Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Analyzing the Performance of Malaysian Logistics Companies using Principal Component Analysis (PCA) and Data Envelopment Analysis (DEA)

Khaleeda Qeestina Muhamad Farid[1], Farah Najwa Mahmadi[1], Noor Asiah Ramli[1,2,*], S. Sarifah Radiah Shariff[1,2], Nur Huda Nabihan Md Shahri[1]

[1] School of Mathematical Sciences, College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
[2] Research Interest Group (RIG) Logistics Modelling, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

ABSTRACT

Logistics operations play a crucial role in the overall business activities, especially in today's global and interconnected marketplace. A well-implemented logistics operation leads to cost savings and enhanced market responsiveness, resulting in an overall business competitiveness in today's challenging business environment. This study used Principal Component Analysis (PCA) and Data Envelopment Analysis (DEA) to analyze the performance of 14 well-known logistics companies in Malaysia from the year 2010 until 2020. The most crucial factor affecting logistics performance was identified using PCA to reduce data dimensionality. PCA results showed that current assets, net fixed assets, current liabilities, operating income, and revenue significantly affected the performance of logistics companies. Then, the efficiency frontier was evaluated using DEA, which considered current assets, net fixed assets and current liabilities as input while operating income and revenue as output variables. In the DEA process, Lingkaran Trans Kota Holdings Berhad is the only company that maintained a full efficiency score of 100 percent throughout the entire period, indicating the efficient utilization of its resources. On the other hand, MISC Berhad was the least efficient, with an average efficiency score of 32.17 percent. This study's findings can be used to increase organizational competitiveness by optimizing performance and boosting efficiency.

Keywords:
Logistics companies; Principal Component Analysis; Data Envelopment Analysis

1. Introduction

In addition to the strong growth of science and technology and the trend of globalisation, logistics operations from production to consumption play an increasingly important role in improving business competitiveness in the manufacturing industry for both services and the overall economy [1]. Logistics is a term that refers to "the process of planning, implementing and controlling the efficient, effective flow and storage of goods, services, and related information from point of origin to point of consumption to meet customer requirements" (Council of Logistics Management, 1998).

* *Corresponding author.*
E-mail address: asiah2206@uitm.edu.my

https://doi.org/10.37934/araset.55.2.151166

Logistics is responsible for everything from warehousing and transportation to integrating the logistics operations of the entire supply chain.

The Malaysian logistics sector has received international recognition [2,3]. Malaysia was in the 32nd spot among 166 countries in the World Bank's 2016 Global Logistics Performance Index (LPI). The LPI score for 2016 was 3.43 [4]. The index is a ranking of the performance of various countries' logistics sectors. Next, Malaysia fell to the 41st spot in 2018 with an LPI score of 3.22. And recently, for 2023, Malaysia managed to get a score of 3.6 with a rank of 26^{th}, where it can be seen that the performance has improved from year to year [5]. In response to the situation, the Transport Ministry said that several initiatives would be implemented to position Malaysia as the preferred logistics gateway to Asia, as reported in *TheStar* by Keebler *et al.*, [6].

Besides implementing the initiatives, it is important to also measure the performance of logistics companies [7,8], and one of the good assessments is using the DEA model [9]. A company might want to assess the performance of its logistics system based on efficiency scores for some fundamental reasons. Companies can use these measures to drive revenue growth and thus increase shareholder value by lowering operating costs [6]. By measuring the operating costs, the company can decide whether, when, and where to make operational changes to reduce costs. This is crucial because it identifies the areas that need better asset management. Companies can improve the price-value relationship of the products offered through cost-cutting measures and service enhancements in order to attract and keep loyal customers [6]. However, there is a dearth of comprehensive research in the field of logistics.

Hence, this study assessed the efficiency performance of Malaysian logistics companies by using the Principal Component Analysis and Data Envelopment Analysis (PCA-DEA) method. PCA helps to identify the essential information in the data table and express it as new principal components (PCs). DEA is a non-parametric model that optimizes the efficiency of decision-making units (DMUs) based on various inputs and outputs. Efficiency is measured using the ratio of outputs to inputs. A higher output-to-input ratio indicates better efficiency, as it demonstrates that the system is generating more output for a given set of inputs. On the other hand, a lower output-to-input ratio implies lower efficiency, indicating that more inputs are required to achieve a particular level of output. This approach is valuable for measuring performance and optimizing the efficiency of logistics companies.

1.1 Principal Component Analysis (PCA)

The Principal Component Analysis (PCA) technique is widely employed in machine learning and data analysis for its ability to reduce dimensionality and visualize data [10-12]. The main objective of Principal Component Analysis (PCA) is to convert data with a high number of dimensions into a space with fewer dimensions while preserving a significant portion of the original variance. By reducing the dimensionality, the analysis becomes simpler, and processing efficiency can be enhanced, all while retaining vital information in the data [12,13]. The original variables are transformed into a new set of orthogonal variables, known as PCs, to achieve this reduction. Most of the variance of the original set is usually retained by the first PC, while progressively smaller sections that were not considered by the earlier PCs are retained by the successive PCs. This is how PCs are typically constructed [11].

PCA facilitates better decision-making for the logistic industry by reducing the logistical data, thus highlighting the crucial elements. It boosts efficiency, streamlines routes, and customizes services, assisting logistic organizations in expanding and responding to changing client demands. Azman *et al.*, [14] used PCA to minimize data dimensionality before undertaking a spatial distribution analysis of the logistics development in Xinjiang. This approach made the results easier to comprehend and allowed the researchers to concentrate on the key factors influencing the relationship between logistics development and spatial dispersion. Logistics companies in Malaysia may profit by using PCA statistical technique to optimize many areas of their operations. In a different application area, Mohanta *et al*., [15] conducted an environmental metric study to assess the air quality pattern in Klang Valley, Malaysia. The study used data from five air monitoring stations in Klang Valley for the years 2010 to 2014. PCA was used to analyze the data to identify the significant pollutant parameters contributing to air quality issues in Klang Valley.

Logistics companies can also evaluate their operational efficiency, identify important performance indicators, and compare their performance to industry benchmarks by integrating PCA and DEA methods. PCA can help to enhance the performance of DEA [16]. They used a PCA-DEA approach to combine the benefits of PCA and DEA. The PCA-DEA model enhanced DEA's ability to discriminate between efficient and wasteful variables. The data's dimensionality was also reduced by the PCA-DEA model, increasing its computational effectiveness. The study demonstrated how PCA may be utilised to enhance DEA's functionality, making the PCA-DEA model a promising new method for evaluating the relative effectiveness of DMUs. Another study used PCA-DEA-Tobit to evaluate China's logistics sector under the condition of carbon emission restrictions [17]. Significant geographical differences in logistics efficiency were discovered, with inadequate scale efficiency being the main issue. Efficiency in logistics was positively correlated with regional economic and logistical development, whereas government influence and energy structure were negatively correlated.

