Journal of Korea Port Economic Association
Vol.40, No.1, 2024, pp.87-118.
https://doi.org/10.38121/kpea.2024.3.40.1.87

PCA-DEA 모델을 기반으로 한 중국 주요연안 항만의 운영 효율성 평가*

장해청** · 이향숙***

Evaluation of the operational efficiency of major coastal ports in China based on the PCA-DEA model

Haiqing Zhang · Hyangsook Lee

Abstract

Coastal ports play an essential role in developing a country and a city. Port efficiency is an important factor affecting port trade, and the importance of port efficiency for port performance has been recognized in previous literature. DEA (Data Envelopment Analysis) and SFA (Stochastic Frontier Analysis) are widely used in this field of research. However, these two methods are limited in selecting input and output variables. In addition, the literature studies on Chinese coastal ports mainly focus on the study of port clusters in local areas, which lacks a holistic approach and generally lacks up-to-date data. Therefore, to fill the gap in this area of research, this paper introduces a model combining principal component analysis and data envelopment analysis to analyze the operational efficiency of the top 17 coastal ports in China in terms of throughput based on the most recent data available in 2021. This paper identifies container throughput as the output variable, and 13 second indicators are selected as input variables from four primary indicators: land, capital, labor, and infrastructure. Four principal components were selected from 13 second indicators using PCA.After that, DEA (BBC) and DEA (CCR) were used to analyze the 17 ports, among which five were Shanghai, Ningbo-Zhoushan, Guangzhou, Xiamen, and Dongguan, respectively, DEA efficient, and the remaining 12 ports were non-DEA efficient. Finally, improvement directions for each port are derived, and brief suggestions are made. This paper provides some reference value for developing and constructing coastal ports in China.

Key words: Port Operation, Efficiency, China Coastal Ports, PCA, DEA

▷ 논문접수: 2024. 03. 05. ▷ 심사완료: 2024. 03. 30. ▷ 게재확정: 2024. 03. 31.

* This research was supported by the 4th Educational Training Program for the Shipping, Port and Logistics from the Ministry of Oceans and Fisheries.
** 인천대학교 동북아물류대학원 박사과정, 제1저자, haiqingjo@gmail.com
*** 인천대학교 동북아물류대학원 교수, 교신저자, hslee14@inu.ac.kr

Ⅰ. Introduction

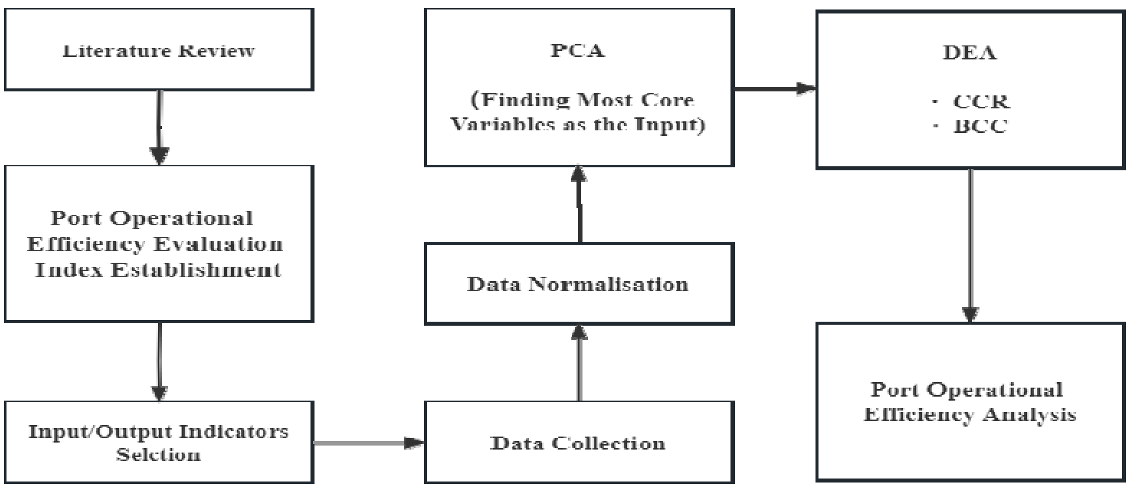
1.1 Research Background and Aims

As an important strategic frontier and infrastructure in the world trade system, the perfection and development level of ports and the efficiency of port operations significantly affect the development of the global economy(Wayne,2006).In recent years, various countries have gradually increased their investment in port construction. The scale of ports is expanding progressively, port facilities and functions are slowly improving, and the service capacity and service level of ports are also gradually improving. For China, in the global supply chain integration environment, Chinese port enterprises have to face the competition from domestic ports and bear the impact of foreign ports(Cullinane,2006). The development of the economy and the improvement of the transportation environment have driven the continuous extension of port functions and the increasingly diverse service requirements of customers, forcing the competition between ports to become white-hot.Therefore, China's port industry faces the development problem of continuing to rapidly improve its ability to cope with competition. How the port industry can improve its development efficiency becomes increasingly critical. To cope with the increasingly fierce competition and obtain more space for development, many domestic port enterprises enhance the competitiveness of ports by building many large berths equipped with advanced machinery and facilities and developing the advantages of the water level and other measures. But the investment in this hardware often needs to achieve better results and also causes idle and wasted resources. Therefore, improving the efficiency of the port enterprise is crucial, and port efficiency is a comprehensive reflection of input and output, which is fundamental to long-term development and competitiveness(Song,2014). At present, both domestic and foreign countries are actively exploring effective measures and methods to improve the efficiency of ports, and it is imperative to objectively and practically measure the current situation of China's port efficiency, and then find ways and means to improve efficiency through efficiency analysis, which also plays a crucial role in enhancing China's international trade strength and economic competitiveness. According to the World Trade Organization (WTO), affected by COVID-19, which began to sweep the world in 2020, global trade fell by 9.2% in 2020, the largest decline since World War II. Global trade volumes did not gradually recover until 2021, when the epidemic receded. During the epidemic, many countries implemented border closures and restrictions, leading to disruptions in the cargo supply chain, and ports, as the core of the maritime supply chain, were directly impacted by port cargo throughput, indirectly affecting port operational efficiency and increasing port operational costs.The port's efficiency represents the port's development level and competitiveness. The efficiency study helps the port enterprises objectively and accurately examine the current situation of port efficiency and their efficiency level, find their shortcomings and the gap between them and excellent ports, optimize the allocation of port resources, and improve the management level, which has far-reaching significance for the long-term development of the port. However, despite extensive literature review, many scholars continue to seek greater objectivity in the selection of input and output variables for DEA analysis. Furthermore, upon reviewing a substantial body of literature on the efficiency of port operations in China, it becomes apparent that while numerous studies have addressed port efficiency, they predominantly focus on specific geographical areas rather than coastal ports as a whole. Moreover, a significant portion of the existing research relies on outdated data, highlighting the need for more recent data for relevant analysis. Based on this starting point, this paper researches the efficiency of Chinese coastal ports and their influencing factors based on the latest data in 2021 use PCA-DEA model. This paper aims to clarify the current situation of China's port efficiency, find the reasons for the unsatisfactory port efficiency, and propose several countermeasures and methods for the long-term sustainable development of Chinese ports.

1.2 Research Framework

In the research framework, a literature review is conducted first. On this basis, port operation efficiency evaluation indexes are established, and then through the process of input and output index selection, data collection and normalization based on port operation efficiency evaluation indexes. Subsequently, the PCA model is introduced to find out the most core and representative variables as input variables for the next DEA (CCR) and DEA (BCC) modeling operations. Finally, the operational efficiency of the port is specifically analyzed based on the results of the DEA model.

Fig 1. Research Framework



II. Theoretical Review

2.1 Port Efficiency Overview

Because the operation of a port is a vast and complex system involving finance, labor, facilities, and many other aspects, more and more scholars have tried to use multiple indicators to measure the efficiency of the port. These include the number of berths, yard area, number of cranes, number of port professionals, port throughput, and port operating profit. For example, (Tongzon,2001) compared the efficiency of international container ports with that of Australian ports by using data such as the number of tugs, cargo throughput, number of berths, yard area, and number of cranes as indicators. (Tongzon 1995) used linear regression to study the impact of various factors on port efficiency. However, the single linear regression model is too simple, ignoring the mutual constraints among the elements and the characteristics of non-linearity that make it difficult to comprehensively portray port efficiency, not fully considering the complex non-linear relationship between port input and output indicators, and having a particular subjectivity in data selection and function setting. Therefore, many scholars began to explore non-parametric analysis methods in subsequent port efficiency-related studies and shifted from a single use of linear regression to a comprehensive evaluation study. Among them, the DEA model is favored by more and more scholars because it is used without considering the relative weights of the selected indicators and, at the same time, can meet the complex mechanism characteristics of multiple inputs and outputs in the current operation of ports. (Roll &Hayuth,1993) advocate using this approach to measure port effectiveness and demonstrate how to obtain relative effectiveness ratings for ports based on hypothetical port data. (Barros,2004) has analyzed the efficiency of the coastal ports of Greece and Portugal using the DEA method as an example.

Study on Port Efficiency in China

The research on port efficiency in China started late compared to other international scholars, but it is developing fast. Most scholars use several methods, such as index analysis, balanced scorecards, and the DEA model in port efficiency assessment. (Pang & Li,2005) pioneered using the balanced scorecard to construct a port performance evaluation system.(Chen,2004) evaluated the operational efficiency of 15 listed ports using the DEA model.(Liu &Jiang,2012) investigated the competitiveness of the top ten ports in China using the DEA evaluation model with three input indicators and two output indicators and proposed improvement countermeasures for inefficient ports. (Yang,2010) analyzed the comprehensive efficiency, scale efficiency, and pure technical efficiency of major coastal ports in Fujian Province through the DEA-CCR model to find out where the ineffectiveness lies and propose countermeasures for improvement. (Liu & Zheng ,1998) studied the influence on port efficiency using the SFA model with internal port infrastructure as the influencing factor index, and the empirical study found that the Number and length of port berths are the positive factors affecting port operation efficiency. (Huang & Yan,2004) selected port efficiency evaluation indicators according to port location conditions and hardware and software facilities and used the fuzzy comprehensive evaluation method to assess the efficiency of container ports in Southeast Asia. (Wang & Bi,2010) chose quay length and the Number of berths as input indicators and cargo throughput, container throughput, and passenger throughput as output indicators to empirically analyze the efficiency of 30 inland river ports in China, and the results showed that inland river ports have inefficiency and a serious waste of port resources.

