Spektrum Industri
Vol. 22, No. 2, 2024, pp. 90-110
ISSN 1693-6590
http://journal3.uad.ac.id/index.php/spektrum

Efficiency Evaluation in Indonesia's Quarrying Industry Using Variable Combinations DEA

Erni Puspanantasari Putri [a,*], Ivan A. Parinov [b], Chuleeporn Wongloucha [c]

[a] Department of Industrial Engineering, Universitas 17 Agustus 1945 Surabaya, Surabaya 60118, Indonesia
[b] I. I. Vorovich Mathematics, Mechanics and Computer Sciences, Southern Federal University, Rostov-on-Don 344006, Russia
[c] Faculty of Economics, Khon Kaen University, Khon Kaen 40002, Thailand
* Corresponding Author: erniputri@untag-sby.ac.id

ARTICLE INFO

Article history
Received June 14, 2024
Revised October 1, 2024
Accepted October 7, 2024

Keywords
Performance evaluation;
Quarrying;
Variable combinations (VCs);
VRS-DEA

ABSTRACT

Data Envelopment Analysis (DEA) is a method considered to evaluate a company's performance. DEA applies multiplies the input and output variables for analyzing the efficiency but does not provide guidance in selecting those variables. As a rule, researchers use several methods. If the number of variables used is too many, it will affect the efficiency value. This will reduce the strength of the efficiency value, which can cause all DMU values to be efficient. DEA and variable selection are important in performance evaluation because DEA aids in determining relative efficiency, whereas variable selection guarantees that the evaluation is based on the most relevant and significant aspects. The purpose of this study is to suggest the variable combination method for subtracting the number of variables that will be utilized in implementing the DEA. The method used in this study is the Average Input Variable Combinations (VCs)-Variable Returns-to-Scale (VRS) DEA. The data were classified, defined, and processed with a view to computing efficiency scores and DMU classifications. The research result indicated that the proposed method (VCs-DEA) treats the variable reduction factor and the average calculation factor to obtain the final result of the efficiency score. These two factors contribute to the accuracy of the efficiency value. Some real-world implications of these findings, such as making better use of resources, streamlining operations, and coming up with new plans, Furthermore, the evidence may be used to benchmark performance as well as help decision-makers in creating more effective policy. This study finds that only 1 out of 12 DMUs is efficient (8%), while the remaining 11 are inefficient (92%). Indonesia quarrying establishment can be classified into 3 categories such as Optimal Category (S-Sand); Middle Category (LS-Lime-Stone; F-Feldspars; Gr-Granite; SA-Stone and Andesite; K-Kaolin; Q-Quartz; and G-Gravel); and Less Category (So-Soil; C-Clay; M-Marble; and O-Others).

1. Introduction

Selecting the ideal input and output weights systematically helps improve the decision-making units' (DMUs') performance evaluation. This method avoids the possible difficulties in coming up with a standard foundation for classifying homogenous units and organizing the output from several effective boundaries. Restricting the weights according to the technical or intrinsic relationships between pertinent inputs and outcomes makes efficient frontier ranking easier. Additionally, it develops a logical structure that unifies different factors throughout all undervalued variables and DMUs. This approach effectively investigates the nuances of the inputs and outputs of homogenous units while evaluating their efficiency using the data envelopment analysis (DEA) method (Kraidi et al., 2024).

DEA is a linear programming-based method that adopts non-stochastic and non-parametric approaches in assessing relative efficiency. It produces the optimal ratios of efficiency across multiple tests that determine a frontier against which peer units are assessed. The scores of the efficiencies of units range from 0 for inefficient units to 1 for efficient units. Partitioning the 0–1 interval helps analyze weak and strong efficiency points and frontier dynamics (Daneshvar & Adesina, 2018a; Daneshvar & Adesina, 2018b). Efficient DMUs are a benchmark for inefficient DMU performance. The efficiency score, therefore, represents an efficient DMU, and DEA acts as a benchmark for the latter. DEA evaluates DMUs by setting efficient ones as standards. The efficient DMUs receive a score of 1, while inefficient DMUs receive a score less than 1. The efficiency scores, therefore, lie between 0 and 1 (Vittal et. al., 2021).

The selection of input-output variables is critical in determining the effectiveness of DEA. However, there are few discussions in the literature on the choice of input or output variables. Most studies just specify the choice of variables in their research. Two major views regarding DEA are the selection of suitable methodologies and data variables (Golany and Roll, 1989). The proper selection of variables is so vital, especially when there are more input-output variables, small weighted models, or inaccuracy within the results (Jenkins and Anderson, 2003). Hence, this reduction of variables is helpful, but there have not been any specific rules to rely on. However, the correlation analysis that has been advocated often leads to variable outcomes, which may not be reproducible. Hence, recently, statistical methods are being employed to refine this variable selection process (Wagner and Shimshak, 2007).

Nataraja and Johnson (2011) reviewed eight different variable selection methodologies and developed guidelines on the selection of an optimal methodology. The methods that come under this category are: statistical testing, efficiency contribution measure, or ECM; bootstrapping; regression-based testing; PCA-DEA; recursive selection; variable reduction; and eigenvalue-based testing. According to that, Pastor et al. (2002) presented a variable relevance assessment method called ECM, or efficiency contribution measure, which designates the variables that affect efficiency. The method is based on two DEA formulations, with and without candidate variables (those to be tested). Then there are some binomial statistical tests that identify whether these variables contribute to efficiency or not. There are also two ways to perform the method: forward selection and backward elimination.

Jenkins and Anderson (2003) suggested that partial correlation should exclude variables that carry minimum information. It denotes information with the help of production unit variance. Variables having zero variance show the same information. Removing highly correlated variables severely affects efficiency scores. The relevance of variables can be analyzed with the support of partial correlation. Highly correlated variables are combined into one input or output with the support of eigenvalues in order to decrease the dimensions of the production function. Daraio and Simar (2007) also recommended integrating those variables through eigenvalues. Ueda and Hoshiai (1997) and Adler and Golany (2001) extended the use of PCA-DEA, whereby PCA is applied to reduce the dimensionality of data. In PCA, weighted linear combinations are used to attain maximum variance among the components while not being interrelated. The main components obtained from linear combinations of the original variables are arranged in order of importance according to the magnitude of the variances. The first component replaces the original variable without any loss of information. This method extends the calculation of efficiency by the DEA method.

Kumar and Singh (2021) present the research on the banking industry, which includes a large number of input/output factors that affect the operational efficiencies. The study applies the standard strategy of DEA for finding out the most influencing input factors and simplifies the model by reducing the redundant variables. The result of this approach turned out to be more accurate performance evaluation of the banks in addition to a more realistic assessment of their operational efficiencies. This, in turn enhances the outcomes of DEA analysis for financial decision-making. Zhang and Wang (2022) establish that variable reduction techniques are some of the most important techniques in DEA studies given the complication of data. They compare factor analysis with principal component analysis; from the comparisons, they established that the latter technique allows for better retention of the important information, hence enhancing the efficiency of the DEA models. The finding is beneficial for both researchers and practitioners in DEA in terms of selecting appropriate variable reduction techniques.

Khan and Ghaffar (2023) discussed using machine learning in DEA to advance variable analysis. Thereafter, insignificantly contributing input variables were identified, and a DEA model was reduced using a machine learning algorithm. The resultant findings showed that machine learning improves the accuracy of variable reduction and uncovers hidden patterns in the data, thus opening new vistas toward dealing with data complexity using modern techniques. Mohammad and Yusof (2020) applied the aggregation method in overcoming the complexity of DEA analyses. This effective combination of DEA with multivariate analysis reduces variant features of input/output variables, and the resulting outputs are more stable and efficient. With this hybrid approach, the DEA model attains a greater degree of flexibility and reliability while performance evaluation is carried out. Narasimhan and Ramakrishnan (2019) presented a solution for reducing the high-dimensionality problem in the DEA analysis by removing extraneous variables. The findings of this dimension reduction demonstrated that it simplified the analysis while improving the accuracy and reliability of the DEA results, offering academics and practitioners practical direction in resolving the model's complexity.

Data Envelopment Analysis and variable selection are among the most important methods in assessing the performances of organizations and decision-making units. DEA is a nonparametric technique that computes the relative efficiency among DMUs working with different configurations of inputs and outputs. Flexible analysis enables inefficient units to adopt optimal practices to improve their performance. Selection of appropriate variables is important; it provides the relevance and preciseness of the DEA model. With the selection of proper variables, we are assured that the factors that are influencing efficiency are reflected in the analysis. Irrelevant variables give biased results; hence, the validity of the analysis decreases. In addition, variable selection helps in avoiding a problem of multicollinearity, which does not give clear implications of the results. Therefore, DEA and variable selection are important in performance evaluation. The reason is that DEA aids in determining relative efficiency, whereas variable selection guarantees that the evaluation is based on the most relevant and significant aspects (Ali & Seiford, 2015; Mardani et al., 2021; Olsen & Karp, 2019).

This study proposes a new approach, variable combinations (VCs)-DEA, which reduces the number of variables in the DEA analysis by considering 21 variable combinations. Variable selection is done irrespective of their order. There are two major factors involved in this model: a reduction in variables factor that would increase the focus towards influential variables and an averaging calculation factor that creates representative efficiency scores. It also compares this with the original DEA method, which does not reduce variables, hence filling the gap in the existing methodologies by offering a systematic approach aimed at improving the accuracy of efficiency results.