More recently, Taletović and Sremac [18] applied the methods of PCA-DEA in their study on transportation companies located in Bosnia and Herzegovina. Initially the study defined six inputs and four outputs parameter. However, after applied PCA, the number of input and output parameters were reduced to two and one respectively before proceeding with efficiency measurement using DEA.

1.2 Technical Efficiency using Data Envelopment Analysis (DEA)

Prior studies have presented many approaches for assessing technical efficiency, including Stochastic Frontier Analysis [19], Neural Network Model [20] and Slack-based Measure [21]. Nevertheless, the most widely used method is Data Envelopment Analysis (DEA). The DEA method is a powerful tool for evaluating the efficiency of organizations [22]. This technique allows for comparisons across units that utilize numerous inputs to produce multiple outputs, resulting in a single measure of total performance. In addition to the measure, the DEA also provides targets for performance, potential improvements achievable through changes in scale, size, and/or resource allocation, and the identification of best practices and benchmark units. DEA has been employed in a wide range of settings, such as financial institutions [23,24], educational institutions [25,26], healthcare facilities [27,28], agriculture development [29,30], and many more.

DEA also does not require making assumptions about the specific form of the production function. Therefore, it can be used to evaluate organizations with different production processes, and this ability is a major advantage in many cases. Another key advantage of DEA lies in its ability to measure multiple dimensions of performance. Therefore, DEA can be used to evaluate various factors affecting performance, such as cost, quality, and customer satisfaction.

Since its invention in 1978, DEA has been extensively applied to the efficiency assessment of multiple input-output problems. Since then, researchers have been working on expanding the DEA model based on various theoretical and reasonable backgrounds. Many models, including radial and non-radial, static and dynamic, and single structure and network structure models, have been proposed. For instance, in a study, DEA and life cycle assessment (LCA) techniques were incorporated into the methodology [31]. The objectives were to evaluate the operational efficiency of 30 grocery undertakings in order to set up operating and environmental benchmarks to optimise performance. Their research has shown that the LCA and DEA methodology can be used to evaluate environmental performance in the services sector.

According to a finding by Tongzon *et al.*, [32], similar countries that are integrating their international shipping sectors can learn from the ASEAN experience. Additionally, it advances knowledge of the many challenges and how these challenges may affect the efficiency of shipping and logistics companies. For example, Li and Qunxi, [33] used the DEA model to study the efficiency of the logistics industry in western China. The researchers found that Inner Mongolia and Ningxia had the highest efficiency, with an average comprehensive technical efficiency of 1. Guizhou, Guangxi, and Qinghai also had favourable development, with comprehensive technical efficiencies above 0.9. However, the remaining regions had comprehensive technical efficiencies below 0.9, indicating relatively low logistics efficiency in many western regions.

In a recent study by Lee *et al.*, [34], researchers used DEA to analyze the performance of logistics companies in Malaysia. They used the methodology of basic indicator approach (BIA) to incorporate operational risk variables in order to increase the operational efficiency and effectiveness of listed logistics companies. The study focused on 27 publicly traded logistics companies on Bursa Malaysia between 2010 and 2021. The researchers were able to determine the areas for improvement in underperforming businesses by using the DEA model. These findings provide insightful information for managers and practitioners. Another study was performed in Malaysia with a focus on corporate courier services [35]. The effectiveness and levels of client satisfaction of various courier services were examined using DEA. The study’s objectives were to assess client satisfaction with the services offered, evaluate the effectiveness of various businesses, and rank them accordingly. Pos Laju, City-Link Express, J&T Express, and GD Express were chosen as the courier services in Hulu Terengganu to be used as the data sources. Pos Laju was found to be the most effective business, followed by City-Link Express and J&T Express in second and third places, respectively. The least effective of the four services, GD Express, received the lowest efficiency rating. The outcomes were primarily affected by elements like assurance, empathy, and tangibility.

Although numerous studies have been conducted on the application of PCA and DEA approaches in various application areas [36-38], so far, scarce research in Malaysia has deployed these approaches to measure performance, particularly in Malaysia’s logistics industry. Therefore, this study fills this gap so that the performance of this industry can be observed.

2. Methodology

2.1 Data Source

This study used secondary data obtained from Bursa Malaysia due to its availability in the yearly financial reports of various logistics companies. To be listed on Bursa Malaysia, these logistics companies are evaluated based on other criteria. Examples of these include a widespread network, a substantial market share, strong financial performance, advanced technological skills, and efficient operations. The data covers the period from 2010 to 2020. The study focused on 14 Malaysian logistics companies as listed in Table 1 below to assess the efficiency and inefficiency of a group of DMUs. The PCA approach was employed to identify the most significant variable(s), and DEA was used to measure efficiency.

Table 1
Listed Malaysian logistics companies 2010–2020

Logistics companies	Decision making unit (DMU)
Ancom Logistics Berhad	DMU1
Hubline Berhad	DMU2
Malaysia Airports Holdings Berhad	DMU3
Harbour-Link Group Berhad	DMU4
Malaysian Bulk Carriers Berhad	DMU5
MISC Berhad	DMU6
MMC Corporation Berhad	DMU7
Tiong Nam Logistics Holdings Berhad	DMU8
CJ Century Logistics Holdings Berhad	DMU9
Complete Logistics Services Berhad	DMU10
Freight Management Holdings Berhad	DMU11
Lingkaran Trans Kota Holdings Berhad	DMU12
See Hup Consolidated Berhad	DMU13
Suria Capital Holdings Berhad	DMU14

2.2 Principal Component Analysis (PCA)

PCA is a statistical method used to identify patterns and relationships in data. PCA is defined as a multivariate approach that examines a data table in which observations are characterised by several interrelated quantitative dependent variables [12,39]. The key data extracted from the table is represented as a collection of new orthogonal variables called PCs, displayed as points on maps to show the pattern of similarity between the observations and the variables [13]. In the end, the PC outperforms the original data, simplifies the index system by reducing the complexity of the original multi-dimensional problem, and increases the efficiency of the study. The main PCA algorithm is as follows [17]:

(1) **Compute the covariance matrix.**
X_{ij} is the *j*-th indicator's observed value from the *i*-th province. In addition to $\overline{X_J}$ and $\sqrt{var\ (X_j)}$ being the sample mean and standard deviation of the *j*-th index, $X_{ij}{}^*$ is the normalized index value.