2.2 Classification of the port efficiency

Specific port efficiency can be divided into scale, technical, and overall efficiency.

(1) Overall efficiency of the port

The overall efficiency of the port mainly refers to the optimal degree that the port can achieve for the reasonable allocation of all resources in the port (including infrastructure, human and material resources, etc.). It measures the overall operational efficiency of the port through port input and output indicators(Ghiara& Tei, 2021). Overall efficiency is mainly related to the daily operation and production activities and management and operation of the port, including the technical efficiency related to port production and the scale efficiency related to port operation and development, which can reflect the operation of a port comprehensively. In the port production and operation activities, the level of overall efficiency can measure the good or bad operation status of the port input and output to put forward constructive suggestions for the overall optimization of port efficiency development.

(2) Scale efficiency of the port

Port scale efficiency is the gap between the existing scale of port construction and the optimal scale of port efficiency. The improvement of port scale efficiency mainly comes from the investment in port expansion, which will enhance the scale efficiency to a certain extent(Pagano,2013). Generally speaking, the higher the scale efficiency of the port, the lower the operating cost of the port, and the higher the cargo throughput and revenue. Of course, when the scale efficiency reaches a certain level, the scale reward may reach a constant or reduced state. It can be seen that if the port input is unreasonable, blind investment in the expansion will lead to scale efficiency not being improved, thus leading to the waste of allocated resources, which is not conducive to the development of port efficiency. The port scale efficiency can be improved by increasing the automation degree of loading and unloading equipment, optimizing cargo flow and storage, optimizing port management and organization, increasing capacity, and improving efficiency.

(3) Technical efficiency of the port

The technical efficiency of a port is the ability to operate all the infrastructure of the port. This efficiency is based purely on the level of technology and does not consider the influence of other factors, such as port management and organization, (Schøyen&Odeck,2013). According to the production frontier surface theory, the technical efficiency of the port is the efficiency value of the port's overall efficiency after excluding the efficiency of the port's scale when the scale reward changes. Overall, port technical efficiency directly relates to the port's production technology and management level(Tongzon,2001). When the production equipment is advanced and the management level is high, it significantly enhances technical efficiency. But when the production technology is backward and cannot meet the port production operation, it will directly lead to the waste of terminal resources and hinder the improvement of port technical efficiency, thus making the port's resource allocation unable to reach the optimum, which in turn affects the subsequent development of the port(Choi,2011). Improving the technical efficiency of the port requires the comprehensive promotion of scientific and technological innovation and technology application, optimization of the internal processes of the port, and improvement of loading and unloading efficiency and cargo transportation efficiency.

The relationship between the three efficiencies of the port with variable returns to scale is shown below:

Overall efficiency = Scale efficiency × Technical efficiency

Ⅲ. Methodology

3.1 Research Modelling: PCA-DEA

The research method introduced in this paper is a combination of the most commonly used DEA method in non-parametric methods and principal component analysis, i.e., PCA-DEA. Since many indicators determine the performance of the port operation and have a strong correlation, the PCA can select the most representative indicators from many indicators. Also, DEA can evaluate indicators of different scales, does not require pre-assignment of weights or knowledge of functional relationships between them, and is very suitable for port enterprises with multiple inputs and outputs. By constructing an evaluation index system, practical factors are selected among several factors affecting port operation efficiency, and a DEA model with non-expected output is used to analyze and study the container port operation efficiency. It will be described in detail in the following.

3.2 Principle Component Analysis (PCA)

K. Pearson first proposed principal component analysis (PCA) for calculating non-random variables.PCA is a multivariate statistical analysis method that mainly uses "dimensionality reduction" to transform multiple indicators into a few unrelated new composite indicators, which we call principal components.This paper uses principal component analysis to select the evaluation index system of port operation to use fewer indicators to explain most of the indicators of the research problem. Since many factors determine the performance of the port operation and have a strong correlation, the principal component analysis can select the most representative indicators from many indicators.Suppose we study a problem with j evaluation indicators, denoted as, Principal component analysis transforms these j evaluation indicators into a linear combination of j indicators to obtain new evaluation indicators called principal components, denoted as These indicators are not related to each other, but they can cover most of the information of the original indicators. The linear equation is expressed as:

$$F_1 = a_{11}c_1 + a_{12}c_2 + \cdots + a_{1j}c_j;$$
$$F_2 = a_{21}c_1 + a_{22}c_2 + \cdots + a_{2j}c_j;$$
....
$$F_j = a_{j1}c_1 + a_{j2}c_2 + \cdots + a_{jj}c_j;$$

Usually, different economic indicators have different scales, and some differ significantly in order of magnitude. When using PCA, the different scales and orders of magnitude may cause new problems in the results. The data should be dimensionless before applying the principal component analysis to eliminate the slight influence of different magnitudes and orders of magnitude.

3.3 Data Envelopment Analysis (DEA)

The Data Envelopment Analysis (DEA) method is a new non-parametric efficiency evaluation method developed by American operations research scientists A. Charnes, W. Cooper, and E. Rhodes based on relative effectiveness.DEA is one of the most widely used methods in port efficiency research. It can evaluate indicators of different scales, does not require pre-assignment of weights or knowledge of functional relationships between them, is widely used in efficiency studies of similar decision units, and is very suitable for port enterprises with multiple inputs and outputs. (Banker,1984) argued that data envelopment analysis is more effective in port performance evaluation compared to mathematical planning methods. (Seiford&Thrall,1990) found that the efficiency frontier of DEA is very stable and suitable for small samples, which is very suitable for the efficiency evaluation of port enterprises. (Roll&Hayuth,1993) applied the DEA method to the port production sector for the first time, which theoretically solved the problem of too single evaluation index. Subsequently, DEA was applied to port efficiency evaluation in many ways, and the CCR model and BBC model, as the most traditional models of DEA methods, were the most widely used. The CCR model, which is based on the theory of a completely ideal state, is the most basic of the DEA models and is characterized by the assumption of constant returns to scale. However, in the actual analytical environment, many times the situation cannot reach the ideal state, i.e., to reach the constant returns to scale. Therefore, in order to measure port efficiency more comprehensively, we introduce another basic model of DEA, the BCC model, which is characterized by the assumption of constant returns to scale, and is closer to the actual application scenario when there is a large elasticity between factors of production. In this paper,by using CCR and BCC model, we can measure the port efficiency comprehensively from both theoretical and practical aspects. The advantage of the DEA model is that there is no need to set explicit expressions between input and output indicators, which excludes certain subjective factors and makes the evaluation conclusion highly objective. It is easier to do an in-depth analysis by building a pairwise model with the following model:

$$
\begin{aligned}
& Min\,\theta \\
& \sum_{i=1}^{n} \lambda_i x_i \le \theta x_0 \\
& \sum_{i=1}^{n} \lambda_i y_i \ge y_0 \\
& \lambda_i \ge 0, i = 1,2,\ldots,n \\
& \theta \quad \text{unbounded}
\end{aligned}
$$

The CCR-DEA model refers to technical efficiency (TE) with constant returns to scale. Technical efficiency is studied to minimize inputs at given outputs or maximize outputs at given inputs. Adding the constraint $\sum_{i=1}^{n} \lambda_i = 1$ to the CCR-DEA model is the BCC-DEA model. Denotes the technical efficiency (TE) under variable payoffs to scale. Technical efficiency measures the ability of a port to have output under specific input conditions. Technical efficiency divided by technical efficiency gives scale efficiency (SE), i.e., SE = TE/TE. Scale efficiency refers to the efficiency that comes with scale. The BCC-DEA model is as follows(Hatami&Adel,2013):

$$
\begin{aligned}
& Min\,\theta^{BCC} \\
& \sum_{i=1}^{n} \lambda_i x_i \le \theta^{BCC} x_0 \\
& \sum_{i=1}^{n} \lambda_i y_i \ge y_0, \sum_{i=1}^{n} \lambda_i = 1 \\
& \theta \quad \text{unbounded}, \quad i = 1,2,\ldots,n
\end{aligned}
$$

Ⅳ. Data Collection

4.1 Port Selection

Both coastal and inland ports are important components of a country's ports, but among them, coastal ports are the most representative of the country's ports. The level of port efficiency and the development direction of a country as a whole can be measured by assessing the efficiency of coastal ports(Li,2022). Therefore, this paper will take China's coastal ports as the research object for an efficiency study.The Ministry of Transport of the People's Republic of China released its annual government work report on January 31, 2022, announcing the yearly port cargo and container throughput data show in 〈Table1〉.According to the data, the top 20 ports in terms of container throughput are Shanghai Port, Ningbo-Zhoushan Port, Shenzhen Port, Qingdao Port, Guangzhou Port, Tianjin Port, Xiamen Port, Suzhou Port (inland), Beibuwan Port, Rizhao Port, Lianyungang Port, Yingkou Port, Dalian Port, Yantai Port, Dongguan Port, Fuzhou Port, Tangshan Port, Foshan Port (inland), Nanjing Port (inland), Jiaxing Port. Suzhou Port, Foshan Port, and Nanjing Port are Chinese inland river ports. Since the scope of this paper is China's coastal ports, these three ports are removed, and the remaining 17 ports are used as research objects.