The objective of this study is to suggest a variable combination method for subtracting the number of variables that will be utilized in implementing the DEA method. We applied 21 variable combinations (VCs). A combination is a mathematical model that specifies the number of proper regulations in an aggregation of variables in which the sequence of the selection does not matter. The idea of combinations is very helpful in determining the number of subsets that can be constructed through a finite variable set. In the combination concept, we will be able to choose the variable subsets in whatever order. Furthermore, we will use the average efficiency score to get the optimal solution. The proposed method (VCs-DEA) treats the variable reduction factor and the average calculation factor to obtain the final result of the efficiency value. These two factors cause the strength of the efficiency value to be more accurate. We compare our proposed method with two existing methods (without and with variable reduction). The existing method, the original DEA, did not utilize variable reduction as (Anouze and Hamad, 2019) had done. Another existing method with variable reduction, the analysis of PCA-DEA, or principal component analysis, was independently expanded by Ueda and Hoshiai (1997) and Adler and Golany (2001).

This research contribution of the proposed method of variable selection research in DEA is as follows: (i) The variable selection approach can improve the measurement of efficiency for the concerned unit by selecting only relevant variables to be considered in the analysis; (ii) Dimensionality Reduction: variable selection has the effect of reducing the number of inputs and outputs that would have been used in the DEA analysis, thus facilitating interpretation and diminishing model complexity; (iii) Relevant Variable Screening: By this study, it will be possible to find out those variables that most contribute to efficiency and performance, therefore giving more transparent insight for decision makers in managing the analyzed unit; (iv) DEA Model Customization: Using the variable selection method will help in customizing the DEA model to suit the unit under analysis much better for relevance and applicability of the model; (v) Better Model Validity: Appropriate variable selection can improve validity in results of analysis and give more confidence to stakeholders that the findings can be relied upon; (vi) Policy Development: The result of this research work will offer useful information for developing policies and strategies to improve efficiency in the organization or sector where the study applies; (vii) New Knowledge in DEA: This research could generate new knowledge in the DEA methodology, especially in the domain of variable selection techniques that may not have been as widely discussed so far, thus opening an avenue for further research in this area; and (viii) Efficiency Analysis Best Practice: Such best practice in applying DEA integrated with variable selection theory will give guidelines for future researchers and practitioners. Therefore, for such contributions, the research on methods of variable selection in DEA is bound to have positive implications in both academic and practical contexts.

2. Methods

2.1. Performance Evaluation

The capacity of a company to assess business performance is a prerequisite for growth and development. Performance evaluations aim to do two things: (i) evaluate an organization's current internal operations; and (ii) compare an organization's performance to industry norms and best practices. Because of this, a company will be in a better position to: (i) assess its benefits and drawbacks; (ii) better organize its operations to meet the demands and desires of clients; and (iii) recognize business possibilities to develop new products, services, and processes that will enhance operations and activities (Putri et al., 2017).

Two popular strategies for simultaneous advancements in methodology are benchmarking and performance evaluation. The methods employed must be especially important if there are no established norms or criteria for the estimation. Verifying the ratio between decision-making units (DMUs) is the main purpose of benchmarking. DMUs include organizations such as enterprises, associations, projects, corporations, etc. Providing information for business decision-making is the aim of performance measurement while continually monitoring the economy and efficiency of the business's operations. Performance evaluation is a commonly employed technique to enhance organizational procedures. If criteria or benchmarks are not provided for evaluation, this approach becomes crucial (Putri & Kusoncum, 2020).

The fruitfulness of an organization relies on a majority of elements, in which performance evaluation is a principal significant to confirm the output realized. Performance evaluation is the manner of specifying the efficiency activity in the companies (Boland, et. al., 2017). The popular-used methods for performance evaluation, such as cost-benefit analysis, regression analysis, ratio analysis, AHP (analytic hierarchy process), fuzzy comprehensive evaluation, the Delphi method, BSC (balanced scorecard), MCDM (multiple-criteria decision making), and data envelopment analysis (DEA). The usefulness in input-output indicates an inconsiderable connection, conducting to complexity in appraising the company performance. DEA can define qualitative problems through quantitative analysis, convert subjective decisions into objective decisions, and extend unbiased weighting or judgments for aggregation. This is different from the other methods discussed above, which require a subjective assessment of indicator weights and create dimensionless data (Cooper et. al., 2004). Moreover, DEA is a convenient and useful multi-criteria evaluation method that is applied in various groups (Olesen and Petersen, 2016). Government ministries, transportation schemes, training institutions, and others, often apply DEA in evaluating the performance of institutions or units (Shao et al, 2021).

2.2. Input-Oriented VRS DEA Envelopment Method

Data Envelopment Analysis (DEA) is a linear programming technique. This method is applied to measure performance in an integrated model. In certain performance evaluations, input and output parameters are employed. Inputs are among the factors that must be minimized. Inputs include things like labor, materials, costs, and other things. The outputs are one aspect that needs optimization. Some examples of outputs include profit and sales. Inputs and outputs undergo categorization before the DEA is implemented. In the estimate, each entity, procedure, and business activity is demonstrated by DEA applying decision-making units (DMUs). There are two different ways that inefficient DMUs are pushed to the periphery of the efficient DMU criteria. There are two main approaches to this criterion, either output- or input-oriented: (i) steps that decrease input to maximize output at current levels; and (ii) steps that increase output to minimize input to maximize output at current levels (Putri et al., 2023).

$$Min\ \theta \quad \text{subjected to} \tag{1}$$

$$\sum_{j=1}^{n} \lambda j\, Xij \;\leq\; \theta\, Xijo\ ;\ \forall i \tag{2}$$

$$\sum_{j=1}^{n} \lambda j\, Yij \;\leq\; \theta\, Yrjo\ ;\ \forall r \tag{3}$$

$$\sum_{j=1}^{n} \lambda \mathrm{j} = 1 \tag{4}$$

$$\lambda j \;\geq\; 0\ ;\quad \forall j, \theta\ free$$

Constant returns to scale (CRS) is the first concept examined by the DEA model. This concept aims to avoid the possibility that distinct DMUs operate at various scales. The VRS (variable returns-to-scale) idea was developed by Banker et al. (1984). In order to overcome this shortcoming, the model was created with the requirement that a DMU can only be associated with other DMUs with identical dimensions. Presume there are n DMUs ($aj = 1, \dots, n$) that utilize m inputs ($xij, i = 1, \dots, m$) and produce s outputs ($yrj, j = 1, \dots, s$) in order to introduce the VRS-DEA model. DMU j_0's technical efficiency is evaluated by DEA in connection with n peer groups of input and output DMUs. Equations (1) through (4) use the DEA formalization to assess DMU j0 based on VRS, where the efficiency of DMU j_0 is equal to the optimal value of u. This concept is considered to be input-oriented (Anouze and Hamad, 2019). Explanation of DEA symbols presented in Table 1 (Putri et al., 2024).

Table 1. Explanation of DEA symbols

Variable	Explanation	Variable	Explanation
DMUj	DMUs are decision-making units (j = 1, …., n)	s	Total number of outputs
X_2	Number of quarrying workers (persons)	DMUo	One of the n DMUs under consideration
X_3	Compensation of quarrying establishment workers (million rupiahs)	Xi_0	DMUo's ith input (i = 1,..., m)
Y_1	Volume of production of quarrying establishment (m^3)	Yr_0	DMUo's rth input (r = 1,..., s)
n	Number of DMUs	λj	Weights not known (j = 1,..., n corresponds to the DMU number)
Xij	jth DMU with ith input (j = 1, …., n; i = 1, …, m)	θ	A decision variable or DEA efficiency score
m	Total number of inputs	$\theta *$	Optimal solution or value
Yrj	Outputs of the jth DMU (r = 1, …, s; j = 1, …., n)		

2.3. Combinations and Subsets

A mathematical model can be defined as a combination. This model indicates the quantity of appropriate rules in a group of items where the selection order poses no issues. In the combination concept, it will be able to choose the items in whatever order. Specified a set S, let 2^S indicate the set of all subsets of S. The 2^S is not a number but a set. For model, $2\ \{a,b,c\} = \{\emptyset,\{a\},\{b\},\{c\},\{a,b\},\{a,c\},\{b,c\},\{a,b,c\}\}$ (Sagan, 2020). The common formulation for the combinations number of n items utilized r in a period is present in Equation (5).

$$\binom{n}{r} = \frac{(n)r}{r!} = \frac{n!}{(n-r)!\,r!} \tag{5}$$

The number of subsets that can be created using a finite set A may be found quite easily with the use of combinations. It should be considered that altering the sequence set item reserves does not produce distinct sets when calculating the number of subsets; for instance, the sets {a, b, e} and {e, a, b} consist of the same set of letters.

The number of unique subsets found by collecting none at all (sets of zero) plus the number of distinct subsets found by collecting one item, two items, and so on together reflect the total number of subsets of a set A containing n items. Lastly, gathering all n items (set A itself) yields the number of distinct subsets. The total number of distinct subsets of A is expressed as shown in Equation (6) (Springer, 2022).

$$\sum_{r=0}^{n} \binom{n}{r} = \sum_{r=0}^{n} \frac{n!}{(n-r)!\,r!} \tag{6}$$

2.4. Research Methodology

The techniques that this study uses to solve the problem include steps 1 (research design and definition), 2 (preparation, data collection, and assessment), 3 (data processing), 4 (result analysis), and 5 (conclusion). Phases involved in data collection, evaluation, and preparation include: (i) classifying data for Indonesian quarrying (IQ) input and output; (ii) assigning IQ's input, output, and decision-making unit (DMU) data; (iii) assigning constraint data; and (iv) assigning efficiency results based on DMUs. Results analysis consists of: (i) efficiency score results; (ii) efficient and inefficient DMUs; (iii) comparison between existing and proposed methods; and (iv) DMU classification. Fig. 1 displays the research method's flowchart.