(2) **Eigenvalues and Eigenvector calculation.**
The eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_7 \geq 0$ and the accompanying orthogonalization eigenvectors u_1, u_2, u_3 are determined for the correlation matrix R for matrix $(X_{ij}{}^*)_{n\times 7}$.

(3) **Ranking and Selection of Principicial Component.**
When $E \geq 80$, the minimum integer of *m* is regarded as the value of *m*, which is *m* when the number of principal components is $E \geq 80$. The cumulative variance eigenvalue contribution rates are calculated as follows: $E = \sum_{m=1}^{p} \lambda_m \ / \sum_{n=1}^{7} \lambda_n$

(4) **Projection of the data.**
Following that, the first *m* main components are isolated: $F = \sum_{J=1}^{7} u_{kj} x_j\ (k = 1,2, \dots m)$

(5) **Interpretation and Analysis.**
The sum of the variance contribution rate and the weight coefficient, where α_k is the variance contribution rate of the *k*-th principal component and y_k is the *k*-th principal component, is then used to calculate the comprehensive evaluation index value $E =$

$\sum_{k=1}^{1} \alpha_k y_k$ for each evaluated object. Finally, the logistics performance in each province/municipality is ranked according to the *F* value.

2.3 Data Envelopment Analysis (DEA)

DEA is a data-driven method for calculating the total factor efficiency of homogenous DMUs [40]. DEA can estimate the efficiency level of DMUs in the form of input and output ratio by using optimization techniques based on objective data of the evaluation object to derive the weight of a set of best input and output variables [41]. The DMUs are assessed for relative efficiency by contrasting the DMUs' actual input-output data against the frontier's projection data. Since it was first proposed, the DEA approach has found extensive use in numerous fields [33,42]. Many researchers have developed novel concepts and models for the DEA approach, such as the cross-efficiency model, super-efficiency model, slack based measure (SBM) model, network structure model, and so on [1,43,44].

The DEA theory approach was first introduced by Charnes *et al*., [41]. The CCR model was named based on the first letters of the three surnames of Charnes, Cooper, and Rhodes. According to the CCR model, scaling returns are constant (CRS). The assessed objects are represented by DMUs in the DEA technique. The CCR model can be thought of as virtually reducing several input/output problems to a single input/output problem based on the weights of the input and output variables. This single efficiency is measured for a specific DMU by using the virtual output to input ratio. The CCR model's linear programming can be described as optimizing a single DMU's efficiency with the restriction that the total efficiency of all DMUs does not exceed one.

Eq. (1) is referred to as the DEA multiplier form, which is the envelope form that can represent its dual model. It is referred to as a comprehensive technical efficiency model since the CCR model assumes that the returns to scale remain constant and the technical efficiency obtained includes the scale efficiency component [41]. DEA models can also be classified as input-oriented, output-oriented, or non-oriented based on how efficiency is measured.

$$\begin{aligned} & \max h_j = \sum_{i=1}^{s} u_i y_{ij} \\ & s.t. \sum_{i=1}^{s} u_i y_{ik} - \sum_{r=1}^{m} v_r x_{rk} \leq 0, \quad k = 1,2,\dots,n \\ & \sum_{r=1}^{m} v_r x_{rj} = 1 \\ & u_i, v_r \geq 0; \quad i = 1,2,\dots,\mathrm{s}; \quad r = 1,2,\dots,\mathrm{m} \end{aligned} \tag{1}$$

where

h_j = efficiency of DMU_j
u_i = weight of output *i*
y_{ij} =output *i* of the DMU_j
vr =weight of input *r*
x_{rj} =input *r* of the DMU_j
y_{ik} =output *i* of the DMU_k
x_{rk}= input *r* of the DMU_k
s = max outputs, *m* = max inputs

This study focused on input orientation with the CCR model. For the input orientation, the assessment was on the movement of the input level towards the frontier via proportional reduction while the output level remained unchanged. In addition, the DEA-CCR model assumed constant returns to scale, enabling all observed production combinations to be proportionally scaled up or down [41].

3. Results

3.1 Principal Component Analysis (PCA) Technique

3.1.1 Identifying the extracted components

One way to determine how many PCs to include in an analysis is to look at the eigenvalues. When comparing separate variables, each component is represented by a unit of variance. This means that when researchers use methods such as PCA, each component accounts for a certain amount of the total variability or dispersion in the dataset [45]. Eigenvalues are a measure of the variance of a PC. Components with eigenvalues greater than or equal to 1 are significant [46]. This means that the corresponding PC contains more information than any single variable alone. Therefore, all PCs with eigenvalues greater than 1 can be interpreted. To make a more informed decision about how many PCs to include, it is also helpful to consider the percentage of total variation explained by each PC. This is helpful in determining how much information each of the PCs is really capturing.

Referring to the results in Table 2 (refer to the Total Initial Eigenvalues column), PCA extracted two PCs for each year from 2010 to 2020, except for 2015 and 2018, where only one PC and three PCs, respectively, were identified. For example, the PC with an eigenvalue greater than 1 explained about 89.125% of the total variance in the original data of 14 Malaysian logistics companies in 2010. The PCA results showed that the first PC (PC1) accounted for 71.735% of the total variance in the data. PC1 had high loadings on the variables of current assets, net fixed assets, current liabilities, operating income, and revenue. It also had low loadings on the variables of net asset per share and earnings per share. Next, the second PC (PC2) accounted for 17.39% of the total variance in the data (further details on these results are elaborated on in Section 4.1.3). Only PC1 and PC2 had eigenvalues greater than one. Hence, this study focused on PC1 and PC2 only as they were the only two PCs that accounted for significant variance in the data. The remaining PCs did not account for enough variance to be considered meaningful, and hence, they were not considered in the analysis.