Table 1. Top 20 Chinese ports in terms of container throughput in 2021

Ranking	Port	Throughput (million TEU)
1	Shanghai Port	4703.34
2	Ningbo-Zhoushan Port	3430.37
3	Shenzhen Port	2876.75
4	Qingdao Port	2371.19
5	Guangzhou Port	2262.88
6	Tianjin Port	2020.39
7	Xiamen Port	1204.64
8	Suzhou Port (inland)	811.72
9	Beibuwan Port	601.19
10	Yingkou Port	520.55
11	Rizhao Port	517.21
12	Lian yungang Port	503.49
13	Foshan Port (inland)	371.52
14	Dalian Port	367.28
15	Yantai Port	365.48
16	Dongguan Port	323.31
17	Fuzhou Port	345.36
18	Tangshan Port	351.49
19	Nanjing Port (inland)	311.89
20	Jiaxing Port	222.91

Data source: MOT, Port Enterprise Annual Report

4.2 Port index system setting

4.2.1 Indicator setting and data sources

When analyzing port efficiency, selecting indicators for evaluating efficiency is significant. Whether the selection of indicators is reasonable determines whether the efficiency value of the port obtained by arithmetic is representative and based. The choice of indicators should be based on the purpose of the port efficiency evaluation and some characteristics of the port before selecting the indicator system. This paper summarizes the literature of scholars as a basis for the establishment of the port efficiency indicator system in this paper is shown in Table 2.

Table 2. The summarised information on the DEA adoption by scholars abroad and domestic China

DMU	Primary Indicators	Second Indicators			
		Input (Numbers)	Authors/Year	Output (Numbers)	Method
4 Australian ports and other 12 container ports	Land; Labour; Capital	Number of cranes; Number of tugs; Number of berths; The terminal area of the ports; The amount of delay time ; The Number of the port authority employees	Tongzon (2001)	Cargo throughput; ship working rate	CCR Additive DEA.
26 Spanish Ports (Divide ports into 3groups by size)	Port activities expenditures	Labor expenditures; Depreciation expenditures; other expenditures;	Martinez-Budri a (1999)	Total cargo moved through the docks; Revenue obtained from the rent of port facilities;	DEA-BCC
India 8 Ports 1993 to 2011	Land	No. of berths Berth length No. of equipment No. of employs	Rajasekar.T & Deo.M. (2018)	Container throughput	DEA-CCR DEA-BBC
12seaports including China's4 Korea's3 and Japan's 5	Capital;Land	Import/Export by customs; GDP by regions; Berth length; Crane numbers	Jiang and Li (2009)	TEU throughputs	DEA-based
Top100 Container Ports of Cargo Systems	Infrastructures	Number of container berths; Number of quay cranes; Container berth length	Dong, Li, and Gajpal (2019)	Throughput of the container ports	CCR-DEA SBM-DEA
9ports inclouding China's8 and Korea's1	Land; Infrastructures	Berth length; Yard area; Number of quay cranes; Number of yard cranes	Zheng and Park (2016)	Container throughput	DEA-BBC DEA-CCR
26 container terminals in Induan	Infrastructure	Draft; Total quay length; Quay cranes; Yard equipment; Yard area	K.Iyera& V.Nanyam (2020)	Throughput in TEUs	DEA-BBC DEA-CCR

6 major ports in West Africa	Land; Infrastructures	Total quay length; Terminal area; Number of quayside cranes; Number of yard gantry cranes; Number of reach stackers	K. Van Dyck (2015)	Container throughput	DEA Window
27 Brazilian ports	Land; Infrastructures	Number of berths; Warehousing area; Yard area; Shipments frequency	Wanke(2013)	Solid bulk throughput; Container throughput	Network-DEA
14 coastal port enterprises and 3 inland port enterprises	Labour; Capital; Cargo uniformity	Staff number; Operational costs; Fixed assets	Sun et al.(2017)	Net profit; Cargo throughput; NOx emissions	DEA-CCR
Shanghai port	Labour; Capital	R&D expenses; Proportion of technical personnel	Xu and Xu(2021)	Business income; TEU	DEA-BBC
17 Container Terminal in China	Land; Infrastructures	Gross crane productivity; Cranes intensity; Berth depth; Berth length	Liu et al.(2022)	Calls; Moves and finish	SBM-DEA Undesirable DEA
15 container ports each of South & Middle Eastern and East Asian region	Land; Infrastructures	Number of berths; Number of cranes,; Total berth length; Berth depth	Mustafa et al.(2021)	TEUs	DEA-CCR DEA-BCC
4 ports for four Indian Ocean island countries	Land; Infrastructures	Quay cranes; Terminal area; Total quay length	Dewarlo(2019)	Cargo throughput	DEA Window
Top18 seaports in Vietnam	Land; Infrastructures	Terminal length; Equipment; Ship calls	Wang and Nguyen (2022)	Cargo throughput; TEU	DEA Malmquist

After summarizing the literature on the construction of port-related indicator systems in table

2 above, it is found that the structure of the relevant indicator systems is similar, with port cargo and container throughput as output variables. The input variables mainly involve four important components, namely land, capital, labor, and infrastructure. For example, the number of terminal berths and yard areas, regional GDP, number of employees, etc. Therefore, in this paper, integrating previous studies, the characteristics of the port itself, and the difficulty of obtaining actual data, four primary indicators of land, capital, labor, and infrastructure were identified, and 13 secondary indicators were selected from them as input indicators for PCA-DEA analysis: Staff numbe r `Proportion of technical personnel `GDP by regions `R&D expenses `Fixed assets `Operational costs `Depreciation expenditures `Yard area ` Total berth length `Total quay length `Number of berths `Number of yard gantry cranes `Number of reach stackers.

Staff number: The number of port staff reflects the labor input of the port, which directly affects the efficiency of port operation and service level. Proportion of technical personnel: The proportion of technical personnel reflects the technical and management level of the port, and a high proportion of technical personnel can help to improve the port's technological innovation ability and operational efficiency. GDP by regions: Regional GDP reflects the economic vitality and trade demand of the region in which the port is located, which is important for assessing the market size and development potential of the port. R&D expenses: R&D expenses represent the extent to which ports are investing in technological innovation and facility improvements, which are critical to enhancing their competitiveness and operational efficiency.Fixed assets: Fixed assets are an important part of a port's infrastructure and are directly related to the port's operational capacity and service level. Operational costs: Operational costs reflect the efficiency of a port's operation and management, and low costs mean higher profitability and competitive advantage. Depreciation expenditures: Depreciation expenditures reflect the useful life and value of port assets, and are important for assessing asset management and financial health.Yard area: Yard area is directly related to the storage and handling capacity of port cargo, and is one of the most important indicators for assessing the operational efficiency of a port. Total berth length: The total berth length reflects the ship docking capacity of the port, which is of great significance in assessing the port's throughput capacity and service level.Total quay length: The total quay length directly affects the ship handling efficiency and service level of the port, and is one of the important indicators to evaluate the operational efficiency of the port.

These 13 indicators (show in Table3) make up for the shortcomings of the previous scholars in the selection of indicators, and can more accurately and comprehensively cover the relevant factors for assessing port efficiency. Secondly, in the summary of output indicators, it is found that the most common output indicator used by the previous scholars is container throughput. Therefore, container throughput is identified as the output indicator in this paper.

Table 3. Port operation performance evaluation index system

DMU	Input			Output
Shanghai; Ningbo-Zhoushan; Shenzhen; Qingdao; Guangzhou ; Tianjin ; Xiamen ; Beibuwan ; Rizhao; Lianyungang; Yingkou; Dalian; Yantai; Dongguan; Fuzhou; Tangshan; Jiaxing;(17)	Primary Indicators	Second Indicators (13)	References	Container throughput (Y)
	Labor	Staff number(persons)(X1)	Sun et al. (2017)	
		Proportion of technical personnel(%)(X2)	Xu and Xu (2021)	
	Capital	GDP by regions(100 million USD)(X3)	Jiang and Li (2009)	
		R&D expenses (ten million) (X4)	Xu and Xu(2021)	
		Fixed assets (100 million RMB)(X5)	Sun et al.(2017)	
		Operational costs (100 million RMB)(X6)	Sun et al. (2017)	
		Depreciation expenditures (ten million RMB)(X7)	Martinez-Budria (1999)	
	Land	Yard area (sqm)(X8)	Wanke (2013)	
		Total berth length(m)(X9)	Mustafa et al. (2021)	
		Total quay length (m)(X10)	KIyera and V.Nanyam (2020)	
	Infrastructure	Number of berths(X11)	Mustafa et al.(2021)	
		Number of yard gantry cranes(X12)	G. K. van Dyck (2015)	
		Number of reach stackers(X13)	G. K. van Dyck (2015)	

Data sources:

Port's container throughput (Y), Staff number (X1), Proportion of technical personnel (X2), R&D expenses (X4), Fixed assets (X5), Operational costs (X6), Depreciation expenditures (X7), Yard area (X8), Total Berth length (X9), Total quay length (X10), Number of berths (X11), Number of yard gantry cranes (X12), Number of The data of GDP by regions (X3) are collected from the official websites of major ports or annual reports of port authorities . The statistical description of the sample data in this chapter is shown in appendix. The GDP of Ningbo-Zhoushan Port is the sum of Ningbo and Zhoushan cities; the GDP of Beibuwan Port is the sum of Nanning, Beihai, Qinzhou, Fangchenggang, Yulin, and Chongzuo cities in the Beibuwan Economic Zone.

Ⅴ. Case Analysis

5.1 Data Normalization

The "data normalization" is completed before the PCA model operation is started. Data normalization aims to lessen the effects of the original data's various magnitudes and units. In this paper, based on "range-normalization," the normalization technique, the range value is set between 0 and 1.