Steps 1: Research Design and Definition

↓

Step 2: Preparation, Data Collection, and Assessment

(i) Classifying data for Indonesian quarrying (IQ) input and output;
(ii) Assign IQ's input, output, and DMU (decision-making unit) data. These consist of three inputs (X_1, X_2, X_3), two outputs (Y_1, Y_2), and twelve DMUs (S, SA, G, LS, Q, M, C, So, F, Gr, K, dan O).
(iii) Assign constraint data. There are 6 constraints (Input-X_1, Input-X_2, Input-X_3, Output-Y_1, Output-Y_2, $\sum\lambda$).
(iv) Assign efficiency results based on DMUs.

↓

Step 3: Data Processing

(i) Specify number of subsets and type of subsets. This study applied the combinations number of n = 5 variables (X_1, X_2, X_3, Y_1, Y_2) and utilized r = 3 in a period (2, 3, and 4).
(ii) Specify 21 types of variable combination (VCs). These VCs applied for establishing efficiency score.
(iii) VCs-VRS DEA Spreadsheet Model. Utilizing the solver's feature to provide efficiency scores, the input-output data were organized in an MS Excel spreadsheet.

↓

Step 4: Results Analysis

(i) Efficiency Score Results.
(ii) Efficient and Inefficient DMUs.
(iii) Comparison between Existing and Proposed Method.
(iv) DMUs Classification.

↓

Step 5: Conclusion

Fig. 1. Flowchart of Research Method

The basic idea of applying mathematical combinations in variable selection within DEA involves exploring possible subsets of input and output variables in finding the combination that maximizes the results of the efficiency analysis. The following are the steps involved in DEA variable selection using mathematical combinations and subsets:

1. List of input and output variables: A list of input and output variables is required in the DEA model. The study implemented 3 inputs, such as number of quarrying establishments, number of quarrying workers, compensation of quarrying establishment workers, and 2 outputs, such as volume of production of quarrying establishments and production value of quarrying establishments. The input variables comprise X_1, X_2, and X_3. The output variables comprise Y_1 and Y_2.
2. Combinations Formation: The use of combinations can form several subsets of input and output variables. Combinations allow different subsets to be selected without regard for the order of selection. For n input or output variables, the total number of combinations that might be formed from r variables is given by Equation (5). This study applied the combination number of n = 5 variables (X_1, X_2, X_3, Y_1, Y_2) and utilized r = 3 in a period (2, 3, and 4). There are ten subsets created by using r = 2 and r = 3, respectively. There are five subsets when r = 4. The subsets are then utilized to run DEA. DEA uses input and output variables (mix variables) in its

implementation. Therefore, the selected subsets are those that have both variables (heterogeneous subsets). Homogeneous subsets will not be used in this study. In total, 25 subsets have been created in this research. It consists of 21 heterogeneous subsets and 4 homogenous subsets.

3. Each of these variable combinations is tested in the DEA model to find out the change in efficiency when different variables are in use. This allows for finding the most suitable input/output combination when any relative efficiency has to be evaluated for the DMU under consideration.
4. Sensitivity Analysis: Sensitivity analysis can be made by evaluating various combinations to find which ones most affect the results of DEA and which ones are not that important.

Fig. 2 presents the steps in DEA variable selection using mathematical combinations and subsets.

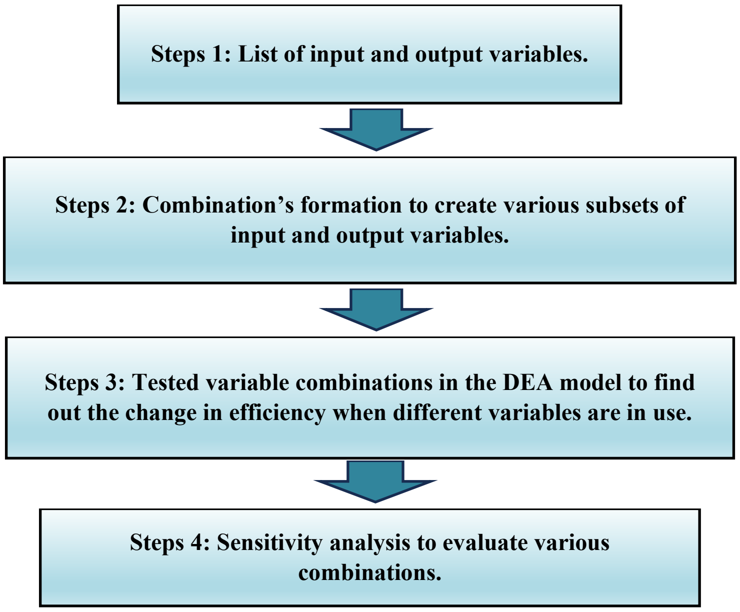

Fig. 2. Steps in DEA Variable Selection using Mathematical Combinations and Subsets

2.5. DEA Variable Selection Criteria Based on Combinations and Subsets in Mathematical Models

DEA can be performed with the use of input and output variables. DEA is a non-parametric method originally used to assess the relative efficiency of DMUs with multiple inputs and outputs. Inputs denote the resources consumed by the DMU, and outputs are the outcome of the utilization of the same resources. DEA assesses how well the DMU generates outputs from its inputs. According to combinations and subsets in the mathematical model, there are two kinds of variables: one based on heterogeneous subsets and another from homogeneous subsets. Heterogeneous Subset-Based Variables: The chosen variables are of heterogeneous subsets, where the input and output variables are selected from a group of data that vary in character or form. The heterogeneous subsets allow for capturing greater variation among DMUs, thus even more comprehensively and accurately enabling an analysis to determine relative efficiency between DMUs. Variable Reduction in Homogenous Subset: A homogeneous subset is a group of variables that represent very similar or identical properties or characteristics. Variables in DEA that are homogeneous—there are no significant variations between DMUs—do not provide much information in the efficiency analysis since they are not helpful enough to distinguish the performance between DMUs. Because of this, variables that are too homogeneous are usually removed or not used since their contribution to the efficiency analysis is low. As a result, input-output variables that differ (heterogeneous) give more informative results in DEA. Conversely, those variables that are uniform (homogeneous) are not utilized because they might be incapable of distinguishing efficiency between the units under consideration (Chen & Delmas, 2022; Khezrimotlagh et. al., 2019; Liu et. al., 2016).

3. Results and Discussion

3.1. Input and Output Data of Quarrying Establishment

The proposed method (the average VCs-VRS DEA envelopment method) can simply be performed by applying a case study. We consider the quarrying establishment data, which consists of three inputs and two outputs, as shown in Table 2. There are 12 kinds of materials and their DMU (decision-making units), as shown in Table 3. The input and output data of quarrying establishments by kind of materials in 2020 are shown in Table 4.

Table 2. Input and output variables for quarrying establishments

Data	Variable	Explanation
Input	X_1	Number of quarrying establishments (units)
	X_2	Number of quarrying workers (persons)
	X_3	Compensation of quarrying establishment workers (million rupiahs)
Output	Y_1	Volume of production of quarrying establishment (m^3)
	Y_2	Production value of quarrying establishment (million rupiahs)

Table 3. Kind of materials and DMUs

DMUs	Kind of Materials	DMUs	Kind of Materials
S	Sand	C	Clay
SA	Stone and Andesite	So	Soil
G	Gravel	F	Feldspars
LS	Lime-Stone	Gr	Granite
Q	Quartz	K	Kaolin
M	Marble	O	Others

Table 4. Input and output data of quarrying establishment

DMUs	Kind of Materials	X_1	X_2	X_3	Y_1	Y_2
S	Sand	60,759	152,682	1,245,282	67,437,616	4,954,429
SA	Stone and Andesite	33,426	87,465	888,574	37,934,605	3,912,987
G	Gravel	5,914	18,390	243,810	14,802,482	1,056,362
LS	Lime-Stone	2,424	9,999	217,070	9,718,944	1,040,036
Q	Quartz	47	894	27,045	1,875,610	165,769
M	Marble	668	3,339	72,477	187,793	335,564
C	Clay	342	3,429	75,596	3,555,291	215,826
So	Soil	930	2,284	46,214	3,133,320	203,733
F	Feldspars	8	316	7,576	110,526	24,669
Gr	Granite	97	1,590	155,602	3,557,268	800,333
K	Kaolin	10	456	18,400	629,247	72,956
O	Others	8,599	39,033	225,297	2,876,868	652,151

3.2. Number and Type of Subsets

The number of combinations is based on 5 variables (X_1, X_2, X_3, Y_1, Y_2) by taking 3 kinds of r (2, 3, and 4) as shown in Table 5. This study applied the combination number of n = 5 variables (X_1, X_2, X_3, Y_1, Y_2) and utilized r = 3 in a period (2, 3, and 4). There are ten subsets created by using r = 2 and r = 3, respectively. There are five subsets when r = 4. Table 4 provides explanations for each type of subset.

There are 21 heterogeneous subsets and 4 homogenous subsets*. The homogenous subsets ((X_1, X_2, X_3)*; (X_1, X_2)*; (X_1, X^3)*; (X_2, X_3)*; and (Y_1, Y_2)*) do not represent the mix variables (input and output variables) at the same time. Therefore, those subsets will not be used as the variable combination for efficiency score calculation.