Table 2
Total Variance Explained from 2010 until 2020

Year	Component	Initial eigenvalues			Extraction sums of squared loadings		
		Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
2010	1	5.021	71.735	71.735	5.021	71.735	71.735
	2	1.217	17.390	89.125	1.217	17.390	89.125
	3	0.599	8.562	97.686			
	4	0.143	2.049	99.735			
	5	0.018	0.250	99.986			
	6	0.001	0.013	99.998			
	7	0.000	0.002	100.000			
2011	1	4.821	68.873	68.873	4.821	68.873	68.873
	2	1.065	15.210	84.083	1.065	15.210	84.083
	3	0.790	11.279	95.362			
	4	0.296	4.225	99.587			

Year	Component	Initial eigenvalues			Extraction sums of squared loadings		
		Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
	5	0.026	0.369	99.955			
	6	0.003	0.041	99.996			
	7	0.000	0.004	100.000			
2012	1	4.544	64.918	64.918	4.544	64.918	64.918
	2	1.292	18.452	83.370	1.292	18.452	83.370
	3	0.659	9.417	92.787			
	4	0.486	6.949	99.736			
	5	0.017	0.245	99.981			
	6	0.001	0.019	100.000			
	7	5.471E-6	7.816E-5	100.000			
2013	1	4.840	69.141	69.141	4.840	69.141	69.141
	2	1.100	15.718	84.860	1.100	15.718	84.860
	3	0.669	9.561	94.421			
	4	0.224	3.206	97.626			
	5	0.164	2.339	99.965			
	6	0.002	0.017	99.999			
	7	6.275E-5	0.001	100.000			
2014	1	4.211	60.151	60.151	4.211	60.151	60.151
	2	1.075	15.356	75.507	1.075	15.356	75.507
	3	0.959	13.696	89.203			
	4	0.718	10.254	99.457			
	5	0.037	0.524	99.980			
	6	0.001	0.019	99.999			
	7	6.820E-5	0.001	100.000			
2015	1	4.680	66.852	66.852	4.680	66.852	66.852
	2	0.969	13.848	80.701			
	3	0.808	11.549	92.250			
	4	0.385	5.502	97.752			
	5	0.151	2.159	99.912			
	6	0.006	0.080	99.992			
	7	0.001	0.008	100.000			
2016	1	4.777	68.243	68.243	4.777	68.243	68.243
	2	1.173	16.764	85.007	1.173	16.764	85.007
	3	0.718	10.262	95.270			
	4	0.244	3.479	98.748			
	5	0.084	1.194	99.942			
	6	0.004	0.057	100.000			
	7	3.303E-5	0.000	100.000			
2017	1	4.750	67.860	67.860	4.750	67.860	67.860
	2	1.078	15.406	83.266	1.078	15.406	83.266
	3	0.948	13.543	96.810			
	4	0.174	2.491	99.301			
	5	0.046	0.663	99.964			
	6	0.002	0.022	99.987			
	7	0.001	0.013	100.000			
2018	1	4.009	57.267	57.267	4.009	57.267	57.267
	2	1.244	17.773	75.040	1.244	17.773	75.040
	3	1.012	14.463	89.503	1.012	14.463	89.503
	4	0.696	9.937	99.440			
	5	0.033	0.474	99.914			

Year	Component	Initial eigenvalues			Extraction sums of squared loadings		
		Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
	6	0.006	0.085	100.000			
	7	3.298E-5	0.000	100.000			
2019	1	4.079	58.274	58.274	4.079	58.274	58.274
	2	1.215	17.364	75.638	1.215	17.364	75.638
	3	0.836	11.940	87.578			
	4	0.735	10.503	98.081			
	5	0.130	1.851	99.932			
	6	0.005	0.065	99.998			
	7	0.000	0.002	100.000			
2020	1	4.387	62.665	62.665	4.387	62.665	62.665
	2	1.311	18.732	81.397	1.311	18.732	81.397
	3	0.702	10.026	91.423			
	4	0.397	5.668	97.090			
	5	0.202	2.887	99.977			
	6	0.001	0.015	99.992			
	7	0.001	0.008	100.000			

3.1.2 Determining the significant PCs

The PCs with the highest variance are typically chosen when performing the PCA. This method of selection aims to identify the PCs that account for the largest proportion of the variability in the initial dataset. The most significant patterns and structures in the data can be identified by choosing the PCs with the highest variance. These PCs show the dimensions along which the data points vary the greatest, enabling the study to minimize the dataset's dimensionality while preserving as much data as feasible. This selection procedure helps in locating the main elements responsible for the dataset's overall variability.

Table 3
Percentage variance from 2010 to 2020

Percentage of variance (%)			
Year	PC1	PC2	PC3
2010	71.735	17.390	x
2011	68.873	15.210	x
2012	64.918	18.452	x
2013	69.141	15.718	x
2014	60.151	15.356	x
2015	66.852	X	x
2016	68.243	16.764	x
2017	67.860	15.406	x
2018	57.267	17.773	14.463
2019	58.274	17.364	x
2020	62.665	18.732	x

Table 3 shows the PCA results from 2010 to 2020 for the whole variable. The analysis of the logistic companies' data using PCA in 2010 revealed that PC1 accounted for the highest percentage of variance with 71.735 percent. PC1 also accounted for the highest percentage of variance in 2011 (68.873 percent), 2012 (64.918 percent), 2013 (69.141 percent), 2014 (60.151 percent), 2016 (68.243

percent), 2017 (67.860 percent), 2018 (57.267 percent), 2019 (58.274 percent), and 2020 (62.665 percent). For 2015, the only PC available for selection was also PC1 (66.852 percent).

3.1.3 Determining the most important variable(s)

Feature selection in PCA is the process of finding the most important features in a dataset for a specific task. PCA helps simplify the dataset by transforming the original features into new ones that are independent and ranked by their importance. This simplification makes the data easier to understand and can enhance the performance of machine learning algorithms. There are various methods for feature selection in PCA. One common approach is to choose the features with the highest loadings. Loadings measure how strongly a feature is correlated with the principal components. Teoh *et al*., [35] found that the variables that are most strongly correlated with the PCs also explain the most variance in the dataset. This means that these variables are the most important ones to consider when interpreting the results of a PCA analysis.

Table 4 shows that five variables consistently displayed positive component loading values, indicating strong correlations across all 5 years from 2010 to 2020. These variables are current assets, net fixed assets, current liabilities, operating income, and revenue. Conversely, two variables, namely net assets per share and earnings per share, exhibited negative component loading values. Consequently, variables with negative values were excluded from the analysis. Thus, for the purpose of identifying technical efficiency in the performance of 14 logistics companies in Malaysia using DEA, this study chose to retain the five variables demonstrating high correlation values.

Table 4
Components loading between 2010 and 2014

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Component	PC1	PC1	PC1	PC1	PC1	PC1	PC1	PC1	PC1	PC1	PC1
Input (I)/Output (O)											
Current assets (I)	0.987	0.976	0.996	0.972	0.993	0.927	0.956	0.962	0.975	0.956	0.927
Net fixed assets (I)	0.978	0.955	0.923	0.902	0.803	0.938	0.949	0.977	0.360	0.487	0.944
Current liabilities(I)	0.991	0.971	0.980	0.983	0.949	0.956	0.948	0.949	0.971	0.955	0.929
Operating income (O)	0.946	0.840	0.732	0.923	0.871	0.926	0.962	0.944	0.976	0.963	0.892
Revenue (O)	0.996	0.990	0.985	0.981	0.878	0.918	0.977	0.992	0.987	0.959	0.907
Net asset per share	-0.221	-0.174	-0.351	-0.403	-0.313	-0.427	-0.302	-0.279	-0.218	-0.302	-0.383
Earnings per share	-0.417	-0.457	-0.335	-0.373	-0.227	-0.379	-0.307	-0.124	-0.098	-0.277	-0.096

3.2 Technical Efficiency Measurement Score

The efficiency of 14 logistics companies in Malaysia was measured using the DEA in R (deaR) online software. The results presented in Table 5 show that the efficiency scores ranged from 0% to 100%, with higher scores indicating greater efficiency and 100% means that a company was operating at full efficiency. Regarding inefficient DMUs, for instance, in 2010, DMU2 was 45.2 percent technically efficient. This finding suggests that DMU2 could reduce its input by approximately 54.8 percent while maintaining the current outputs of operating income and revenue.

