Journal of Decision Systems

ISSN: 1246-0125 (Print) 2116-7052 (Online) Journal homepage: www.tandfonline.com/journals/tjds20

A machine learning driven decision support system for evaluating port performance: development and validation

Leonardo Leoni , Xiaotian Xie , Guoqing Zhao , Yi Wang & Filippo De Carlo

To cite this article: Leonardo Leoni , Xiaotian Xie , Guoqing Zhao , Yi Wang & Filippo De Carlo (2026) A machine learning driven decision support system for evaluating port performance: development and validation, Journal of Decision Systems, 35:1, 2682227, DOI: 10.1080/12460125.2026.2682227

To link to this article: https://doi.org/10.1080/12460125.2026.2682227

JOURNAL OF DECISION SYSTEMS
2026, VOL. 35, NO. 1, 2682227
https://doi.org/10.1080/12460125.2026.2682227

OPEN ACCESS Check for updates

A machine learning driven decision support system for evaluating port performance: development and validation

Leonardo Leoni[a], Xiaotian Xie[b], Guoqing Zhao[c], Yi Wang[d] and Filippo De Carlo[e]

[a]Department of Theoretical and Applied Sciences (DiSTA), eCampus University, Novedrate, Italy; [b]Newcastle University Business School, Newcastle University, Newcastle, UK; [c]School of Management, Swansea University, Swansea, UK; [d]Department is Business School, University of Bedfordshire, University Square, Luton, UK; [e]Department of Industrial Engineering, University of Florence, Firenze, Italy

ABSTRACT

Limited research has examined how ports' big data analytics capability (BDAC) is associated with operational and sustainable performance. In response, this study develops and validates a decision support system (DSS) that integrates expert judgements, fuzzy set theory, unsupervised machine learning (ML), Decision Trees, and Bayesian Network analysis. Data were collected through a Likert-scale questionnaire completed by 158 respondents from 40 major ports. The responses were aggregated using an improved Similarity Aggregation Method, and K-Means clustering was applied to classify ports into performance groups. Decision Trees were then developed to identify performance clusters and key improvement areas, while a Bayesian Network was used to explore relationships among BDAC, port operational performance, and port sustainable performance. The results indicate that ports with stronger BDAC generally achieve better operational and sustainable performance, although other contextual factors may also play important roles.

ARTICLE HISTORY
Received 12 April 2026
Accepted 26 May 2026

KEYWORDS
Big data analytics capability; port operations; port management; decision-making

1. Introduction

Decision-making processes are considered critical factor in the success of any organisation (Bregar, 2024) and they are strongly influenced by the ability to evaluate the organisational performance. However, supply chains have become progressively more challenging due to globalisation (Tiwari et al., 2023), while high levels of volatility and complexity in organisational performance have made decision-making process more difficult (Jraisat et al., 2021). Within this context, Decision Support Systems (DSSs) play a pivotal role in decision-making by highlighting critical factors, defining more convenient strategies, and assessing firm performance. Furthermore, they could be employed at different decision levels (Meski et al., 2021). Consequently, the literature offers a wide variety of DSSs, among which the artificial intelligence-based or machine learning (ML)-based (Cantini et al., 2022; Olan et al., 2024; Tirkolaee & Torkayesh, 2022) and expert-based systems (Bai et al., 2019; Cinelli et al., 2022) are very popular. ML-based DSSs usually employ typical ML algorithms (e.g. neural network) to define the current conditions or predict future conditions of parameters. ML has become increasingly popular because of its analytical power and advances in computational devices (Uddin et al., 2024). However, enterprises frequently face difficulties in adopting ML, due to the lack of required unbiased data and the need to continuously retrain the model (DeBrusk, 2018). Furthermore, data may be of poor quality (e.g. due to manual entry errors) or may be stored in many distinct databases, leading to long pre-processing times (Rapaccini et al., 2022). Finally, ML must be used for measurable parameters or variables; otherwise, it will be abused when data are artificially generated by using mathematical model (Cantini et al., 2022). On the other hand, EDSSs are often characterised by the adoption of multi-criteria decision methods (MCDM) (Lorenz et al., 2022; Udokporo et al., 2021), which can also be integrated with fuzzy theory (Jain et al., 2020). Some MCDMs may be time-consuming (e.g. DELPHI) or inconsistent when several comparisons are

CONTACT Guoqing Zhao guoqing.zhao@swansea.ac.uk School of Management, Swansea University, Swansea SA1 8EN, UK

considered (Gohari et al., 2022). Furthermore, when multiple experts are consulted, it is required to deal with conflicts among experts' opinions (Z. Zhang & Li, 2022), which is not a trivial task. Finally, expert-based DSSs usually require experts to express opinions, and the information they provide is generally limited to the field of application. Indeed, MCDMs such as the Analytical Hierarchy Process (AHP) can be influenced by subjectivity (Yadav et al., 2022). Thus, the results cannot always be generalised and are often limited to the needs and priorities of the firm adopting the tool. Based on the these considerations, it is important to develop DSSs that are easy to interpret, quick to use, and able to provide general results.