Table 5. Number and type of subsets

No.	Number of Subsets	Type of Subsets
1.	$\binom{5}{2}^{\cdot} = \frac{(5)2}{2!} = \frac{5!}{(5-2)!\,2!} = 10$	(X_1, Y_1); (X_2, Y_1); (X_3, Y_1); (X_1, Y_2); (X_2, Y_2); (X_3, Y_2); (X_1, X_2)*; (X1, X_3)*; (X_2, X_3)*; (Y_1, Y_2)*
2.	$\binom{5}{3}^{\cdot} = \frac{(5)3}{3!} = \frac{5!}{(5-3)!\,3!} = 10$	(X_1, X_2, Y_1); (X_1, X_3, Y_1); (X_1, X_3, Y_1); (X_1, X_2, Y_2); (X_1, X_3, Y_2); (X_2, X_3, Y_2); (X_1, Y_1, Y_2); (X_2, Y_1, Y_2); (X_3, Y_1, Y_2); (X_1, X_2, X_3)*
3.	$\binom{5}{4}^{\cdot} = \frac{(5)4}{4!} = \frac{5!}{(5-4)!\,4!} = 5$	(X_1, X_2, X_3, Y_1); (X_1, X_2, X_3, Y_2); (X_1, X_2, Y_1, Y_2); (X_1, X_3, Y_1, Y_2); (X_2, X_3, Y_1, Y_2)

3.3. Types of Variable Combination for Efficiency Score Calculation

There are 4 types of VCs based on the combination of variables, namely: VCs – 2 variables, VCs – 3 variables, VCs – 4 variables, and VCs – 5 variables. Type of VCs – 2 variables consisting of VC_{-16}, VC_{-17}, VC_{-18}, VC_{-19}, VC_{-20}, and VC_{-21}. Type of VCs – 3 variables consisting of VC_{-5}, VC_{-6}, VC_{-7}, VC_{-9}, VC_{-10}, VC_{-11}, VC_{-13}, VC_{-14}, and VC_{-15}. Type of VCs – 4 variables consisting of VC_{-2}, VC_{-3}, VC_{-4}, VC_{-8}, and VC_{-12}. Type of VCs – 5 variables have only one subset, namely VC_{-1}. The variable combinations for each type of VCs are presented in Table 6.

Table 6. Types of variable combination (VCs) for efficiency score calculation

Input-Output		Types of Variable Combinations (VCs)										
		VC_{-1}	VC_{-2}	VC_{-3}	VC_{-4}	VC_{-5}	VC_{-6}	VC_{-7}	VC_{-8}	VC_{-9}	VC_{-10}	VC_{-11}
Input		X_1	-	X_1	X_1	X_1	-	-	X_1	-	X_1	X_1
		X_2	X_2	-	X_2	-	X_2	-	X_2	X_2	-	X_2
		X_3	X_3	X_3	-	-	-	X_3	X_3	X_3	X_3	-
Output		Y_1	Y_1	Y_1	Y_1	Y_1	Y_1	Y_1	Y_1	Y_1	Y_1	Y_1
		Y_2	Y_2	Y_2	Y_2	Y_2	Y_2	Y_2	-	-	-	-
Input-Output		**Types of Variable Combinations (VCs)**										
		VC_{-12}	VC_{-13}	VC_{-14}	VC_{-15}	VC_{-16}	VC_{-17}	VC_{-18}	VC_{-19}	VC_{-20}	VC_{-21}	
Input		X_1	-	X_1	X_1	X_1	-	-	X_1	-	-	
		X_2	X_2	-	X_2	-	X_2	-	-	X_2	-	
		X_3	X_3	X_3	-	-	-	X_3	-	-	X_3	
Output		-	-	-	-	Y_1	Y_1	Y_1	-	-	-	
		Y_2	Y_2	Y_2	Y_2	-	-	-	Y_2	Y_2	Y_2	

3.4. VCs-VRS DEA Spreadsheet Model

Utilizing the solver's feature to provide efficiency scores, the input-output data were organized in an MS Excel spreadsheet. This idea is based on a linear programming model. There are four (4) components in this spreadsheet, such as the cells for (a) decision variables (λ and θ); (b) objective function (efficiency, θ); (c) the formulation of reference set (constraints in the right-hand-side); and (d) the formulation of efficiency for DMU under evaluation (constraints in the left-hand-sided) (Putri et al, 2016). Spreadsheet model of VCs-VRS DEA shown in Table 7. Type of variable combination 1 (VC-1) consists of 3 input variables (X1, X2, X3) and 2 output variables (Y1, Y2). The result of efficiency score is as follows: S (1), SA (1), G (1), LS (1), Q (1), M (0.85), C (1), So (1), F (1), Gr (1), K (1), and O (0.56). Spreadsheet model of VC-2 to VC-21-VRS DEA can be done in the same way of VC-1.

3.5. Efficiency Score Results

The average input variable combination-oriented VRS DEA envelopment method is used in this study to determine the efficiency score. The result of efficiency scores based on the types of variable combinations (VC_{-1} to VC_{-21}) and its average as shown in Table 8.

Table 7. VC-1-VRS DEA spreadsheet model

DMUs	X_1	X_2	X_3	Y_1	Y2	λ	Efficiency
S	60,759	152,682	1,245,282	67,437,616	4,954,429	0	1
SA	33,426	87,465	888,574	37,934,605	3,912,987	0	1
G	5,914	18,390	243,810	14,802,482	1,056,362	0	1
LS	2,424	9,999	217,070	9,718,944	1,040,036	0	1
Q	47	894	27,045	1,875,610	165,769	0.73	1
M	668	3,339	72,477	187,793	335,564	0	0.85
C	342	3,429	75,596	3,555,291	215,826	0	1
So	930	2,284	46,214	3,133,320	203,733	0	1
F	8	316	7,576	110,526	24,669	0	1
Gr	97	1,590	155,602	3,557,268	800,333	0.27	1
K	10	456	18,400	629,247	72,956	0	1
O	8,599	39,033	225,297	2,876,868	652,151	0	0.56

Constraints	Reference set		DMU under Evaluation	12	Efficiency 0.56
Input-X_1	60	≤	566		
Input-X_2	1,080	≤	2,831		
Input-X3	61,444	≤	61,444		
Output-Y_1	2,325,584	≥	187,793		
Output-Y_2	335,564	≥	335,564		
∑λ	1	=	1		

Table 8. Results of efficiency score

DMUs	Efficiency Score (VCs)											Average Input VCs-VRS DEA
	VC-1	VC-2	VC-3	VC-4	VC-5	VC-6	VC-7	VC-8	→	VC-20	VC-21	
S	1	1	1	1	1	1	1	1		1	1	1
SA	1	1	1	1	1	1	1	0.90		1	1	0.96
G	1	1	1	1	1	1	1	1		0.47	0.89	0.90
LS	1	1	1	1	1	1	1	1		0.82	0.98	0.98
Q	1	1	1	1	1	1	1	1		0.61	1	0.94
M	0.85	0.25	0.85	0.25	0.06	0.25	0.85	0.12		0.25	0.85	0.38
C	1	0.46	1	0.46	0.28	0.46	0.71	1		0.18	0.49	0.55
So	1	0.62	1	0.62	0.09	0.62	1	1		0.27	0.75	0.66
F	1	1	1	1	0.37	1	1	1		1	1	0.97
Gr	1	1	1	1	1	1	1	1		1	1	0.97
K	1	1	1	1	1	1	0.77	1		0.87	0.77	0.95
O	0.56	0.03	0.56	0.03	0.01	0.03	0.56	0.19		0.03	0.56	0.23

3.6. Efficient and Inefficient DMUs

The average efficiency score of an efficient DMU is 1, but the average efficiency score of an inefficient DMU is less than 1. Based on the average input VCs - VRS DEA envelopment method, the results of the DMUs are both inefficient and efficient, as seen in Table 9. There are 1 efficient DMU out of 12 (8%) and 11 inefficient quarrying DMUs out of 12 (92%).

Table 9. Efficient and inefficient DMUs based on Average Input VCs-VRS DEA

DMUs	Average Efficiency Score	Remark	DMUs	Average Efficiency Score	Remark
S	1	Efficient	C	0.55	Inefficient
SA	0.96	Inefficient	So	0.66	Inefficient
G	0.90	Inefficient	F	0.97	Inefficient
LS	0.98	Inefficient	Gr	0.97	Inefficient
Q	0.94	Inefficient	K	0.95	Inefficient
M	0.38	Inefficient	O	0.23	Inefficient

An efficient DMU manages its inputs and outputs optimally, attaining better technical and allocative efficiency compared with other DMUs. It may be able to maximize output with fewer or more appropriate use of resources. In addition, differences in scale of operation may also be a factor. An efficient DMU may be operating at an optimal scale, while others are operating at too large or too small a scale, resulting in inefficiencies. In addition, exogenous factors, such as especially favorable market conditions, better access to technology, or regulatory support, may also influence the efficiency of a DMU. Other reasons may be internal, such as better management, process innovation, and using higher technology, which might be the reason for the efficiency of only one DMU. It may also be argued that the inefficient DMUs are victims of inefficiencies due to poor management or inefficient use of technology and problems in supply chains, thus worsening their performance. With these factors in view, it thus makes sense why only 1 DMU is efficient, while the rest may face various internal and external challenges that affect their efficiency.