The efficiency of 14 logistics companies in Malaysia was analyzed from 2010 to 2020. Lingkaran Trans Kota Holdings Berhad (DMU12) is the only company that maintained a full efficiency score of 100 percent throughout the entire period, indicating the efficient utilization of its resources. This finding is supported by the research done by Rusli *et al*., [47]. Despite using a different approach, the finding is still the same whereby Lingkaran Trans Kota Holding Berhad maintained full efficiency every year. One of the justifiable reasons for Lingkaran Trans Kota Holdings Berhad's being fully efficient is its continuous and strong financial performance throughout the research period. The company has additionally broadened its market footprint, both within its own country and abroad, by developing a robust network of operations. In addition, Lingkaran Trans Kota Holdings Berhad has demonstrated a dedication to achieving operational excellence by streamlining transportation, warehousing, and distribution processes. Effective logistics operations enhance the efficiency and cost-effectiveness of services.

Table 5
Percentage of technical efficiency of 14 logistics companies in Malaysia 2010−2020

DMUs	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	Average(%)
DMU1	**100**	**100**	**100**	93.68	**100**	**100**	**100**	**100**	**100**	**100**	**100**	99.43
DMU2	45.20	54.80	50.52	64.00	19.89	63.31	31.28	44.42	**100**	71.79	88.25	57.59
DMU3	53.31	57.75	**100**	**100**	37.68	56.93	36.07	35.76	44.14	44.40	35.36	54.67
DMU4	72.85	98.43	78.70	65.62	81.94	64.02	95.12	72.28	**100**	**100**	27.11	77.82
DMU5	**100**	**100**	48.44	49.62	42.56	31.50	21.70	34.44	19.00	44.44	59.21	50.08
DMU6	34.35	43.98	36.10	32.04	36.02	37.84	18.88	25.05	33.33	27.74	28.52	32.17
DMU7	33.30	40.66	41.58	31.19	35.59	90.28	23.84	24.72	38.35	46.11	55.44	41.91
DMU8	56.53	56.76	35.65	25.12	47.17	38.58	23.33	25.72	51.22	41.76	50.45	41.13
DMU9	98.21	80.49	80.72	58.22	61.52	72.47	69.32	75.43	73.21	84.16	80.58	75.85
DMU10	71.48	**100**	98.58	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	97.28
DMU11	100	97.98	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	99.82
DMU12	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	100
DMU13	**100**	**100**	**100**	84.66	87.01	100	82.68	85.91	70.31	80.38	98.73	89.97
DMU14	68.28	50.22	81.18	72.18	46.81	98.84	55.17	83.02	**100**	60.96	59.35	70.55
Average(%)	73.82	77.22	75.10	69.74	64.03	75.27	61.24	64.77	73.54	71.55	70.21	

Five companies achieved a perfect efficiency score of 100 percent for 2010. The least efficient company in 2010 was MMC Corporation Berhad (DMU7), with an efficiency score of 33.30 percent. In 2011, some companies like Ancom Logistics Berhad (DMU1), Malaysian Bulk Carriers Berhad (DMU5), Lingkaran Trans Kota Holdings Berhad (DMU12), and See Hup Consolidated Berhad (DMU13) remained fully efficient with a score of 100%. Complete Logistics Services Berhad (DMU10) improved to a perfect 100 percent score from the previous year. However, CJ Century Logistics Holdings Berhad (DMU9), Freight Management Holdings Berhad (DMU11), and Suria Capital Holdings Berhad (DMU14) became less efficient with scores of 80.49 percent, 97.98 percent, and 50.22 percent, respectively.

Some companies showed improvement from 2019 to 2020, including Malaysian Bulk Carriers Berhad (DMU5), MMC Corporation Berhad (DMU7), CJ Century Logistics Holdings Berhad (DMU9), and See Hup Consolidated Berhad (DMU13). However, the performance of others like Hubline Berhad (DMU2), MISC Berhad (DMU6), and Suria Capital Holdings Berhad (DMU14) declined. The rest of the companies maintained their scores.

Figure 1 presents the average efficiency score of each company for 11 years from 2010 to 2020. Lingkaran Trans Kota Holdings Berhad (DMU12) stands out as the most efficient logistics company throughout the years, maintaining an average efficiency score of 100 percent.

Conversely, MISC Berhad (DMU6) was the least efficient, with an average efficiency score of 32.17 percent. This discovery offers a comprehensive framework for identifying elements that could potentially contribute to a company being seen as underperforming. According to the data analyzed in this study, the company is experiencing financial difficulties, including a decline in operating income. Difficulties in effectively managing financial resources may have a negative impact on the overall functioning of the company. Furthermore, this company may be experiencing operational inefficiencies, resulting in delays, blunders, or escalated expenditures in its logistic operations.

Inadequate operational performance has the potential to impact customer happiness and overall competitiveness, thus making it the least efficient.