5.2 PCA

The PCA analysis was implemented by SPSS software, and the data in the appendix were entered into SPSS software to obtain the variance decomposition principal component extraction analysis table 〈Table 4〉.

Table 4. KMO and Bartlett test

KMO		0.778
Bartlett test	Approx. Chi-Square	205.314
	df	78
	p-value	0.000

The suitability of the study data for principal component analysis was first analyzed to use principal component analysis for information enrichment studies. From Table 4, it can be seen that: the KMO is 0.778, which is greater than 0.6, which meets the prerequisite requirements for principal component analysis, implying that the data can be used for principal component analysis research. Meanwhile, the data passed Bartlett's test ($p<0.05$), which indicates that the research data are suitable for principal component analysis. Following the operation of the PCA model, four principal components with the variance proportion of 65.492% (PC1), 9.420% (PC2), 7.537% (PC3), and 6.499% (PC4) are created based on the original 13 input variables, forming the cumulative variance of 88.948%(show in Table5). Their corresponding weighted variance explained (weights) are:

65.492/88.948 = 73.63%;

9.420/88.948 = 10.59%;

7.537/88.948 = 8.47%;

6.499/88.948 = 7.31%

Table 5. Total Variance Explained

PC	Eigenvalues			% of variance		
	Eigen	% of Variance	Cum. % of Variance	Eigen	% of Variance	Cum. % of Variance
1	8.514	65.492	65.492	8.514	65.492	65.492
2	1.225	9.420	74.913	1.225	9.420	74.913
3	0.980	7.537	82.449	0.980	7.537	82.449
4	0.845	6.499	88.948	0.845	6.499	88.948
5	0.420	3.232	92.180	-	-	-
6	0.331	2.544	94.724	-	-	-
7	0.261	2.004	96.728	-	-	-
8	0.151	1.162	97.890	-	-	-
9	0.103	0.789	98.679	-	-	-
10	0.076	0.585	99.264	-	-	-
11	0.064	0.489	99.752	-	-	-
12	0.019	0.145	99.897	-	-	-
13	0.013	0.103	100.000	-	-	-

Note: obtained by SPSS PCA analysis

Table 6 shows the information extraction of the principal components for the research items and the correspondence between the principal components and the research items. From Table 6, it can be seen that the communalities value of all research items is higher than 0.4, which means that there is a strong correlation between the research items and the principal components, and the principal components can extract the information effectively. After ensuring that the principal components can extract most of the information of the research items, the correspondence between the principal components and the research items is analyzed (the absolute value of the loading coefficient is greater than 0.4, which means that the item has a correspondence with the principal components).

Table 6. Loading Information of PCA

Attributes	Loadings				Communalities
	PC 1	PC 2	PC 3	PC 4	
Staff number(X1)	0.896	0.149	-0.215	0.036	0.873
Proportion of technical personnel(X2)	0.211	0.849	0.445	0.117	0.976
GDP by regions(X3)	0.860	0.001	0.044	-0.159	0.767
R&D expenses expenses (X4)	0.734	-0.159	-0.013	0.541	0.857
Fixed assets(X5)	0.850	0.062	-0.348	0.169	0.877
Operational costs(X6)	0.783	-0.271	0.378	-0.205	0.871
Depreciation expenditures(X7)	0.898	0.163	-0.279	0.104	0.923
Yard area(X8)	0.645	-0.376	0.532	0.205	0.884
Total Berth length(X9)	0.873	-0.280	-0.197	0.067	0.884
Total quay length(X10)	0.914	0.165	-0.052	-0.151	0.888
Number of berths(X11)	0.762	-0.032	-0.048	-0.601	0.945
Number of yard gantry cranes(X12)	0.928	-0.104	0.253	0.009	0.935
Number of reach stackers(X13)	0.888	0.307	-0.000	0.003	0.884

Note: obtained by SPSS PCA analysis

In order to use the principal component scores for comprehensive evaluation, it is necessary to use the "linear combination coefficient matrix" to establish the relationship equation between the principal components and the study items (based on the standardized data to establish the relationship expression), as shown in Table 7. On this basis, the equations for PC1, PC2, PC3 and PC4 can be derived as follows:

PC1Score=0.307(X1)+0.072(X2)+0.295(X3)+0.252(X4)+0.291₩(X5)+0.268(X6)+0.308(X7)+0.221(X8)+0.299(X9)+0.31(X10)+0.261(X11)+0.318(X12)+0.3049X13)

PC2Score=0.134(X1)+0.767(X2)+0.001(X3)-0.143(X4)+0.056(X5)-0.245(X6)+0.148(X7)-0.340(X8)-0.253(X9)+0.149(X10)-0.029(X11)-0.094(X12)+0.278(X13)

PC3Score=-0.217(X1)+0.450(X2)+0.044(X3)-0.013(X4)-0.352(X5)+0.382(X6)-0.282(X7)+0.538(X8)-0.200(X9)-0.052(X10)-0.048(X11)+0.255(X12)-0.000(X13)

PC4Score=0.040(X1)+0.127(X2)-0.173(X3)+0.588(X4)+0.183(X5)-0.22(X6)+0.114(X7)+0.223(X8)+0.073(X9)-0.165(X10)-0.654(X11)+0.010(X12)+0.004(X13)

Table 7.Linear combination coefficient matrix

Items	Component			
	Component 1	Component 2	Component 3	Component 4
staff number(persons)	0.307	0.134	-0.217	0.040
Proportion of technical personnel(%)	0.072	0.767	0.450	0.127
GDP by regions(100 million USD)	0.295	0.001	0.044	-0.173
R&D expenses (million USD)	0.252	-0.143	-0.013	0.588
Fixed assets(million USD)	0.291	0.056	-0.352	0.183
Operational costs(million USD)	0.268	-0.245	0.382	-0.223
Depreciation expenditures(million USD)	0.308	0.148	-0.282	0.114
Yard area(million sqm)	0.221	-0.340	0.538	0.223
Total Berth length(m)	0.299	-0.253	-0.200	0.073
Total quay length(m)	0.313	0.149	-0.052	-0.165
Number of berths	0.261	-0.029	-0.048	-0.654
Number of yard gantry cranes	0.318	-0.094	0.255	0.010
Number of reach stackers	0.304	0.278	-0.000	0.004

Besides, the composite score is calculated by accumulating the variance explained and the component scores after multiplying them. The formula for the current data is:

65.492*PC1Score+9.420*PC2Score+7.537*PC3Score + 6.499*CPC4 Score)/88.948

The final is: 0.736*PC1Score + 0.106*PC2Score 2 + 0.085*PC3Score + 0.073*PC4Score

Based on the reading of the relevant literature, it is known that the selection of principal components of PCA is decided on the basis of the cumulative information ratio of the principal components.After sorting the absolute load factor values under each PC in descending order, the following results can be obtained:

First, PC1 has a strong correlation with several indicators, including Staff number (X1) and Number of yard gantry cranes (X12), Number of reach stackers (X13), Total quay length (X10), Depreciation expenditures (X7), etc. However, in general, only in port facilities-related indicators Number of berths (X11), Number of yard gantry cranes (X12), and Number of reach stackers (X13) accounted for the largest share. Among them, the Number of yard gantry cranes(X12) has the highest correlation coefficient of 0.928. Therefore, PC1 primarily plays the role of facility indicator and can be considered a facility factor. For PC2, the Proportion of technical personnel (X2) shows the highest correlation so PC2 can be regarded as the labor factor. PC3 is highly correlated with Yard area (X8), so PC3 can be identified as the land factor. Finally, PC4 presents the highest positive correlation in R&D expenses expenses (X4), so PC4 is considered the capital factor.

Because the principal components of the PCA analysis need to be used next for DEA analysis, a mathematical function to collect the Number of input and output variables is also used(Cullinane,2006):

$$n \geq k(p+q)$$

This function can be interpreted as follows.Where n is the Number of all DMUs, p is the Number of inputs, and q is the Number of outputs. Here, the default value of k is usually 2 or 3.When this function is used in the scenario described in this paper, the Number of DMUs is fixed to 17, and the output variable is set to 1 (container throughput). If the Number of inputs is set to 4, then the value of k is determined to be 3, the formula calculates the result as 15, which is less than or equal to the value of n. If the p-value is 5, the result is greater than n, which is not in line with the logic of the formula; if the p-value is 3, the result is 12, which is less than the value of n but not maximized, the results take a smaller value, indicating that if set to 3 input indicators, there is a lack of a certain overall representation. Therefore, a p-value is the most logical choice.Considering the proportion of information content of the original variables represented by the PCs and the reasonableness of the choice of the Number of inputs revealed by the mathematical function, and considering that the cumulative variance explained by the 4 PCs has reached 88.948%, state is reached where most of the original information can be adequately interpreted.Therefore, in this paper,the Number of inputs is set to 4 in this paper, namely "Number of yard gantry cranes (X12)", "Proportion of technicians (X2)" and "Yard area (X8)", "R&D expenses costs (X4)".The Number of yard gantry cranes (X12)", "Proportion of technicians (X2)," and "Yard area (X8)", "R&D expenses costs (X4)" under facilities, labor, land, and infrastructure. Table 8 shows the information on DMU, outputs, and selected inputs by PCA on the raw data.