Efficiency is a great tool in many industries, as it ascertains the usage of capital, labor, or raw material resources in the production of maximum output. Efficiency scores generally measure, within an industrial context, the use of an organization or sector in exploiting resources to achieve maximum results with minimal input. This score can be determined using the DEA method. DEA compares the performances of one entity with other entities within the same sector. Efficiency scores are important in the following ways: they pinpoint areas for improvement, indicating less efficient parts of the operations and thus helping in improving; they are a competitive marker in that companies with high efficiency scores tend to have low costs and large profit margins; they optimize resources and increase productivity through optimal use of labor and raw materials; they aid strategic decision-making for the management on investment decisions and technology development; and they enhance sustainability by reducing waste and excess use of resources, hence supporting environmental sustainability (Akhmetova & Suleimanova, 2022; Yusoff & Chedid, 1978; Zhu & Zhou, 2021).

The importance of this research for industry is as follows: (i) The variable selection approach can improve the measurement of efficiency for the concerned unit by selecting only relevant variables to be considered in the analysis; (ii) Dimensionality Reduction: variable selection has the effect of reducing the number of inputs and outputs that would have been used in the DEA analysis, thus facilitating interpretation and diminishing model complexity; (iii) Relevant Variable Screening: By this study, it will be possible to find out those variables that most contribute to efficiency and performance, therefore giving more transparent insight for decision makers in managing the analyzed unit; (iv) DEA Model Customization: Using the variable selection method will help in customizing the DEA model to suit the unit under analysis much better for relevance and applicability of the model; (v) Better Model Validity: Appropriate variable selection can improve validity in results of analysis and give more confidence to stakeholders that the findings can be relied upon; (vi) Policy Development: The result of this research work will offer useful information for developing policies and strategies to improve efficiency in the organization or sector where the study applies; (vii) New Knowledge in DEA: This research could generate new knowledge in the DEA methodology, especially in the domain of variable selection techniques that may not have been as widely discussed so far, thus opening an avenue for further research in this area; and (viii) Efficiency Analysis Best Practice: Such best practice in applying DEA integrated with variable selection theory will give guidelines for future researchers and practitioners (Bai & Sarkis, 2017; Emrouznejad & Yang, 2018; Wang & Wu, 2017).

The 92% inefficiency means that most of the resources utilized by the mining firms are not well utilized. These resources include labor, energy, and capital. Negative impacts brought to companies as a result of these inefficiencies include reduced profitability and increased operational costs because inefficient production consumes more time and resources. Additionally, inefficiency makes companies lose their competitiveness, more so if compared to other companies operating at higher efficiency levels. This will badly affect the company's position in the global market as its capacity to compete at prices would go down.Such inefficiency at the industry level can also translate into commodity price volatility. The aggregate production costs rise when numerous mining companies are inefficient, thus probably having an impact on raw material prices. This, in turn, has an effect on the supply chain, as the companies that make use of the raw materials may face delays in delivery or have to pay a higher price for the raw materials. Secondly, the effect on the environment is also greater because inefficient usage of resources frequently speeds up the exploitation of natural resources, results in higher levels of greenhouse gas emissions, and causes environmental degradation. This inefficiency problem calls for various stakeholders to act. Company management may undertake more sophisticated technologies like automation and digital sensors, which can monitor operational efficiency in real time. They can also take regular audits to identify the points that need rectification. It is also very possible to include mechanisms for incentives to the government for those companies that increase efficiency or apply "green" technologies, with simultaneous stringent regulation that allows more efficient energy use. Investors themselves may contribute by ensuring the control of the companies' activities for efficiency and sustainability and by facilitating the adoption of ESG standards. Communities and non-governmental organizations can also join in the awareness campaigns to ensure better and more environmentally friendly mining practices. The companies may adopt a concrete step to utilize the benchmarking methods against other more efficient mining companies and adopt lean management in order to minimize production waste. Besides, companies can consider the use of renewable energy sources like solar or wind to reduce dependence on expensive and environmentally unfriendly fossil fuels. The positive consequences would be that if these steps are taken, it would allow mining companies to increase their efficiency, lower costs of production, and enhance their competitiveness in the global market with minimal negative impacts on the environment and surrounding communities.

3.7. Comparison between Existing and Proposed Methods

Our proposed method refers to the analysis conducted by Madhanagopal and Chandrasekaran (2014). DEA applies multiples of input and output variables for analyzing efficiency but does not provide guidance in selecting those variables. As a rule, researchers use several methods. Based on this analysis, our proposed method applied variable reduction using the variable combination (VCs) method. VCs is a method to subtract the number of variables that will be utilized in implementing the DEA method. A combination is a mathematical model that specifies the number of proper regulations in an aggregation of variables in which the sequence of the selection does not matter. In the combination concept, we will be able to choose the variable subsets in whatever order. Furthermore, we will use the average efficiency score to get the optimal solution. By applying the average, the results of efficiency will be more accurate. Some of the underlying things are: (a) the average is the central value for the data; (b) the average is referred to by the value of each point in the sequences. Hence, the average can be mentioned to actually reflect the central tendency of the data, and (c) the average is reliable in the sense that it does not differ too much when the repeated sample is taken from a large population for approximation purposes.

Data Envelopment Analysis (DEA) is a non-parametric method applied for the estimation of the relative efficiency of decision units (DMUs) within a multivariate analysis context. Since DEA does not rely on any specific assumptions concerning data distribution, in contrast to parametric methods, there are no requirements regarding formal statistical tests, such as t-tests or ANOVA, in the context of the results from using DEA. However, even though DEA does not require statistical tests by itself, there are considerations to be validated: results validation, sensitivity of the analysis, confidence intervals, and comparison to other methods. Thus, even though DEA in itself does not require statistical tests, additional approaches can be used to enhance the analyses and give better context for the results obtained (Banker et al., 1984; Charnes et al., 1978).

This research compares the proposed method (VCs-DEA) with two existing methods. The existing method, the original DEA, did not utilize variable reduction as Anouze and Hamad (2019) had done. Another existing method with variable reduction, the analysis of principal component analysis (PCA)-DEA, was separately extended by Ueda and Hoshiai (1997) and Adler and Golany (2001). Table 10 presents the efficiency scores of the existing (Original DEA and PCA-DEA) and proposed methods (VCs-DEA). Efficient DMUs have an efficiency score of 1, while inefficient

DMUs have an efficiency score of less than 1. Based on the value of the efficiency score (ES), it can be determined the number of efficient DMU, the number of inefficient DMU, efficient DMU (%), and inefficient DMU (%). VCs-DEA has (i) the smallest number of efficient DMUs and percentage of efficient DMUs (1 and 8%, respectively); and (ii) the largest number of inefficient DMUs and percentage of inefficient DMUs (11 and 92%, respectively). Original DEA has (i) the largest number of efficient DMUs and percentage of efficient DMUs (10 and 83%, respectively); and (ii) the smallest number of inefficient DMUs and percentage of inefficient DMUs (2 and 17%, respectively). PCA-DEA has an intermediate value between VCs-DEA and Original DEA, namely the number of efficient DMUs (3), percentage of efficient DMUs (25), number of inefficient DMUs (9), and percentage of inefficient DMUs (75%). Fig. 3 presents the distribution of efficiency scores (ES) of existing (Original DEA and PCA-DEA) and proposed methods (VCs-DEA). Original DEA tends to have DMUs with ES values equal to 1. Both PCA-DEA and VCs-DEA have DMUs with fluctuating ES distributions. However, PCA-DEA has a lower fluctuating ES distribution than VCs-DEA.

Table 10. Efficiency score of existing and proposed methods

DMUs	Existing Method		Proposed Method
	Original DEA	PCA-DEA	VCs-DEA
S	1	0.20	1
SA	1	0.19	0.96
G	1	0.36	0.90
LS	1	0.43	0.98
Q	1	1	0.94
M	0.85	0.20	0.38
C	1	0.46	0.55
So	1	0.61	0.66
F	1	0.37	0.97
Gr	1	1	0.97
K	1	1	0.95
O	0.56	0.03	0.23
Number of Efficient DMU	10	3	1
Number of Inefficient DMU	2	9	11
Efficient DMU (%)	83	25	8
Inefficient DMU (%)	17	75	92

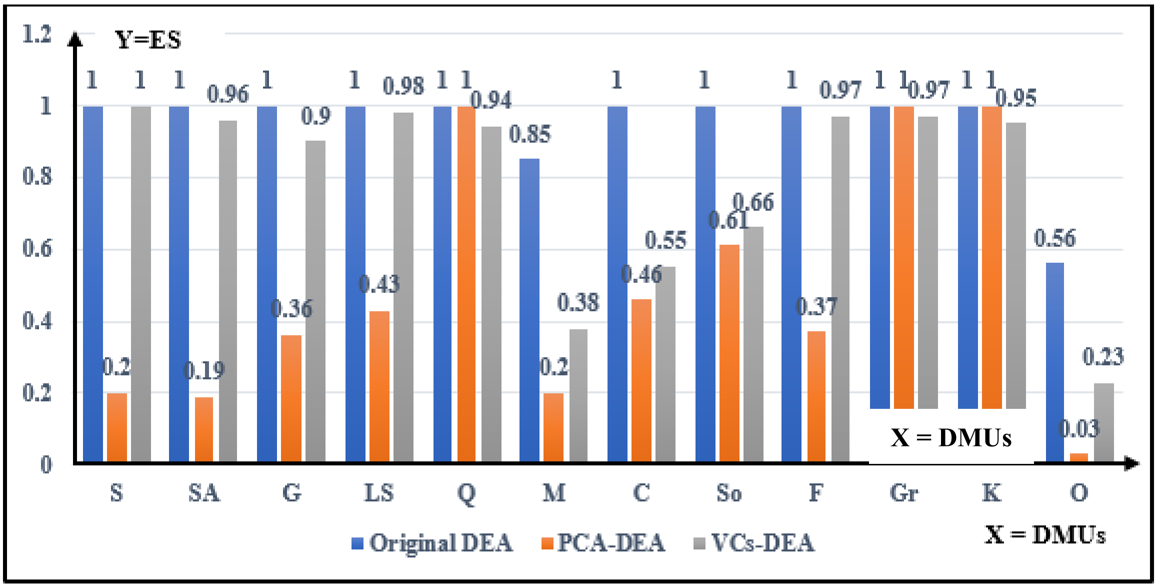

Fig. 3. Efficiency score of existing and proposed methods

The result of the original DEA indicates that there are 10 efficient DMUs out of 12 (83%) and 2 inefficient DMUs out of 12 (17%). The original DEA does not treat variable reduction. This affects the power efficiency results, which causes the number of efficient DMUs to be very large (83%). The result of PCA-DEA indicates that there are 3 efficient DMUs out of 12 (25%) and 9 inefficient DMUs out of 12 (75%). The PCA-DEA applied variable reduction. Therefore, the result of PCA-DEA is more accurate than the original DEA.