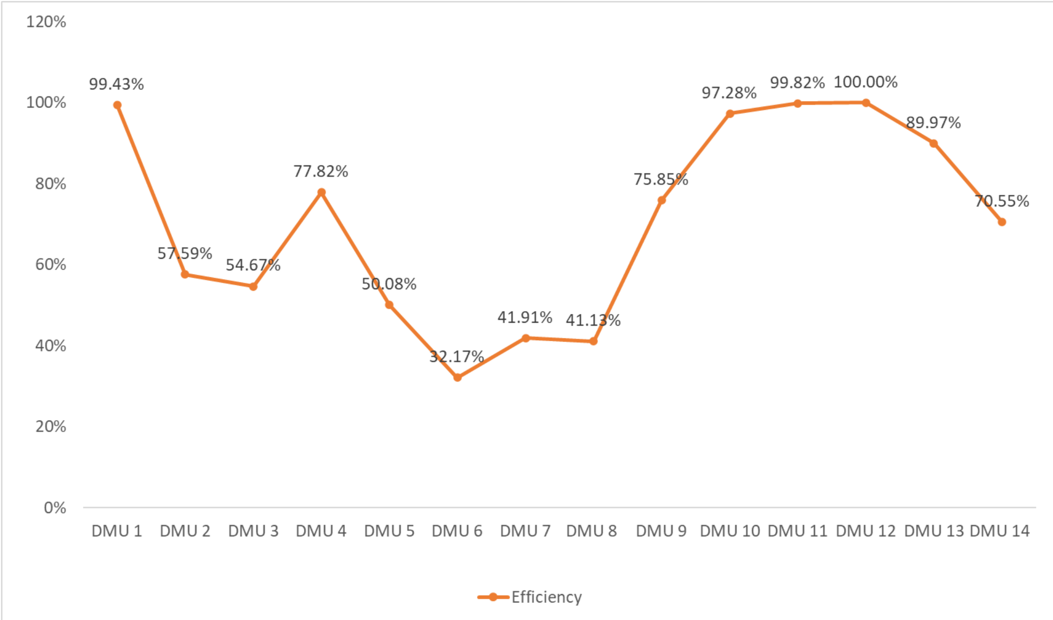

Fig. 1. Average efficiency score of 14 logistics companies in Malaysia

The findings can also be interpreted by looking at the yearly trend. Figure 2 depicts the shifts in the average technical efficiency scores from 2010 until 2020. The yearly scores fell in the range of 61−77 percent throughout the years. The lowest efficiency score was recorded in 2016 with 61.24 percent, whereas the highest efficiency was reported in 2011 with 77.22 percent.

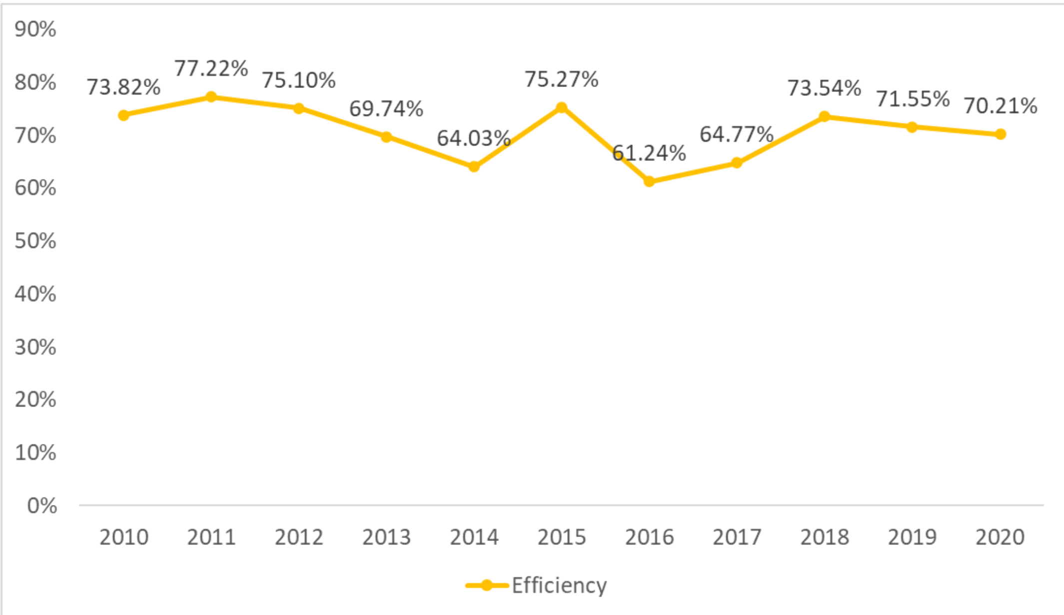

Fig. 2. Average efficiency score yearly between 2010 and 2020

4. Conclusion

Two analyses were performed to evaluate the efficiency of the logistic businesses in Malaysia. In the first step, PCA was used to identify the most important factors contributing to the Malaysian logistics industry's performance. DEA was also employed to measure the technical efficiency of the Malaysian logistics industry. In PCA, we used the highest variance percentage, which was the first PC. From 2010 to 2020, the most significant variables in the first PC were current assets, net fixed assets, operating income, and revenue. The DEA results showed that Lingkaran Trans Kota Holdings Berhad scored 100 percent in terms of efficiency throughout the study period. MMC Corporation Berhad reported the lowest efficiency score of 36.46 percent. The efficiency scores for 2010 until 2020 showed that 13 out of the 14 logistics companies were inefficient, with uneven efficiency scores throughout the study period. Based on the improvement table's results, current assets, net fixed assets, and current liabilities will all need to be minimized to maintain the same number of outputs. Since the study was input-oriented, the areas where the input variables should be minimized were considered in order to achieve maximum efficiency.

This work may provide a significant empirical contribution to the existing literature on logistic systems. Previous research has shown minimal integration between Principal Component Analysis (PCA) and Data Envelopment Analysis (DEA), making it particularly valuable in the Malaysian setting. Furthermore, this empirical evidence can provide valuable insights into the policies, rules, and regulations pertaining to enhancing organizational performance.

For future research, the researchers wish to embark on determining the influence factors on the efficiency score. It can be conducted by using another model, such as regression analysis. Due to the ever-changing nature of the logistics industry, it is advantageous to examine how external factors, such as advancements in technology, changes in regulations, or shifts in the global economy, may have influenced the efficiency scores. This analysis will provide a more comprehensive understanding of logistic performance as a whole.

Acknowledgement

We are grateful for the financial support provided by the Universiti Teknologi MARA (UiTM) under the Institut Pengangkutan Malaysia (Mitrans-EK) with project code: RS12021GRN79RN001 that enabled us to carry out this study effectively.

References

[1] Ho, Nguyen-Nhu-Y., Phuong Mai Nguyen, Thi-Minh-Ngoc Luu, and Thi-Thuy-Anh Tran. "Selecting partners in strategic alliances: An application of the SBM DEA model in the Vietnamese logistics industry." *Logistics* 6, no. 3 (2022): 64. https://doi.org/10.3390/logistics6030064

[2] Bakar, Mohd Azlan Abu, and Harlina Suzana Jaafar. "Malaysian logistics performance: A manufacturer's perspective." *Procedia-Social and Behavioral Sciences* 224 (2016): 571-578. https://doi.org/10.1016/j.sbspro.2016.05.442

[3] Sharif, K. I. M., M. T. B. Yiaw, M. A. Ismail, Q. U. Khand, M. G. Hassan, and S. A. Shahron. "Leveraging technology for sustainable logistics: logistics performance index in inland ports with eco strategy and sustainable practices." *International Journal of Intelligent Systems and Applications in Engineering* 12, no. 1 (2024): 666-675.