DSSs are adopted for different purposes and in various fields, including maritime operations. In this context, DSSs are used to evaluate port performance or identify the critical factors influencing performance. Indeed, maritime shipping is considered the principal transport mode for goods, accounting for about 90% of worldwide trade volume (Wang et al., 2023). This has resulted in the continuous development of DSSs to assess port and maritime operational performance (Gao et al., 2019; Wong et al., 2021). Furthermore, due to their economic importance and widespread use, ports should maintain an appropriate level of sustainability by considering all the factors that characterise the Triple Bottom Line (TBL): economic, environmental, and social. To this end, Port Sustainable Performance (PSP) has attracted significant attention (Garg et al., 2022; Kong et al., 2022; Tseng & Pilcher, 2019; Tseng & Yip, 2020; Tsolakis et al., 2022), leading to the proposal of different DSSs. Meanwhile, the advent of COVID-19 has led ports to invest in digital technologies (Alamoush et al., 2022), including big data platforms and cloud technologies. Information technologies and Big Data Analytics Capability (BDAC) can help organisations better manage massive datasets, achieve effective and efficient decision-making, and improve organisational performance (Chen et al., 2022; Di Vaio & Varriale, 2020; Simoni et al., 2022). In recent years, the ability to handle big data has improved drastically (Bhattacharya et al., 2024), and BDAC is usually regarded as a game-changer (Raut et al., 2021). In this regard, some studies highlight the benefits of BDAC on organisational and operative performance (Mikalef et al., 2020; Wamba & Akter, 2019). BDAC is also seen as a smart supportive layer for the development of smart ports (H. Li et al., 2023). However, insights into how BDAC influences organisational performance are still limited (Orero-Blat et al., 2025). Moreover, little attention has been devoted either to evaluating ports' BDAC or to accessing its impact on both port operative (PP) and PSP. In fact, big data are mainly used to predict parameters, detect anomalies, and automate operations, while the literature linking BDAC and environmental considerations remains scarce for container terminals (Raeesi et al., 2022). Finally, although many DSSs have been developed for environmental and operational performance, a DSS for evaluating port BDAC is still lacking. Thus, this paper first aims to propose a method for developing a DSS based on the integration of experts' judgements and ML. Specifically, experts from different organisations are asked to evaluate the performance of their organisation. The collected opinions are used to train an unsupervised ML model. Accordingly, the resulting tool can be considered general because it is created from information collected from different firms. This aspect helps overcome the specificity of results typical of multi-criteria methods based on expert judgements. Furthermore, the DSS can be used without collecting extensive data or implementing algorithms, thereby overcoming the limitations of ML applications. Second, this paper aims to evaluate ports' BDAC and its relationships with PP and PSP. Developing a tool for evaluating port performance could help highlight the strengths and weaknesses that characterise ports.

The remainder of this paper is structured as follows. Section 2 presents the theoretical background regarding DSSs, unsupervised machine learning and Fuzzy Set Theory (FST) in the context of port operations. Section 3 illustrates the tools and steps of the framework to develop the DSS, while Section 4 describes the main results. Section 5 discusses the findings. Finally, Section 6 presents the conclusions, limitations, and avenues for future research

2. Theoretical background

In this section, first, the background related to DSSs developed for port operations is presented. Subsequently, the section is expanded by focusing on the adoption of unsupervised learning and Fuzzy Set Theory (FST) for maritime performance evaluation. Finally, the conceptual framework of the present study is described in the last subsection.

2.1. Decision support system in port operations

In the context of port operations, different approaches are used to develop tools capable of guiding decisions, which, in turn, may be related to different factors. Examples of DSSs for port operations and management include risk-based models for natural hazards (Banan-Dallalian et al., 2023; Cremen et al., 2022), systems that support logistics (Rotunno et al., 2023; Zhen et al., 2022), and models that support maintenance activities (Ebrahim et al., 2022).