The result of VCs-DEA indicates that there is 1 efficient DMU out of 12 (8%) and 11 inefficient quarrying DMUs out of 12 (92%). The VCs-DEA treats the variable reduction factor and the average calculation factor to obtain the final result of the efficiency value. These two factors cause the strength of the efficiency value to be more accurate. These two factors contribute to the accuracy of the efficiency value. Hence, this method is unique or superior. This can be seen in the number of efficient DMUs, which is only 8%. The original DEA has the highest number of efficient DMUs (83%), and PCA-DEA has the highest number of efficient DMUs (25%). The VCs-DEA provides a more accurate efficiency value than the original DEA and PCA-DEA. Therefore, the proposed method is better than the existing method to evaluate the performance of Indonesian quarrying.

Average Input Variable Combinations (VCs)-DEA is more accurate because of its intelligent approach in reducing the input variables without losing the important information. Essentially, VCs-DEA works by taking an average combination of various input variables, which actually simplifies the model but still retains the essence of each input. This step reduces the complexity of the analysis and stabilizes the results since the combination of variables reduces the sensitivity of the model towards unstable variables or outliers. The reduction of the number of variables is carried out with due care to ensure that any loss of information remains as minimal as possible; hence, the efficiency assessment outcome remains accurate. On the contrary, the PCA-DEA method tries to reduce input variables using a statistical approach. While effective in reducing the number of input variables used, this approach, in which PCA selects the number of principal components of input variables to explain most of the variation in data, poses a risk of losing important information that may not be captured by the principal components. For example, when principal components explain only a part of the variation in data, other less important components may also carry information that can be relevant in efficiency assessment. Thus, though PCA-DEA reduces problems of dimensionality and multicollinearity, there is a potential loss of detail, which can reduce accuracy in the final results. Meanwhile, Original DEA does not use any variable reduction techniques and analyzes all available input variables directly. That sounds simple, but the use of too much input can create certain problems—for example, the "curse of dimensionality." If the number of input variables is too large, then the DEA model may turn out to be highly sensitive to small changes in the dataset, which can again result in incorrect efficiency evaluations. Another important point is the fact that great correlation between input variables may lead to problems of multicollinearity, where the model cannot tell how much each variable contributes, and, with respect to that, it makes results unstable and less reliable. One of the major strengths of VCs-DEA is that it can at least partially transcend these problems through the aggregation process. The information redundancy could be reduced and balance the influential strength of each variable in efficiency assessment by aggregating correlated input variables. Let's say, for instance, a firm's efficiency analysis involves up to so many inputs, such as the number of workers, capital, and raw materials. VCs-DEA can combine highly correlated variables into one combination variable, thereby reducing the risk of distortion due to the usage of excessive variables. Variable reduction plays a significant role in striking an optimal balance between the model's complexity and precision. If the number of input variables is large, then a model can run into unsteadiness and a loss of interpretability. By variance reduction, on the other hand, with methods such as VC-DEA, we ensure higher accuracy and stability of the model by trying to minimize the influence of irrelevant or redundant variables. However, in respect to complex efficiency analysis, VCs-DEA outperforms others because it can make a trade-off between model simplification and retaining important information; hence, it could be more accurate and stable than PCA-DEA and Original DEA (García-Sánchez et al., 2019; Cooper et al., 2004; Tavana et al., 2020).

The comparative analysis between methods (VCs-DEA (Average Input Variable Combinations DEA), PCA-DEA (Principal Component Analysis DEA), and Original DEA): VCs-DEA, PCA-DEA, and Original DEA can then indicate a deeper understanding of the operational efficiencies of quarrying companies in Indonesia with different but complementary approaches. The VCs-DEA averages these input variables, which are usually fluctuating—for example, fuel usage or daily productivity affected by weather or technical conditions—to get more stable analysis results. This is very relevant for quarrying companies because the fluctuations of these variables often make the efficiency analysis difficult by just using unstable raw data. Meanwhile, the PCA-DEA method supports an easy analysis by reducing the data dimension, putting attention only on the variables having the most significant effect on the efficiency. PCA-DEA helps companies to focus more on the main factors of influence in such complex quarrying operations because many variables, such as equipment life, production capacity, energy consumption, and labor costs, come into play. For example, if PCA-DEA identifies heavy capacity equipment and operational age as two of the most crucial factors in determining efficiency for a quarry, then management can become more focused on allocating resources or investments to improve the performance of such tools. On the other hand, Original DEA uses all the input and output variables without simplification, thus allowing the companies to get a full picture of the operational efficiency. This will be less effective, however, if there are a lot of irrelevant or collinear variables that actually mask the results of the analysis. Even then, for quarry wanting to see efficiency holistically—from energy consumption and equipment productivity to environmental factors—Original DEA is still useful to get a comprehensive assessment. Comparing these three methods will enable quarrying companies to identify, more precisely, points of inefficiency and design more focused strategies towards optimization of resource use in such a way that the operational efficiency is more effective and that management decisions are based on more in-depth analysis relevant to the company's real situation (Cooper et al., 2007; Emrouznejad & Yang, 2018; Nalbantian & Toroslu, 2015).

The major uniqueness of the comparative analysis between the VCs-DEA, PCA-DEA, and Original DEA methods is that these three methods introduce different dimensions in measuring the operational efficiency of quarrying companies. VCs-DEA provides more stable and consistent results than conventional methods when raw data are used. PCA-DEA helps the company in strategic resource allocation by paying more attention to the most influential variables. Original DEA may view all the important aspects related to mining operations—from energy consumption to environmental impacts—in a holistic manner. Therefore, one of the unforeseen findings/novelties of this analysis is its flexibility in providing options of methods that may be fitted to specific needs in mining companies. In this way, companies will be able to choose the most appropriate one, depending on data stability or focusing on key variables, whether or not they have the need to obtain a holistic view. In this respect, the contribution of the present study is novel in the literature on operational efficiency in the mining sector because firms will be able to make more targeted decisions based on the most relevant analysis method compared to real conditions in the field.

3.8. DMUs Classification

Fig. 4 presents the distribution of DMU average efficiency scores (AES) based on VCs-DEA. The best DMU is S (AES = 1). The least efficient DMU is O (AES = 0.23). Seven DMUs have an average efficiency score in the range of 0.9, namely: LS (AES = 0.98), F (AES = 0.97), Gr (AES = 0.97), SA (AES = 0.96), K (AES = 0.95), Q (AES = 0.94), and G (AES = 0.9). Three DMUs have efficiency values in the range of 0.66–0.38, namely: So (AES = 0.66), C (AES = 0.55), and M (AES = 0.38).

Based on the distribution of DMU average efficiency scores (AES) in Fig. 4, the Indonesia quarrying establishment can be classified into 3 categories, as shown in Table 11. The optimal category (AESR = 0.99 - 1) has a percentage of the DMUs number of 8% (1/12x100%). The other two categories are the middle category (AESR = 0.70–0.98) and the less category (AESR = 0.20–0.69), which have a percentage of the DMUs number of 58% (7/12x100%) and 33% (4/12x100%), respectively.

Fig. 5 presents a comparison of DMU classification based on average efficiency score range (AESR) and percentage (%). The optimal category (AESR = 0.99 - 1) has a percentage of 8%. The medium category (AESR = 0.70–0.98) and low (AESR = 0.20–0.69) have percentages of 58% and 33%, respectively.

