Network Technology Conference, 24-26 December 2011, Harbin, available at: https://ieeexplore.ieee.org/document/6182087 (accessed 6 April 2019).

Further reading

Park, R.K. and De, P. (2015), *'An Alternative Approach to Efficiency Measurement of Seaports', Port Management, Palgrave Readers in Economics*, Palgrave Macmillan, London.

Corresponding author

Thi Quynh Mai Pham can be contacted at: phammai090691@gmail.com