[4] Alwadood, Zuraida, Norlenda Mohd Noor, Nurul Fadhila Ahmad Salehin, and Nur Izzati Mahadhir. "Performance measure of Malaysian seaports using Data Envelopment Analysis (DEA)." *Journal of advanced research in computing and applications* 14, no. 1 (2019): 1-9.

[5] Jean-François Arvis, Lauri Ojala, Ben Shepherd, Daria Ulybina, and Christina Wiederer. "The Logistics Performance Index and Its Indicators. " World Bank Publications. (2023).

[6] Keebler, James S., and Richard E. Plank. "Logistics performance measurement in the supply chain: a benchmark." *Benchmarking: An International Journal* 16, no. 6 (2009): 785-798. https://doi.org/10.1108/14635770911000114

[7] Karabacak, Ebubekir, and Hüseyin Ali Kutlu. "Evaluating the Efficiencies of Logistics Centers with Fuzzy Logic: The Case of Turkey." *Sustainability* 16, no. 1 (2024): 438. https://doi.org/10.3390/su16010438

[8] Dong, Changrong, Gongmin Zhao, Yuanhao Wang, Yongjie Wu, and Haimeng Liang. "Evaluation of logistics-industry efficiency and enhancement path in China's Yellow River Basin under dual carbon targets." *Sustainability* 15, no. 17 (2023): 12848. https://doi.org/10.3390/su151712848

[9] Krmac, Evelin, and Mozhgan Mansouri Kaleibar. "A comprehensive review of data envelopment analysis (DEA) methodology in port efficiency evaluation." *Maritime Economics & Logistics* 25, no. 4 (2023): 817-881. https://doi.org/10.1057/s41278-022-00239-5

[10] Gwelo, Abubakari S. "Principal components to overcome multicollinearity problem." *Oradea Journal of Business and Economics* 4, no. 1 (2019): 79-91. https://doi.org/10.47535/1991ojbe062

[11] Ghojogh, Benyamin, Mark Crowley, Fakhri Karray, and Ali Ghodsi. "Principal component analysis." In *Elements of Dimensionality Reduction and Manifold Learning*, pp. 123-154. Cham: Springer International Publishing, 2023. https://doi.org/10.1007/978-3-031-10602-6_5

[12] Wold, Svante, Kim Esbensen, and Paul Geladi. "Principal component analysis." *Chemometrics and intelligent laboratory systems* 2, no. 1-3 (1987): 37-52. https://doi.org/10.1016/0169-7439(87)80084-9

[13] Ghazali, Norhafizah, and Zalina Mohd Ali. "Principal Component Analysis Approach in Klang River Water Quality Index Modelling." (2023). https://doi.org/10.13189/eer.2023.110112

[14] Azman Shah Ismail, Azman Shah Ismail, Ahmad Makmom Abdullah Ahmad Makmom Abdullah, and A. S. Mohd Armi. "Environmetric study on air quality pattern for assessment in Northern region of Peninsular Malaysia." (2017): 186-196. https://doi.org/10.3923/jest.2017.186.196

[15] Mohanta, Kshitish Kumar, Deena Sunil Sharanappa, and Abha Aggarwal. "Efficiency analysis in the management of COVID-19 pandemic in India based on data envelopment analysis." *Current Research in Behavioral Sciences* 2 (2021): 100063. https://doi.org/10.1016/j.crbeha.2021.100063

[16] Adler, Nicole, and Boaz Golany. "Including principal component weights to improve discrimination in data envelopment analysis." *Journal of the Operational Research Society* 53, no. 9 (2002): 985-991. https://doi.org/10.1057/palgrave.jors.2601400

[17] Deng, Fumin, Lin Xu, Yuan Fang, Qunxi Gong, and Zhi Li. "PCA-DEA-tobit regression assessment with carbon emission constraints of China's logistics industry." *Journal of Cleaner Production* 271 (2020): 122548. https://doi.org/10.1016/j.jclepro.2020.122548

[18] Taletović, M., and S. Sremac. "PCA-DEA model for efficiency assessment of transportation company." *International Journal of Management and Decision Making* 2, no. 1 (2023): 11-20. https://doi.org/10.56578/jimd020102

[19] Arsad, Roslah, Zaidi Isa, Ezzah Suraya Sarudin, Budi Halomoan Siregar, and Mohammad Nasir Abdullah. "Evidence of Malaysian Company Performance using Copula and Stochastic Frontier Analysis During the COVID-19 Pandemic." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 33, no. 1 (2023): 256-266. https://doi.org/10.37934/araset.33.1.256266

[20] Santin, Daniel, Francisco J. Delgado, and Aurelia Valino. "The measurement of technical efficiency: a neural network approach." *Applied Economics* 36, no. 6 (2004): 627-635. https://doi.org/10.1080/0003684042000217661

[21] Duho, King Carl Tornam. "Intellectual capital and technical efficiency of banks in an emerging market: a slack-based measure." *Journal of Economic Studies* 47, no. 7 (2020): 1711-1732. https://doi.org/10.1108/JES-06-2019-0295

[22] Xin-gang, Zhao, and Wei Zhen. "The technical efficiency of China's wind power list enterprises: An estimation based on DEA method and micro-data." *Renewable Energy* 133 (2019): 470-479. https://doi.org/10.1016/j.renene.2018.10.049

[23] Antunes, Jorge, Abdollah Hadi-Vencheh, Ali Jamshidi, Yong Tan, and Peter Wanke. "Cost efficiency of Chinese banks: Evidence from DEA and MLP-SSRP analysis." *Expert Systems with Applications* 237 (2024): 121432. https://doi.org/10.1016/j.eswa.2023.121432

[24] Ramli, Noor Asiah, Siti Shaliza Mohd Khairi, and Nor Azizah Razlan. "Performance measurement of islamic and conventional banking in Malaysia using two-stage analysis of DEA model." *International Journal of Academic Research in Business and Social Sciences* 8, no. 4 (2018): 1185-1197. https://doi.org/10.6007/IJARBSS/v8-i4/4152

[25] Liu, Jiani, Kim Jungyin, Shim Jaewoo, Lee Heechul, and Wasi Ul Hassan Shah. "Evaluating the efficiency, productivity change, and technology gaps of China's provincial higher education systems: A comprehensive analytical framework." *Plos one* 19, no. 1 (2024): e0294902. https://doi.org/10.1371/journal.pone.0294902