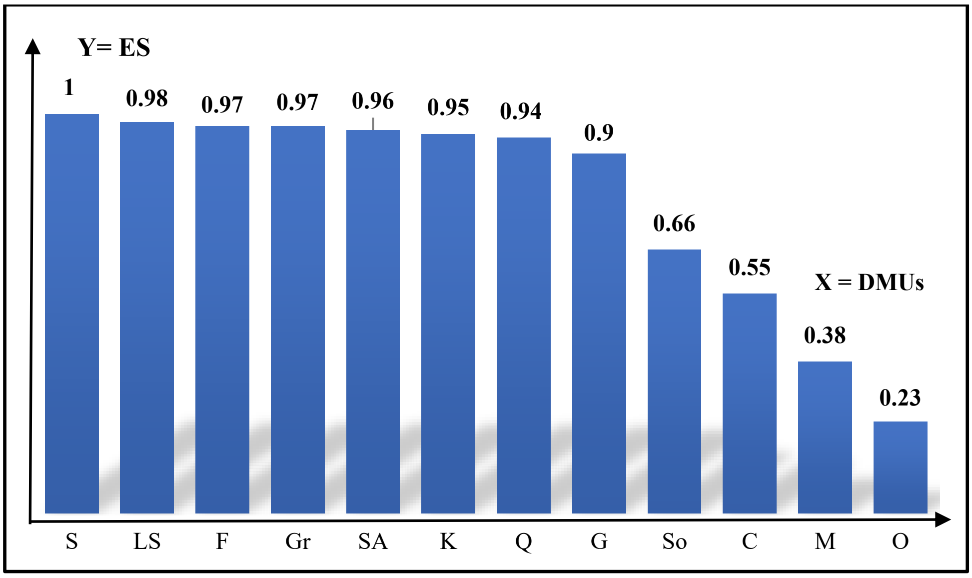

Fig. 4. DMU Average Efficiency Score Based on VCs-DEA

Table 11. DMUs classification

Category	Average Efficiency Score Range (AESR)	DMU	Kind of Materials	Average Efficiency Score (AES)	Percentage (%)
Optimal	0.99 - 1	S	Sand	1	8
Medium	0.70 - 0.98	LS	Lime-Stone	0.976	58
		F	Feldspars	0.970	
		Gr	Granite	0.969	
		SA	Stone and Andesite	0.959	
		K	Kaolin	0.954	
		Q	Quartz	0.937	
		G	Gravel	0.901	
Low	0.20 - 0.69	So	Soil	0.659	33
		C	Clay	0.548	
		M	Marble	0.380	
		O	Others	0.234	

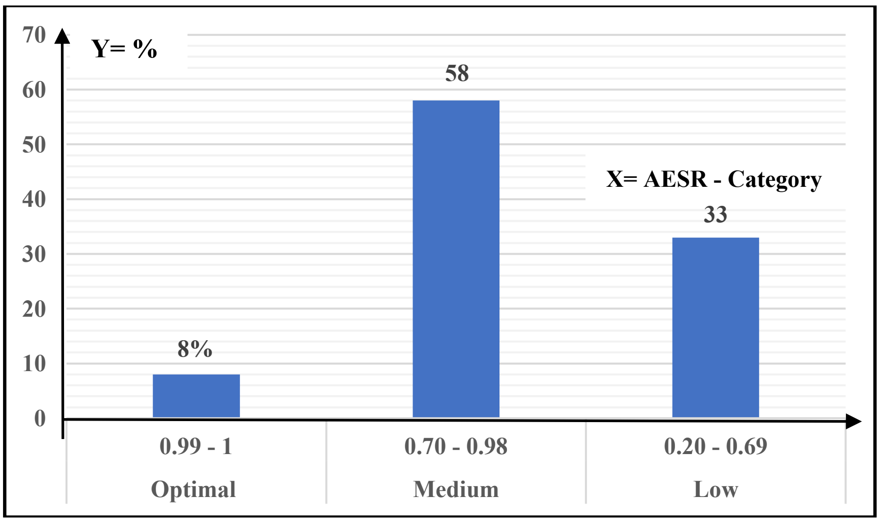

Fig. 5. DMU Classification Comparison Based on AESR and Percentage (%)

Fig. 6 presents the classification of DMUs in terms of AESR, DMU type, and efficiency score (ES). The optimal category (AESR = 0.99 - 1) only has 1 DMU, namely S (ES = 1). The medium category (AESR = 0.70–0.98) consists of 7 DMUs, namely: LS (ES = 0.976), F (ES = 0.970), Gr (ES = 0.969), SA (ES = 0.959), K (ES = 0.954), Q (ES = 0.937), and G (ES = 0.901). The low category (AESR = 0.20–0.69) consists of 4 DMUs, namely: So (ES = 0.659), C (ES = 0.548), M (ES = 0.380), and O (ES = 0.234).

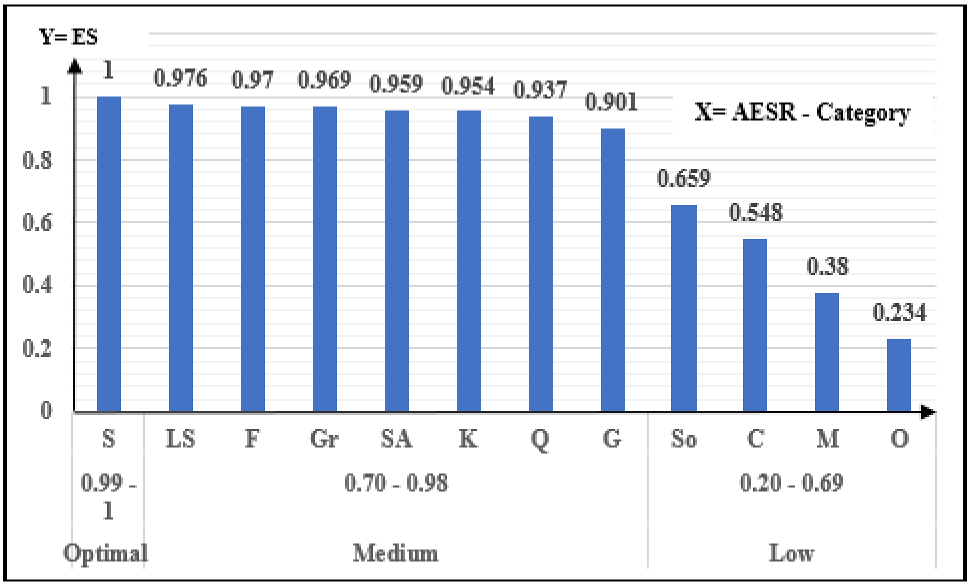

Fig. 6. DMU Classification Comparison Based on AESR, DMU Types, and Efficiency Scores

The comparative analysis between methods (VCs-DEA (Average Input Variable Combinations DEA), PCA-DEA (Principal Component Analysis DEA), and Original DEA) is described as the following: VCs-DEA, PCA-DEA, and Original DEA can then indicate a deeper understanding of the operational efficiencies of quarrying companies in Indonesia with different but complementary approaches. The VCs-DEA averages these input variables, which are usually fluctuating—for example, fuel usage or daily productivity affected by weather or technical conditions—to get more stable analysis results. This is very relevant for quarrying companies because the fluctuations of these variables often make the efficiency analysis difficult by just using unstable raw data. Meanwhile, the PCA-DEA method supports an easy analysis by reducing the data dimension, putting attention only on the variables having the most significant effect on the efficiency. PCA-DEA helps companies to focus more on the main factors of influence in such complex quarrying operations because many variables, such as equipment life, production capacity, energy consumption, and labor costs, come into play. For example, if PCA-DEA identifies heavy capacity equipment and operational age as two of the most crucial factors in determining efficiency for a quarry, then management can become more focused on allocating resources or investments to improve the performance of such tools. On the other hand, Original DEA uses all the input and output variables without simplification, thus allowing the companies to get a full picture of the operational efficiency. This will be less effective, however, if there are a lot of irrelevant or collinear variables that actually mask the results of the analysis. Even then, for quarry wanting to see efficiency holistically—from energy consumption and equipment productivity to environmental factors—Original DEA is still useful to get a comprehensive assessment. Comparing these three methods will enable quarrying companies to identify, more precisely, points of inefficiency and design more focused strategies towards optimization of resource use in such a way that the operational efficiency is more effective and that management decisions are based on more in-depth analysis relevant to the company's real situation (Brodie & Hsu, 2018; García-Sánchez et al., 2019; Sullivan & Mounsey, 2010).

4. Conclusion

The study evaluates the efficiency of Indonesian quarrying operations using three methods: VCs-DEA, PCA-DEA, and Original DEA. Among these, VCs-DEA is identified as the most accurate, classifying 8% of Decision-Making Units (DMUs) as efficient due to its unique approach of variable reduction and averaging, which stabilizes fluctuating inputs. PCA-DEA offers simplification by focusing on key variables, aiding resource allocation and planning, while Original DEA provides comprehensive evaluations but struggles with irrelevant or collinear data. Future research suggests integrating these methods into a unified framework sensitive to real-time data and industry conditions, exploring case studies, and addressing external factors like policies and climate. It emphasizes training and collaboration for practical application, aiming to optimize efficiency and sustainability across mining and other sectors.

Author Contribution: All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

Akhmetova, S. A., & Suleimanova, A. G. (2022). Analyzing the efficiency of green technologies in manufacturing enterprises. *Sustainability*, 14(2), 947, https://doi.org/10.3390/su14020947.

Ali, M. & Seiford, L. M. (2015). The use of data envelopment analysis in performance measurement: A review of the literature. *Journal of the Operational Research Society*, 66(3), 355-370, https://doi.org/10.1057/jors.2012.63.

Anouze, A. L. M., & Hamad, I. B. (2019). Data envelopment analysis and data mining to efficiency estimation and evaluation. *International Journal of Islamic and Middle Eastern Finance and Management*, *12*(2), 169-190, https://doi.org/10.1108/IMEFM-11-2017-0302.

Bai, Y., & Sarkis, J. (2017). A data-driven variable selection approach for performance measurement in supply chain management. *International Journal of Production Research*, 55(10), 2826-2841, https://doi.org/10.1080/00207543.2016.1200312.

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, https://doi.org/10.1287/mnsc.30.9.1078.

Boland L, Légaré F, Perez MM, Menear M, Garvelink MM, McIsaac DI, Painchaud Guérard G, Emond J, Brière N, & Stacey D. (2017). Impact of home care versus alternative locations of care on elder health outcomes: An overview of systematic reviews. *BMC Geriatrics*, *17*(1), 1-20, https://doi.org/10.1186/s12877-016-0395-y.