Table 8. The information of DMU, outputs, and selected inputs by PCA

DMU	Inputs selected by PCA (the initial parameters of the most core mapping)				Output
	PC1(Infrastructure Factor) Number of yard gantry cranes(X12)	PC2 (Labour Factor)Proportion of technical personnel(X2),%	PC3(Land Factor)Yard area(X8), million sqm	PC4(Capital Factor) R&D expenses expenses(X4),million USD	Container Throughput(Y), million TEU
Shanghai port	586	0.12	778.32	20.89	4703.34
Ningbo-Zhoushan port	317	0.09	566.93	20.04	3430.37
Shenzhen port	369	0.14	704.25	22.61	2876.75
Qingdao port	243	0.14	589.97	12.04	2371.19
Guangzhou port	210	0.29	340.68	4.17	2262.88
Tianjin port	343	0.22	429.3	21.04	2020.39
Xiamen port	274	0.09	725.53	1.02	1204.64
Beibuwan port	160	0.12	590.31	1.20	601.19
Rizhao port	112	0.26	401.57	1.48	517.21
Lian yungang port	152	0.09	570.69	13.55	503.49
Yingkou port	109	0.18	185.21	0.77	520.55
Dalian port	169	0.15	172.63	0.57	367.28
Yantai port	88	0.06	208.93	12.10	365.48
Dongguan port	87	0.02	299.58	0.45	323.31
Fuzhou port	91	0.03	107.58	2.77	345.36
Tangshan port	67	0.14	371.95	8.09	351.49
Jiaxing port	36	0.01	307.48	1.18	222.91

5.3 DEA

DEA is an established statistical technique for measuring the relative efficiency of units for which simple efficiency measures are challenging to obtain (Farrell &Charnes,1978).DEA is characterized by the ability to perform multiple inputs and outputs simultaneously. We will next analyze the two components of DEA, CCR, and BCC separately.

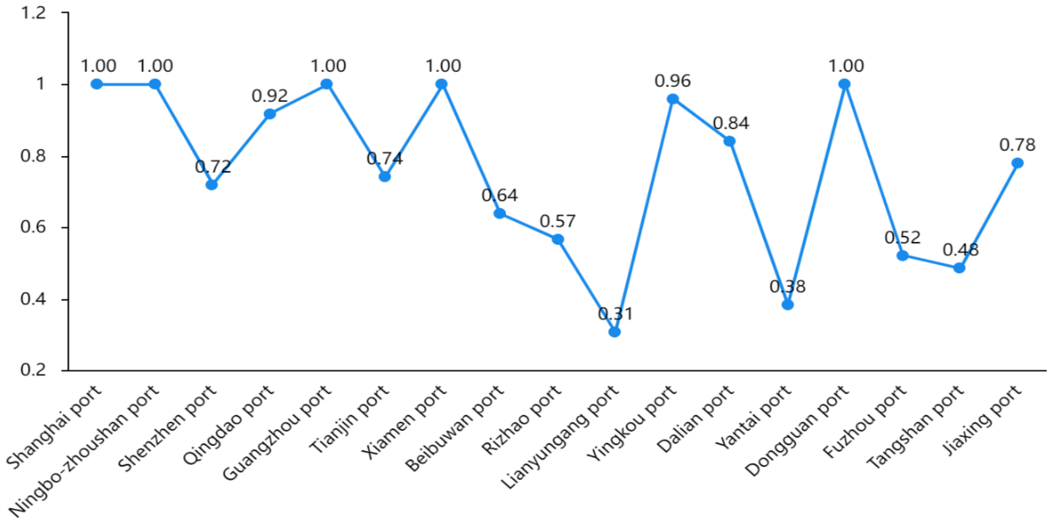
5.3.1 DEA(CCR)

The DEA (CCR) model investigates the efficiency of inputs and outputs when the payoffs of scale are constant. First, the overall efficiency (OE (θ)) is used to measure the overall efficiency of the effectiveness of the decision unit, and the value should be less than or equal to 1. The slack variable S+ means "how much output can be increased to achieve the target efficiency," and the slack variable S- means "how much input can be reduced to achieve the target efficiency." After the CCR analysis of the four PC indicators for 17 ports by SPSSAU, 〈Table 9〉 presents the analysis results.〈Fig.2〉 shows the corresponding integrated efficiency line graph. The following charts show that only "Shanghai Port," "Ningbo-Zhoushan Port," "Guangzhou Port," Xiamen Port," and "Dongguan Port" have an overall efficiency value of 1, and the slack variables S+ and S- are both 0, which achieves DEA efficiency. The overall efficiency values of the remaining 12 ports are all less than 1; thus, DEA efficiency cannot be obtained. Secondly, according to the significance of the slack variable S-, the larger the value, the more significant the gap between the input investment and the target efficiency. According to the graph, the most critical gap is in "Jiaxing Port," with a value of 199.423, followed by "Qingdao Port" and "Beibuwan Port," with a value of 159.997 and 130.835, respectively. The smallest gap is 8.504 for the "Fuzhou port."

Table 9. DEA (CCR) Calculation Results Effectiveness Analysis

DMU	Overall Efficiency (OE (θ))	Slackening variables S-	Slackening variables S+	Efficiency
Shanghai port	1.000	0.000	0.000	DEA Efficiency
Ningbo-Zhoushan port	1.000	0.000	0.000	DEA Efficiency
Guangzhou port	1.000	0.000	0.000	DEA Efficiency
Xiamen port	1.000	0.000	0.000	DEA Efficiency
Dongguan port	1.000	0.000	0.000	DEA Efficiency
Tianjin port	0.741	75.067	0.000	No DEA Efficiency
Qingdao port	0.919	159.997	0.000	No DEA Efficiency
Beibuwan port	0.638	130.835	0.000	No DEA Efficiency
Rizhao port	0.567	97.883	0.000	No DEA Efficiency
Lianyungang port	0.306	92.699	0.000	No DEA Efficiency
Yingkou port	0.961	26.848	0.000	No DEA Efficiency
Dalian port	0.840	81.247	0.000	No DEA Efficiency
Yantai port	0.384	22.306	0.000	No DEA Efficiency
Shenzhen port	0.721	33.930	0.000	No DEA Efficiency
Fuzhou port	0.522	8.504	0.000	No DEA Efficiency
Tangshan port	0.485	124.156	0.000	No DEA Efficiency
Jiaxing port	0.779	199.423	0.000	No DEA Efficiency

Fig 2. Line chart of Overall Efficiency (OE (θ)) calculated by DEA (CCR) model



Note: obtained by SPSS DEA (CCR) analysis

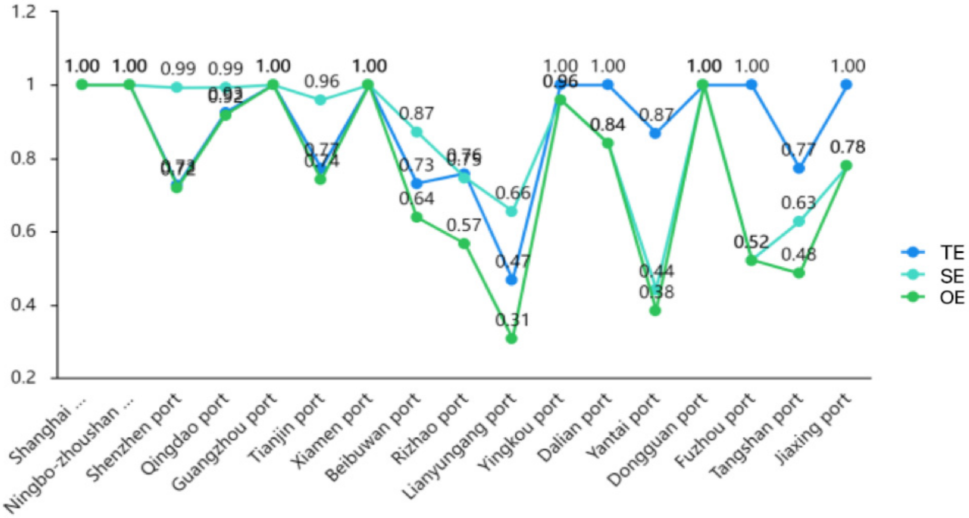
5.3.2 DEA(BCC)

The DEA (CCR) model investigates the efficiency of inputs and outputs when the payoffs of scale are variable. First, the BCC model divides the overall efficiency (OE (θ)) into two types of efficiency: technical efficiency (TE) and scale efficiency (SE). TE reflects the efficiency due to technological factors, and the value is equal to 1, indicating the rational use of factors. Conversely, when the value is less than 1, it suggests that the technical efficiency of the factor has yet to be improved.SE measures the efficiency of scale; when the value is equal to 1, it means that the returns to scale are constant (optimal state); when the value is less than 1, it means that the returns to scale are increasing (due to the small scale can be expanded to increase the benefits) if the value is greater than 1, it means that the returns to scale are decreasing (due to the large scale can be reduced to increase the benefits). Second, the overall efficiency (OE) reflects the efficiency of the decision-making unit (DMU) elements. The value equals the multiplication of technical and scale efficiency, and the value should be less than or equal to 1. Also, as in the CCR analysis, the slack variable S- means "the target efficiency is achieved by reducing the number of inputs"; the slack variable S+ means "the target efficiency is achieved by increasing the number of outputs."The TE, SE, and OE results calculated by the BCC model are shown in 〈Table 10〉, and the corresponding line graphs are shown in 〈Fig3〉. The chart shows that among the 17 ports, only "Shanghai Port," "Ningbo-Zhoushan Port," "Guangzhou Port, "Xiamen Port" and "Dongguan Port" are DEA efficient, while the remaining 13 ports are non-DEA efficient. It should be noted that although the scale efficiency (SE) and overall efficiency (OE) of the remaining 12 ports are not optimal, "Yingkou Port," "Dalian Port," "Fuzhou Port" and "Jiaxing Port" are in the optimal state of technical efficiency. In addition, the slack variable S- results for this BCC analysis are the same as those of the CCR model.