[26] Zuluaga, Rohemi, Alicia Camelo-Guarín, and Enrique De La Hoz. "Assessing the relative impact of Colombian higher education institutions using fuzzy data envelopment analysis (fuzzy-DEA) in state evaluations." *Journal on Efficiency and Responsibility in Education and Science* 16, no. 4 (2023): 299-312. https://doi.org/10.7160/eriesj.2023.160404

[27] Andrews, Antony, and Grigorios Emvalomatis. "Efficiency measurement in healthcare: the foundations, variables, and models–a narrative literature review." *Economics* 18, no. 1 (2024): 20220062. https://doi.org/10.1515/econ-2022-0062

[28] GAO, Dian, Lushaobo SHI, Jinhui LIN, Xingmin WANG, and Dong WANG. "Research on the Efficiency and Equity of Rural Medical and Health Resources Allocation in China Based on DEA-GIS Methodology." *Chinese General Practice* 27, no. 07 (2024): 849. https://doi.org/10.12114/j.issn.1007-9572.2023.0413
[29] Shah, Wasi Ul Hassan, Yuting Lu, Jianhua Liu, Abdul Rehman, and Rizwana Yasmeen. "The impact of climate change and production technology heterogeneity on China's agricultural total factor productivity and production efficiency." *Science of The Total Environment* 907 (2024): 168027. https://doi.org/10.1016/j.scitotenv.2023.168027
[30] Zhang, Ziyi, Juntao Du, Zhiyang Shen, Hassan El Asraoui, and Malin Song. "Effects of modern agricultural demonstration zones on cropland utilization efficiency: An empirical study based on county pilot." *Journal of environmental management* 349 (2024): 119530. https://doi.org/10.1016/j.jenvman.2023.119530
[31] Álvarez-Rodríguez, Cristina, Mario Martín-Gamboa, and Diego Iribarren. "Combined use of Data Envelopment Analysis and Life Cycle Assessment for operational and environmental benchmarking in the service sector: A case study of grocery stores." *Science of the Total Environment* 667 (2019): 799-808. https://doi.org/10.1016/j.scitotenv.2019.02.433
[32] Tongzon, Jose L., and Sang-Yoon Lee. "Achieving an ASEAN single shipping market: shipping and logistics firms' perspective." *Maritime Policy & Management* 43, no. 4 (2016): 407-419. https://doi.org/10.1080/03088839.2015.1105393
[33] Li, Yao, and Qunxi Gong. "Efficiency Measurement of Logistics Industry in Western China Based on Data Envelopment Analysis Model." In *2021 International Conference on Big Data and Intelligent Decision Making (BDIDM)*, pp. 65-69. IEEE, 2021. https://doi.org/10.1109/BDIDM538[17].2021.00020
[34] Lee, Pei Fun, Weng Siew Lam, and Weng Hoe Lam. "Performance evaluation of the efficiency of logistics companies with data envelopment analysis model." *Mathematics* 11, no. 3 (2023): 718. https://doi.org/10.3390/math11030718
[35] Teoh, Yeong Kin, Suzanawati Abu Hasan, and Siti Nur Atikah Zahari. "Data envelopment analysis for Malaysia courier services performance and customer satisfaction/Teoh Yeong Kin, Suzanawati Abu Hasan, and Siti Nur Atikah Zahari." *Journal of Computing Research and Innovation (JCRINN)* 5, no. 3 (2020): 25-33. https://doi.org/10.24191/jcrinn.v5i3.135
[36] Omrani, Hashem, Zijiang Yang, Arash Karbasian, and Tamara Teplova. "Combination of top-down and bottom-up DEA models using PCA: A two-stage network DEA with shared input and undesirable output for evaluation of the road transport sector." *Socio-Economic Planning Sciences* 89 (2023): 101706. https://doi.org/10.1016/j.seps.2023.101706
[37] Taletović, M., and S. Sremac. "PCA-DEA model for efficiency assessment of transportation company." *International Journal of Management and Decision Making* 2, no. 1 (2023): 11-20. https://doi.org/10.56578/ijmd020102
[38] Rasinojehdehi, Reza, and Seyyed Esmaeil Najafi. "Integrating PCA and DEA techniques for strategic assessment of network security." *Computational Algorithms and Numerical Dimensions* 2, no. 1 (2023): 23-34. https://doi.org/10.22105/cand.2023.424893.1076
[39] Abdi, Hervé, and Lynne J. Williams. "Principal component analysis." *Wiley interdisciplinary reviews: computational statistics* 2, no. 4 (2010): 433-459. https://doi.org/10.1002/wics.101
[40] Cooper, William W., Lawrence M. Seiford, and Joe Zhu. "Data envelopment analysis: History, models, and interpretations." *Handbook on data envelopment analysis* (2011): 1-39. https://doi.org/10.1007/978-1-4419-6151-8_1
[41] Charnes, Abraham, William W. Cooper, and Edwardo Rhodes. "Measuring the efficiency of decision making units." *European journal of operational research* 2, no. 6 (1978): 429-444. https://doi.org/10.1016/0377-2217(78)90138-8
[42] Ersoy, Yusuf. "Performance evaluation in distance education by using data envelopment analysis (DEA) and TOPSIS methods." *Arabian Journal for Science and Engineering* 46, no. 2 (2021): 1803-1817. https://doi.org/10.1007/s13369-020-05087-0
[43] Chung, Yangho H., Rolf Färe, and Shawna Grosskopf. "Productivity and undesirable outputs: a directional distance function approach." *journal of Environmental Management* 51, no. 3 (1997): 229-240. https://doi.org/10.1006/jema.1997.0146
[44] Wu, Huaqing, Jie Wu, Liang Liang, and Yongjun Li. "Efficiency assessment of Chinese logistics firms using DEA." *International journal of shipping and transport logistics* 4, no. 3 (2012): 212-234. https://doi.org/10.1504/IJSTL.2012.047485
[45] Gárate-Escamila, Anna Karen, Amir Hajjam El Hassani, and Emmanuel Andrès. "Classification models for heart disease prediction using feature selection and PCA." *Informatics in Medicine Unlocked* 19 (2020): 100330. https://doi.org/10.1016/j.imu.2020.100330
[46] Kaiser, Henry F. "The application of electronic computers to factor analysis." *Educational and psychological measurement* 20, no. 1 (1960): 141-151. https://doi.org/10.1177/001316446002000116

[47] Rusli, Nur Annisa Ezzaty, Noor Asiah Ramli, S. Sarifah Radiah Shariff, Zalina Zahid, and Siti Aida Sheikh Hussin. "Evaluating the efficiency and productivity of Malaysian logistics companies using epsilon-based measure and Malmquist index during the Covid-19 pandemic." *Journal of Industrial Engineering and Management* 15, no. 4 (2022): 521-537. http://dx.doi.org/10.3926/jiem.4022