Brodie, R., & Hsu, C. (2018). Exploring the Relationship Between Mining Policy and Operational Efficiency. *Resources Policy*, 55, 1-10, https://doi.org/10.1016/j.resourpol.2018.02.008.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429-444, https://doi.org/10.1016/0377-2217(78)90138-8.

Chen, Z., & Delmas, M.A. (2022). Measuring corporate social responsibility performance in China: A data envelopment analysis approach. *Journal of Business Ethics*, 180, 713-732, https://doi.org/10.1007/s10551-021-04915-9.

Cooper, W.W., Seiford, L.M., & Zhu, J. (2004). Data Envelopment Analysis. In: Handbook on Data Envelopment Analysis. *International Series in Operations Research & Management Science*, Springer, Boston, MA., 71, 1-39, https://doi.org/10.1007/1-4020-7798-X_1.

Daneshvar, S. & Adesina, K. (2018a). Integrated data envelopment-thermoexergetic optimization framework for multicomponent distillation system with multiexergetic response in the robust parameter design procedures. *Energy Sources*, *40*, 1491–1507, https://doi.org/10.1080/15567036.2018.1477876.

Daneshvar, S. & Adesina, K. (2018b). Modified variable return to scale back propagation neural network robust parameter optimization procedure for multi-quality processes. *Eng. Optim*, *50*, 1352–1369, https://doi.org/10.1080/0305215X.2018.1524463.

Emrouznejad, A., & Yang, G. L. (2018). A survey and analysis of the state-of-the-art in DEA variable selection. *Journal of Productivity Analysis*, 49(1), 1-23, https://doi.org/10.1007/s11123-017-0465-3.

García-Sánchez, I.-M., & Rodríguez-Domínguez, L. (2019). Environmental Performance and Economic Efficiency in the Mining Industry: Evidence from the World. *Resources, Conservation and Recycling*, 148, 1-11, https://doi.org/10.1016/j.resconrec.2019.05.016.

Golany, B., & Roll, Y. (1989). An application procedure for DEA. *Omega*, *17*(3), 237-250, https://doi.org/10.1016/0305-0483(89)90029-7.

Jenkins, L., & Anderson, M. (2003). A multivariate statistical approach to reducing the number of variables in data envelopment analysis. *European Journal of Operational Research*, *147*(1), 51-61, https://doi.org/10.1016/S0377-2217(02)00243-6.

Khan, A. I., & Ghaffar, A. (2023). Reducing input variables in DEA: A novel approach using machine learning techniques. *Operations Research Perspectives*, 10, 100247, https://doi.org/10.1016/j.orp.2023.100247.

Khezrimotlagh, D., Salleh, S., & Mohsenpour, Z. (2019). A refined data envelopment analysis model for measuring the performance of decision-making units. *Computers & Industrial Engineering*, 128, 482-489, https://doi.org/10.1016/j.cie.2018.12.058.

Kraidi, A. A., Daneshvar, S., & Adesina, K. A. (2024). Weight-restricted approach on constant returns to scale DEA models: Efficiency of internet banking in Turkey. *Heliyon*, *10*(10), 1-15, https://doi.org/10.1016/j.heliyon.2024.e31008.

Kumar, A., & Singh, P. (2021). Application of DEA with reduced variables: An empirical analysis in the banking sector. *International Journal of Financial Studies*, 9(4), 45, https://doi.org/10.3390/ijfs9040045.

Liu, W., Zhu, Q., & Shang, J. (2016). A data envelopment analysis model for eco-efficiency evaluation of China's industrial sectors. *Journal of Cleaner Production*, 142, 2802-2811, https://doi.org/10.1016/j.jclepro.2016.10.154.

Mardani, A., Jusoh, A., & Zavadskas, E. K. (2021). Data envelopment analysis and performance evaluation: A literature review. *Operations Research Perspectives*, 8, 100167, https://doi.org/10.1016/j.orp.2021.100167.

Mohammad, A., & Yusof, M. (2020). A hybrid DEA approach for variable reduction and performance evaluation. *Journal of the Operational Research Society*, *71(6)*, 1023-1035, https://doi.org/10.1080/01605682.2019.1672310.

Nalbantian, L., & Toroslu, I. (2015). Assessing the efficiency of mining firms using DEA. *Resources Policy*, 46, 33-45, https://doi.org/10.1016/j.resourpol.2014.09.003.

Narasimhan, K., & Ramakrishnan, K. (2019). Reducing dimensionality in DEA models: A practical approach. *Computational and Mathematical Organization Theory*, *25(4)*, 507-525, https://doi.org/10.1007/s10588-019-09350-1.

Olesen, O. B., & Petersen, N. C. (2016). Stochastic data envelopment analysisi-A review. *European Journal of Operation Research*, *251*(1), pp.2-21, https://doi.org/10.1016/j.ejor.2015.07.058.

Olsen, T. L. & Karp, L. S. (2019). Data envelopment analysis and the role of variables in measuring efficiency: A review. *Journal of Productivity Analysis*, 51(2), 117-135, https://doi.org/10.1007/s11123-019-00523-5.

Putri, E. P. & Kusoncum, C. (2020, October). Performance evaluation using PCA-CRS input oriented DEA method. A case study of East Java exports in Indonesia to ASEAN countries. In *Proceedings of the 2019 International Conference on "Physics, Mechanics of New Materials and Their Applications"*. 445-452, SCI055000. Nova Science Publishers, Inc, https://novapublishers.com/shop/proceedings-of-the-2019-international-conference-on-physics-mechanics-of-new-materials-and-their-applications/.

Putri, E. P., Aduldaecha, S., Putra, B. R. S. P., & Agatha Hannabel Avnanta Puteri, A. H. A. (2024). Performance evaluation of Indonesia's large and medium-sized industries using data envelopment analysis method. *OPSI: Jurnal Optimasi Sistem Industri*, *17*(1), 118-134, https://doi.org/10.31315/opsi.v17i1.11785.

Putri, E. P., Arief, Z., & Yuwono, I. (2023, February). Performance evaluation using input-oriented envelopment DEA method: A case study of micro and small industry in Indonesia. In *Physics and Mechanics of New Materials and Their Applications, 2021 – 2022*. 289-304, SCI055000. Nova Science Publishers, Inc, https://doi.org/10.52305/QLWW2709.

Putri, E. P., Chetchotsak, D., Jani, M. A., & Hastijanti, R. (2017). Performance evaluation using PCA and DEA: A case study of the micro and small manufacturing industries in Indonesia. *ASR: CMU Journal of Social Sciences and Humanities*, *4*(1), 37-56, https://doi.org/10.12982/CMUJASR.2017.0003.

Putri, E. P., Chetchotsak, D., Ruangchoenghum, P., Jani, M. A., & Hastijanti R. (2016). Performance evaluation of large and medium scale manufacturing industry clusters in East Java Province, Indonesia. *International Journal of Technology*, *7*(7), 1117-1127, https://doi.org/10.14716/ijtech.v7i7.5229.

Shao, Q., Yuan, J., Lin, J., Huang, W., Ma, J., & Ding, H. (2021), A SBM-DEA based performance evaluation and optimization for social organizations participating in community and home-based elderly care services. *PloS ONE*, *16*(3), 1-25, https://doi.org/10.1371/journal.pone.0248474.

Springer, 2022. Math Review: Sets, Functions, Permutations, Combinations, and Notation, Available Online at: https://link.springer.com/content/pdf/bbm%3A978-1-4612-3988-8%2F1.pdf, Accessed on January 25, 2022

Sullivan, L. J., & Mounsey, R. A. (2010). The Role of Technology and Innovation in the Mining Sector. *Journal of Cleaner Production*, 18(11), 1186-1195, https://doi.org/10.1016/j.jclepro.2010.01.020.

Tavana, M., Khalili-Damghani, K., & Mina, H. (2020). A new dynamic network DEA model for time-series efficiency measurement using goal-directed benchmarking. *Measurement*, 165, 108145, http://doi.org/10.1016/j.measurement.2020.108145.

Vittal, B., Nellutla, R., & Reddy, M. K. (2021). Selection and analysis of input-output variables using data envelopment analysis of decision making units - Indian private sector banks. *International Journal of Engineering and Advanced Technology*, *10*(5), 119-127, https://doi.org/10.35940/ijeat.E2674.0610521.

Wagner, J.M., & Shimshak, D. G. (2007). Stepwise selection of variables in data envelopment analysis: Procedures and managerial perspectives. *European Journal of Operational Research*, *180*(1), 57-67, https://doi.org/10.1016/j.ejor.2006.02.048.

Wang, J., & Wu, S. (2017). Variable selection for DEA models using a data-driven approach. *International Journal of Production Economics*, 193, 48-56, https://doi.org/10.1016/j.ijpe.2017.07.007.

Yusoff, N. A., & Chedid, R. (2023). The relationship between innovation, efficiency, and firm performance in the Malaysian SMEs. *Journal of Small Business Management*, 61(1), 105-123, https://doi.org/10.1080/00472778.2021.1978420.

Zhang, Y., & Wang, Y. (2022). Variable reduction techniques in DEA: A comparative study of factor analysis and principal component analysis. *Journal of Productivity Analysis*, *58(3)*, 239-255, https://doi.org/10.1007/s11123-022-00780-9.

Zhu, H., & Zhou, Y. (2021). Measuring operational efficiency and its influencing factors in the logistics industry: Evidence from China. *Journal of Transport Geography*, 90, https://doi.org/10.1016/j.jtrangeo.2020.102897.