Table 10. DEA (BCC) Calculation Result

DMU	Technical efficiency (TE.)	Scale efficiency (SE (k))	Overall efficiency (OE(θ))	Slackening variables S-	Slackening variables S+	Efficiency
Shanghai port	1.000	1.000	1.000	0.000	0.000	DEA Efficiency
Ningbo-Zhoushan port	1.000	1.000	1.000	0.000	0.000	DEA Efficiency
Guangzhou port	1.000	1.000	1.000	0.000	0.000	DEA Efficiency
Xiamen port	1.000	1.000	1.000	0.000	0.000	DEA Efficiency
Dongguan port	1.000	1.000	1.000	0.000	0.000	DEA Efficiency
Tianjin port	0.773	0.959	0.741	75.067	0.000	No DEA Efficiency
Qingdao port	0.925	0.994	0.919	159.997	0.000	No DEA Efficiency
Beibuwan port	0.731	0.873	0.638	130.835	0.000	No DEA Efficiency
Rizhao port	0.758	0.747	0.567	97.883	0.000	No DEA Efficiency
Lianyungang port	0.467	0.655	0.306	92.699	0.000	No DEA Efficiency
Yingkou port	1.000	0.961	0.961	26.848	0.000	No DEA Efficiency
Dalian port	1.000	0.840	0.840	81.247	0.000	No DEA Efficiency
Yantai port	0.869	0.442	0.384	22.306	0.000	No DEA Efficiency
Shenzhen port	0.726	0.992	0.721	33.930	0.000	No DEA Efficiency
Fuzhou port	1.000	0.522	0.522	8.504	0.000	No DEA Efficiency
Tangshan port	0.773	0.627	0.485	124.156	0.000	No DEA Efficiency
Jiaxing port	1.000	0.779	0.779	199.423	0.000	No DEA Efficiency

Fig 3. Line chart of TE,SE and OE calculated by DEA(BCC)model



Note: obtained by SPSS DEA (BCC) analysis

5.4 Efficiency Result Analysis and Discussion

In general, according to the results of DEA(BCC) and DEA(CCR), it is known that the BCC analysis additionally divides the combined efficiency into technical efficiency (TE) and scale efficiency (SE), which are analyzed separately. Thus, the results for the combined efficiency (OE) and the slack variable S- are the same, despite the differences between the constant and variable returns to scale. Specifically, "Shanghai Port," "Ningbo-Zhoushan Port," "Guangzhou Port, "Xiamen Port," and "Dongguan Port" are DEA efficient, while the remaining 12 ports are non-DEA efficient. This is because the reward of scale analysis (ROS) and the input redundancy analysis will give a more precise and detailed picture of port efficiency. Therefore, to further analyze the port efficiency, we will discuss these two analyses next.

5.4.1 Return-to-Scale (ROS)

Return-to-Scale (ROS) refers to the extent to which an increase in factors of production (e.g., labor and capital) results in a corresponding rise in output (e.g., goods and services). In other words, ROS measures the extent to which an increase in factors of production affects output. Since port operations are by nature dynamic and complex (Valentine&Gray, 2001), it isn't easy to assume that ROS will always remain constant when assessing the actual operational efficiency of a port. Therefore, the discussion of ROS in this paper is based on the variable case, according to the BBC model assuming variable payoffs to scale and regarding the payoff coefficient of scale (lambda value) to investigate the payoffs to scale. There are three situations: firstly, the payoff coefficient is equal to 1, which means that the returns to scale are constant (optimal); secondly, the payoff coefficient is less than 1, which means that the returns to scale are increasing (due to small scale, the benefits of increasing scale can be increased); and finally, the payoff coefficient is greater than 1, which means that the returns to scale are decreasing (due to large scale, the benefits of increasing scale can be reduced).According to ⟨Table 11⟩, "Shanghai port" ˋ"Ningbo-Zhoushan port" ˋ"Guangzhou port " ˋ"Shanghai port" ˋ "Ningbo-Zhoushan port" ˋ"Guangzhou port" ˋ "Xiamen port" ˋ"Dongguan port" has a scale payoff coefficient of 1, and the returns to scale are unchanged. The remaining 12 ports have scale payoff coefficients of less than one and increasing returns to scale. In other words, except for the five ports with constant returns to the scale mentioned above, the remaining 12 ports must be expanded to improve the target operating efficiency.

Table 11. ROS type for 17 Ports in China

DMU	ROS Coefficient	Type
Shanghai port	1.000	Constant
Ningbo-Zhoushan port	1.000	Constant
Guangzhou port	1.000	Constant
Xiamen port	1.000	Constant
Dongguan port	1.000	Constant
Tianjin port	0.751	Increase
Qingdao port	0.780	Increase
Beibuwan port	0.399	Increase
Rizhao port	0.273	Increase
Lianyungang port	0.147	Increase
Yingkou port	0.316	Increase
Dalian port	0.239	Increase
Yantai port	0.107	Increase
Shenzhen port	0.858	Increase
Fuzhou port	0.093	Increase
Tangshan port	0.102	Increase
Jiaxing port	0.056	Increase

Note: obtained by SPSS DEA (ROS) analysis

5.4.2 Investment Redundancy Rate

The BCC and CCR models were used to analyze the operational efficiency of 17 Chinese coastal ports. The slack variable S- results are used to determine the direction of improvement for the ports because the meaning of the slack variable S- is "How much input reduction is needed to achieve the target efficiency."In other words, the port is not achieving its target efficiency due to "over-investment". This is because "over-investment" means that the port is over-invested and wasteful in some aspects of its operations.Next, Investment redundancy rate is introduced to find better ways to reduce inputs. Because the investment redundancy rate analysis can give details of the specific amount of input reduction needed to achieve the DEA target efficiency in each of the four input indicators for each of the 12 ports that did not achieve DEA efficiency because the value of the slack variable S- was greater than 0.Investment redundancy rate refers to the ratio of 'over-investment' to 'invested,' where a more significant value means more 'over-investment.'〈Table 12〉 is the slack variable S- and investment redundancy rate data for 17 ports. In this paper, the Proportion of technical personnel (X2), R&D expenses expenses (X4), Yard area (X8), and Number of yard gantry cranes (X12) are four input indicators.First, in terms of labor factor, the investment redundancy rate is 8.8% for Shenzhen Port, 14.9% for Beibu Gulf Port, 33.5% for Rizhao Port, 15.9% for Lianyungang Port, 65.6% for Yingkou Port, 22.4% for Yantai Port, 59.7% for Dalian Port, and 41.9% for Tangshan Port. The eight ports need to reduce the corresponding input units of Shenzhen port 0.012, Beibuwan port 0.018, Rizhao port 0.087, Lianyungang port 0.014, Yingkou port 0.118, Dalian port 0.09, Yantai port 0.013, Tangshan port 0.059 to achieve the desired operational efficiency. Regarding the capital factor, the investment redundancy rate is 38.5% for Tianjin Port, 8.9% for Lianyungang Port, 20.7% for Yantai Port, and 20.1% for Tangshan Port. The four ports must reduce the corresponding input units to 8.099 for Tianjin Port, 1.206 for Lianyungang Port, 2.509 for Yantai Port, and 1.869 for Tangshan Port, respectively, to achieve the desired operational efficiency. In terms of land factor, the investment redundancy rate is 4.8% for Shenzhen Port, 27.1% for Qingdao Port, 22.2% for Beibuwan Port, 22.4% for Rizhao Port, 16% for Lianyungang Port, 9.5% for Yantai Port, 32.9% for Tangshan Port, and 64.9% for Jiaxing Port. The eight ports need to reduce the corresponding input units of Shenzhen Port 33.918, Qingdao Port 159.997, Beibuwan Port 130.817, Rizhao Port 97.796, Lianyungang Port 91.478, Yantai Port 19.784, Tangshan Port 122.229, Jiaxing Port 199.423. Regarding facility factors, the investment redundancy rate is 19.5% for Tianjin Port, 24.5% for Yingkou Port, 48% for Dalian Port, and 9.3% for Fuzhou Port. Then these four ports should reduce the corresponding input units by 66.968 for Tianjin port, 26.73 for Yingkou port, 81.158 for Dalian port, and 8.504 for Fuzhou port, respectively. Finally, by adding up all the input elements of the same port, the results are as follows: Shenzhen port needs to reduce the corresponding input by 33.93, Tianjin port by 75.067, Qingdao port by 159.997, Beibuwan port by 130.835, Rizhao port by 97.883, Lianyungang port by 92.699, Yingkou port by 26.848, Dalian port by 81.247, Yantai port by 22.306, Fuzhou port by 8.504, Tangshan port by 124.156, Jiaxing port by 199.423. The Port of Yantai is 22.306, the Port of Fuzhou is 8.504, the Port of Tangshan is 124.156, and the Port of Jiaxing is 199.423.

Table 12. Date on Slack Variable S− and Input Redundancy rate

DMU	Slack Variable S-					Investment Redundancy Rate			
	Proportion of technical personnel(X2)	R&D expenses (X4)	Yard area (X8)	Number of yard gantry cranes(X12)	Sum	Proportion of technical personnel(X2)	R&D expenses (X4)	Yard area (X8)	Number of yard gantry cranes (X12)
Shanghai port	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Ningbo-Zhoushan port	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Guangzhou port	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Xiamen port	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Dongguan port	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Tianjin port	0.000	8.099	0.000	66.968	75.067	0.000	0.385	00	0.195
Qingdao port	0.000	0.000	159.997	0.000	159.997	0.000	0.000	0.271	0.000
Beibuwan port	0.018	0.000	130.817	0.000	130.835	0.149	0.000	0.222	0.000
Rizhao port	0.087	0.000	97.796	0.000	97.883	0.335	0.000	0.244	0.000
Lian yungang port	0.014	1.206	91.478	0.000	92.699	0.159	0.089	0.160	0.000
Yingkou port	0.118	0.000	0.000	26.730	26.848	0.656	0.000	0.000	0.245
Dalian port	0.090	0.000	0.000	81.158	81.247	0.597	0.000	0.000	0.480
Yantai port	0.013	2.509	19.784	0.000	22.306	0.224	0.207	0.095	0.000
Shenzhen port	0.012	0.000	33.918	0.000	33.930	0.088	0.000	0.048	0.000
Fuzhou port	0.000	0.000	0.000	8.504	8.504	0.000	0.000	0.000	0.093
Tangshan port	0.059	1.869	122.229	0.000	124.156	0.419	0.231	0.329	0.000
Jiaxing port	0.000	0.000	199.423	0.000	199.423	0.000	0.000	0.649	0.000

Note: obtained by SPSSAU DEA (Investment Redundancy) analysis

VI. Conclusions and Recommendations

6.1 Conclusions

With the development of economic globalization, ports play an increasingly critical role in the global flow of goods and trade between countries. The port is a complex system essential to the international logistics network. Although China's ports have been developing rapidly in recent years, especially the ports of Shanghai and Ningbo-Zhoushan, which have been occupying the top three positions of global ports for years, there are still many seaports along China's coast, and the overall level of China's ports still has much room for improvement. The operational efficiency of ports directly indicates competitiveness. It is one of the critical factors affecting port performance and competitiveness and has also attracted the attention of many scholars. Many scholars use two of the most common frontier models, DEA and SFA, because of their practicality in terms of efficiency. However, after summarizing much-related literature, most scholars still need more objectivity in selecting input and output variables in DEA. Secondly, after translating a large amount of literature on port operation efficiency in China, it was found that although there are many port efficiency-related studies, they mainly focus on analyzing port efficiency in a single or small area. Only a few studies specifically focus on the operational efficiency of coastal ports. In addition, a large amount of the research literature has old data, and there needs to be more literature that uses data from recent years for relevant analysis. Therefore, to fill the research gap in this area, this paper introduces a combination of PCA and DEA methods to analyze the suitable port efficiency using the latest data from 2021 and the top 17 coastal ports in China regarding container throughput. When conducting PCA-DEA analysis, the object needs to be selected first. In this paper, 17 ports along the coast of China are used as DMUs for analysis. Secondly, input and output indicators need to be selected, and container throughput is identified as the output indicator in this paper. In terms of input indicators, this paper summarizes the literature of several related studies, concludes four major scopes, namely, labor, land, capital, and infrastructure, and selects 13 indicators around these four scopes.

The 13 indicators were analyzed by principal component analysis. Four core input indicators were chosen from them. "Proportion of technical personnel (X2)", "R&D expenses (X4)", "Yard area (X8)", and "Number of yard gantry cranes (X12)" are the four indicators that are loaded with the highest information variables, representing labor, capital, land, and infrastructure, respectively. After identifying the core indicators, we moved on to the DEA (BBC) and DEA (CCR) analyses. The overall efficiency (OE) results were consistent despite the constant and variable payoffs of scale for BBC and CCR, respectively. From the results of both analyses, it is jointly shown that five ports—Shanghai Port, Ningbo-Zhoushan Port, Guangzhou Port, Xiamen Port, and Dongguan Port are DEA effective. At the same time, the lambda value is equal to 1, which keeps the ROS constant. In contrast, the remaining 12 ports are non-DEA efficient with lambda values less than 1, maintaining the ROS increasing state, and need to increase efficiency by scaling up. In addition, the slack variable (S-) and the investment redundancy rate can be used to derive how much input reduction is needed to achieve the desired efficiency for the remaining 12 non-DEA efficient ports. This is because according to the definition of slack variable (S-), the target efficiency is achieved when the value of the slack variable (S-) is 0; when the value of the slack variable (S-) is greater than 0, it means that the port has excess inputs and needs to reduce the inputs to achieve the target efficiency. Further, the concept of input redundancy is introduced to explain the specific indicators that need to be reduced when the port needs to reduce inputs to achieve the target efficiency. Finally, the redundancy analysis shows that except for the five ports that achieve DEA, the remaining 12 ports need to reduce the inputs in the four input indicators to different degrees to achieve the target efficiency. Input redundancy provides a direction for improvement in port efficiency. This paper analyzes port efficiency through a port operation efficiency evaluation model with a hybrid PCA-DEA approach. It proposes improvement directions for developing China's coastal ports, providing some guidance for developing China's port industry.

In a related study with similar results, (Cullinane &Wang,2006) used DEA to analyze sample data from the world's top container ports and found that there was a significant amount of waste in container port production, while the sample ports showed a mixture of diminishing, increasing, and constant returns to scale in terms of efficiency. (Cullinane&Wang,2010) applied DEA analysis to 25 major container ports and revealed considerable waste in the production of container ports. It also provides a basis for assessing the competitiveness of container ports, benchmarking best practices, and identifying specific sources or causes of inefficiencies. Both studies and this paper found significant waste in the production of the ports, thus affecting the ports' ability to achieve their target efficiencies. At the same time, port efficiency shows three states of decreasing, increasing, and constant revenue size, on the basis of which this paper analyzes the directions for improvement. (Pjevevi, 2012) measured and analyzed the efficiency of ports on the Danube River through DEA methodology to identify sources of inefficiency and to develop recommendations to improve the services of these ports and their operations. Total warehouse area, quay length, number of cranes, and port throughput were found to be measures of port efficiency in Serbian river ports. This paper uses PCA analysis to identify similar indicators that affect the efficiency of each port, namely, the proportion of technical personnel, R&D expenses, yard area, Number of yard gantry cranes, which are derived from four different dimensions, a more comprehensive The four indicators are from four different dimensions, which is a more comprehensive measure of port efficiency.

However, this paper still has limitations in some aspects. Firstly, this paper only uses two models, DEA (BCC) and DEA (CCR), and does not use other methods, such as Malmquist, that can sample different years, thus lacking inter-decade comparisons and change analysis. Secondly, this paper only studies the main 17 seaports along the coast of China, which has a long coastline and many ports distributed along the coast. The sample size selected in this paper needs to be bigger, which may affect the model's accuracy. Third, the impact of COVID-19 leads to the possibility of varying degrees of fluctuations in the various conditions affecting the efficiency of port operations, which in turn affects the measurement of port efficiency. However, port operations consist of a large, dynamic system that requires the cooperation of many parties to accomplish. The effect of COVID-19 on port operational efficiency can hardly be reflected uniformly in all aspects. Due to the limitations of the study period and data collection, this paper is unable to fully reflect the impact of the 2021 epidemic on port efficiency. Although the study in this paper was not able to cover the impact of the 2021 epidemic, it recognizes the importance of this factor to a full understanding of port efficiency and encourages future studies to explore and analyze this aspect in greater depth. Fourth, the study in this paper covers 17 seaports along the coast of China, with a large number of ports and detailed data on specific indicators. The dataset mainly comes from reliable official data and public information, and there are different degrees of difficulties in obtaining the data due to the limitations of China's relevant policies. To ensure the authenticity of the data, the latest data are only updated to 2021.

6.2 Recommendations

According to the above analysis, among the 17 ports, except for five ports, namely Shanghai Port, Ningbo-Zhoushan Port, Guangzhou Port, Xiamen Port, and Dongguan Port, the remaining 12 ports have room for improvement in port operation efficiency. Therefore, because of the above analysis, some efficiency improvement suggestions are made for these ports: 1. Each port should increase the construction of route flights, promote the development of foreign trade business based on the national development strategy, and vigorously develop domestic trade business. 2. Insist on innovation as the core, enhance the development of new dynamic energy, improve the degree of comprehensive intelligence, and expand the depth and breadth of port terminal intelligence, digitalization, and automation. 3. Dalian and Yingkou ports need to improve port operations and decision-making. Strengthen the connection between port clusters, improve the overall integrated technical efficiency, and enhance port construction. Fuzhou Port, Tianjin Port, Yingkou Port, and Dalian Port should complete the renewal of port facilities and equipment, berth renovation, and merger as soon as possible to improve the technical efficiency of the ports. At the same time, it is suggested that the ports not rely on more resources for blind expansion but instead use existing resources to consolidate the foundation. For Tangshan Port, build an intelligent port and improve port technology by accelerating the renewal of facilities and equipment and creating a big data platform. Shenzhen Port, Beibuwan Port, Rizhao Port, and Lianyungang Port should make more appropriate deployments of technical personnel, maximize the rational use of scientific and technological resources, and improve comprehensive technical efficiency.

Reference

Talley, W. K. (2006). *Port performance: an economics perspective. Research in Transportation Economics,* 17, 499-516.

Cullinane, K., & Wang, T. F. (2006). Port governance in China. *Research in Transportation Economics,* 17, 331-356.

Song, L., & van Geenhuizen, M. (2014). Port infrastructure investment and regional economic growth in China: Panel evidence in port regions and provinces. *Transport Policy,* 36, 173-183.

(WTO)World Trade Organization(2022) *World Trade Report 2021* https://www.wto.org/english/res_e/reser_e/wtr_e.htm

(MOT) Ministry of Transport of the People's Republic of China. *Annual report 2021* https://xxgk.mot.gov.cn/2020/ndbg/2021/

Shanghai International Port (Group) Co., Ltd. (SIPG).

Port Annual Report 2021 *https://www.portshanghai.com.cn/xxpl/index.jhtm*
l?index=2

Ningbo Zhoushan Port Group Co. Ltd. *Port Annual Report 2021* https://www.zjseaport.com/jtwwen/gywm/jtjj/

Guangzhou Port Company Limited. *Port Annual Report 2021* https://www.gzport.com/foreground/index.html?id=e5e4ecaf19b14bd798705bf1d77f286a

Shenzhen Yantian Port Group Co.,Ltd. *Port Annual Report 2021* https://www.ytport.com/tzgg/

Xiamen Port Holding Group Co.Ltd. *Port Annual Report 2021* http://port.xm.gov.cn/

Dongguan Port Group. *Port Annual Report 2021* https://www.dgportgroup.com/

Tianjin Port (Group) Co.,Ltd. *Port Annual Report 2021* https://www.ptacn.com/channels/12.html

Beibu Gulf *Port. Annual Report 2021* https://www.bbwgw.com/bbwport/index

Qingdao Port International Co.Ltd. *Annual Report 2021* https://www.qingdao-port.com/portal/en

Rizhao Port Group Co., Ltd. *Annual Report 2021* http://en.rzport.com/zoujingangkou.jsp

Jiangsu Lianyungang Port Co.Ltd. *Port Annual Report 2021* https://www.sbgx123.com/lyggk/index.html

LiaoNing Port Group. *Annual Report 2021* http://www.liaoningport.com/html/en/index.html

Yantai Port. *Annual Report 2021* https://www.yantaiport.com.cn/

Tangshan Port Group Co.Ltd. *Annual Report 2021* http://www.jtport.com.cn/

Fuzhou Port Group Co.,Ltd *Annual Report 2021* https://www.fzport.com/contents/list?tid=144

Li, H., Jiang, L., Liu, J., & Su, D. (2022). Research on the evaluation of logistics efficiency in Chinese coastal ports based on the four-stage DEA model. *Journal of Marine Science and Engineering,* 10(8), 1147..

Ghiara, H., & Tei, A. (2021). Port activity and technical efficiency: determinants and external factors. *Maritime Policy & Management,* 48(5), 711-724.

Pagano, A. M., Wang, G. W., Sánchez, O. V., & Ungo, R. (2013). Impact of privatization on port efficiency and effectiveness: results from Panama and US ports. *Maritime Policy & Management,* 40(2), 100-115.

Schøyen, H., & 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 & Logistics,* 15, 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,* 35(2), 107-122.

Choi, Y. (2011). The efficiency of major ports under logistics risk in Northeast Asia. *Asia-Pacific Journal of Operational Research*, 28(01), 111-123.

Sachish, A. (1996). Productivity functions as a managerial tool in Israeli ports. *Maritime policy and Management*, 23(4), 341-369.

Tongzon, J. L. (1995). Determinants of port performance and efficiency. *Transportation Research Part A: Policy and Practice*, 29(3), 245-252.

Roll, Y., & Hayuth, Y. E. H. U. D. A. (1993). Port performance comparison applying data envelopment analysis (DEA). *Maritime policy and Management*, 20(2), 153-161.

Barros, C. P., & Athanassiou, M. (2004). Efficiency in European seaports with DEA: evidence from Greece and Portugal. *Maritime Economics & Logistics*, 6, 122-140.

Pang, L.Port Performance Evaluation and Analysis Discussion. *Port Economics* .05(2005):43-44.

Huang, J. Y., & Yan, Y. X. (2004). Design scheme of comprehensive evaluation index system of port container transport competitiveness. *Shipp Manag*, 26(9), 8-10.

Wang, L., & Bi, Z. W. (2010). Evaluating the operational efficiency of major inland ports in China using the 3-Stage DEA model. *Ind Econ Res*, 4, 40-48.

Bendall, H. B., & Stent, A. F. (1987). On measuring cargo handling productivity. *Maritime Policy and Management*, 14(4), 337-343.

Talley, W. K. (1994). Performance indicators and port performance evaluation. *Logistics and Transportation Review*, 30(4), 339.

Liu, Z. (1995). The comparative performance of public and private enterprises: the case of British ports. *Journal of Transport Economics and policy*, 263-274.

Cullinane, K., Song, D. W., & Gray, R. (2002). A stochastic frontier model of the efficiency of major container terminals in Asia: assessing the influence of administrative and ownership structures. *Transportation Research Part A: Policy and Practice*, 36(8), 743-762.

Cullinane, K., & Song, D. W. (2003). A stochastic frontier model of the productive efficiency of Korean container terminals. *Applied economics,* 35(3), 251-267.

Cullinane, K., Song, D. W., Ji, P., & Wang, T. F. (2004). An application of DEA windows analysis to container port production efficiency. *Review of network Economics*, 3(2).

Cullinane, K., Ji, P., & Wang, T. F. (2005). The relationship between privatization and DEA estimates of efficiency in the container port industry. *Journal of economics and business*, 57(5), 433-462.

Cullinane, K., Song, D. W., & Wang, T. (2005). The application of mathematical programming approaches to estimating container port production efficiency. *Journal of productivity analysis*, 24, 73-92.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, 30(9), 1078-1092.

Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin philosophical magazine and journal of science,* 2(11), 559-572.

Figueroa-García, J. C., & Castro-Cabrera, C. E. (2015, August). On Solving CCR-DEA Problems Involving Type-2 Fuzzy Uncertainty Using Centroid-Based Optimization. *In International Conference on Intelligent Computing* (pp. 187-195). Cham: Springer International Publishing.

Hatami-Marbini, A., Tavana, M., Saati, S., & Agrell, P. J. (2013). Positive and normative use of fuzzy DEA-BCC models: A critical view on NATO enlargement. *International Transactions in Operational Research,* 20(3), 411-433.

Martinez-Budria, E., Diaz-Armas, R., Navarro-Ibanez, M., & Ravelo-Mesa, T. (1999). A study of the efficiency of Spanish port authorities using data envelopment analysis. *International Journal of Transport Economics/Rivista internazionale di economia dei trasporti,* 237-253.

Rajasekar, T., & Deo, M. (2018). The size effect of Indian major ports on its efficiency using Dea-Additive models. *Int. J. Adv. Manag.* Econ, 1, 12-18.

Jiang, B., & Li, J. (2009). DEA-based performance measurement of seaports in northeast Asia: Radial and non-radial approach. *The Asian Journal of Shipping and Logistics*, 25(2), 219-236.

Dong, G., Zhu, J., Li, J., Wang, H., & Gajpal, Y. (2019). Evaluating the environmental performance and operational efficiency of container ports: An application to the maritime silk road. *International journal of environmental research and public health,* 16(12), 2226.

Zheng, X. B., & Park, N. K. (2016). A study on the efficiency of container terminals in Korea and China. *The Asian Journal of Shipping and Logistics*, 32(4), 213-220.

Iyer, K. C., & Nanyam, V. N. (2021). Technical efficiency analysis of container terminals in India. *The Asian Journal of Shipping and Logistics*, 37(1), 61-72.

van Dyck, G. K. (2015). Assessment of port efficiency in West Africa using data envelopment analysis. *American Journal of Industrial and Business Management*, 5(04), 208.

Xu, L., & Xu, S. (2021, September). Evaluation and Forecast of Port Enterprise Innovation Efficiency Based on ES-DEA. *In 2021 5th Annual International Conference on Data Science and Business Analytics* (ICDSBA) (pp. 362-365). IEEE.

Liu, S., Park, S. H., Choi, Y. S., & Yeo, G. T.

(2022). *Efficiency evaluation of major container terminals in the top three cities of the Pearl River Delta using SBM-DEA and undesirable DEA*. The Asian Journal of Shipping and Logistics, 38(2), 99-106.

Mustafa, F. S., Khan, R. U., & Mustafa, T. (2021). Technical efficiency comparison of container ports in Asian and Middle East region using DEA. *The Asian Journal of Shipping and Logistics*, 37(1), 12-19.

Cullinane, K., & Wang, T. F. (2006). Data envelopment analysis (DEA) and improving container port efficiency. *Research in Transportation Economics*, 17, 517-566.

Pjevčević, D., Radonjić, A., Hrle, Z., & Čolić, V. (2012). DEA window analysis for measuring port efficiencies in Serbia. *Promet-Traffic & Transportation*, 24(1), 63-72.

Cullinane, K., & Wang, T. (2010). The efficiency analysis of container port production using DEA panel data approaches. OR *spectrum*, 32(3), 717-738.

Dewarlo, O. (2019). DEA window analysis for measuring port performances efficiency of four islands countries located in west Indian ocean countries. American *Journal of Industrial and Business Management*, 9(12), 2098-2111.

Wang, C. N., Nguyen, P. H., Nguyen, T. L., Nguyen, T. G., Nguyen, D. T., Tran, T. H., ... & Phung, H. T. (2022). A two-stage DEA approach to measure operational efficiency in Vietnam's port industry. *Mathematics,* 10(9), 1385.

PCA-DEA 모델을 기반으로 한 항만의 운영 효율성 평가

장해청 · 이향숙

국문요약

국내 운송을 담당하는 연안항들은 국가와 도시의 발전에 필수적인 역할을 하고 있다. 항만 효율성은 항만 경쟁력을 확보하기 위한 중요한 요소로, 이에 대한 연구는 기존 문헌에서 지속적으로 진행되었다. 중국의 경우 주로 지방의 항만 클러스터에 대한 연구에 초점을 맞추고 있는데, 미시적 관점에서만 접근하고 있으며, 최신 자료도 부재한 상황이다. 따라서 본 연구에서는 최신 자료를 활용하여 중국 상위 17개 연안항의 운영 효율성을 종합적으로 분석하고자 한다. 본 연구에서는 컨테이너 처리량을 산출변수로 선정하고, 크게 토지, 자본, 노동, 인프라에 속하는 13개 지표로부터 PCA(Principal Component Analysis) 분석을 통해 4개의 투입변수를 최종 선정하였다.그런 다음 17개 항구의 운영 효율성을 DEA (데이터 포위 분석)로 분석했습니다. 분석 결과, 상하이, 닝보-저우산, 광저우, 샤먼, 둥관의 5개 항만이 효율적인 반면, 나머지 12개의 항만은 상대적으로 비효율적인 것으로 나타났다. 본 연구는 중국의 연안항을 보다 거시적 관점에서 비교·분석한 것으로 이를 통해 상대적 비교가 가능하며, 향후 항만의 발전 전략 및 정책 수립을 위한 기초자료로 활용할 수 있을 것이다.

주제어: 항만 운영, 효율성, 중국 연안항, PCA, DEA