In addition to these fields, DSSs are also used to evaluate port performance, such as port resilience (Zhou et al., 2021). In this context, researchers employ different tools, among which ML and multi-criteria models are very popular. For instance, Wong et al. (2021) proposed a DSS based on a mathematical model combined with optimisation algorithms (such as deep neural network) to predict shipping network yields. In another recent study, Moscoso-López et al. (2021) employed different ML algorithms to predict perishable cargo flows, while H. Li et al. (2023) compared different ML algorithms for trajectory prediction. Another work by Gao et al. (2019) used a recurrent neural network to predict daily container flows. Cankaya et al. (2021) adopted distinct ML techniques to predict and classify vessel activities. According to the previous paragraphs, ML is generally employed to predict future conditions, which, in turn, could support decision-making process. However, ML may not always be a viable option for firms and organisations because of several challenges, such as a lack of data or know-how barriers (Pan et al., 2022). Regarding multi-criteria model-based DSSs used to evaluate maritime performance and related activities, several techniques are adopted, such as the Analytic Hierarchy Process (Attardi et al., 2015; Lirn et al., 2013) and the Technique for Order of Preference by Similarity to Ideal Solution (X. Zhang et al., 2022). These techniques usually employ expert judgements, which are associated with uncertainty. To address this uncertainty, FST can be exploited (Celik & Akyuz, 2018; Salleh et al., 2014). In other words, these DSSs could be used to evaluate unknown or latent variables; however, even though the tools are general (i.e. they can be employed by any user), the results they provide are usually case-specific (i.e. they are valid only for the port that uses them).

Based on the above considerations, it could be beneficial to merge the advantages of ML-based and expert-based DSSs to obtain a DSS capable of providing general results and being used quickly without the requiring specific algorithmic or mathematical knowledge.

2.2. Unsupervised learning and fuzzy set theory to evaluate port performance

Among unsupervised learning approaches, clustering observations into groups is very popular in supply chain and operations management. Clustering algorithms are used for several purposes, such as grouping suppliers into different segments to define a strategy for each supplier (Coşkun et al., 2022), evaluating the performance of similar supply chains in different regions (Sabouhi et al., 2021), and dividing enterprises into different categories of environmental practices (Pahuja et al., 2024).

Unsupervised learning is also employed in the context of port operations; however, it has seen far fewer applications than supervised learning (Filom et al., 2022). One possible reason for this trend is that unsupervised learning has an unknown output for each observation, unlike supervised learning. Thus, even though unsupervised learning does not require experts to label the available data, it requires a deeper understanding of the results. Typical unsupervised learning algorithms used for port operations include K-Means and Density-Based Spatial Clustering (DBSCAN). The first technique is used by Liu et al. (2020) and AbuAlhaol et al. (2018) to study port congestion. Port congestion and traffic have also been studied by Xin et al. (2022), Peng et al. (2022) and Wang et al. (2022) through DBSCAN and its practical evolutions. It is worth mentioning that unsupervised machine learning is also used for other purposes in the context of port operations. For instance, Fuentes (2021) employed DBSCAN to rank ports based on their relative importance, while Cabral and de Sousa Ramos (2014) adopted K-Means, hierarchical clustering, and partitioning around medoids to divide 17 ports into different groups based on a set of competitiveness criteria. Azzam et al. (2021) also divided ports into different clusters through K-Means, considering twenty-five operational dimensions such as port size. In a more recent work by Mansouri et al. (2021), K-Means was used to classify a group of eight ports based on their ability to accommodate autonomous ships. K-Means and, more generally, unsupervised learning are used in the context of port operations to evaluate port congestion or classify ports based on their characteristics, mainly operational characteristics. However, little attention has been devoted to classifying ports based on their ability to analyse Big Data.

Another popular solution for analysing port performance, characteristics, or features is to exploit expert judgements. Indeed, experts can provide information in cases of data scarcity. In this context, it is pivotal to aggregate judgements on the same subject from different experts, who may hold different opinions. Moreover, it is fundamental to address the subjectivity of the respondents. A widespread approach for performing these tasks is the adoption of fuzzy set theory, especially for risk analysis and risk assessment related to port operations (Alyami et al., 2019; Sarkar et al., 2022; Wen-Xin & Guo-Ping, 2021). Fuzzy set theory and related techniques are also popular for evaluating port performance (Bray et al., 2015; Castellano et al., 2019) and, subsequently, comparing different ports or potential port locations (Pamucar & Görçün, 2022; Raad et al., 2022). Although most research has traditionally focused on operational port performance, green and sustainability aspects have progressively become more important; therefore, recent works have incorporated them into their studies (Zhao et al., 2021). Green aspects are usually considered either as a standalone class (Garg et al., 2022; Jaidee et al., 2022; Tseng & Pilcher, 2019) or as part of broader port performance evaluation (Tseng & Yip, 2020; Zhao et al., 2021). In the present study, green performance is regarded as a standalone category that includes environmental, economic, and social factors (i.e. the pillars of the Triple Bottom Line concept). Furthermore, most current research has concentrated on identifying the factors that influence the port choice and evaluating investment conditions in ports (Nguyen et al., 2022). In contrast, there is a shortage of studies related to port performance and sustainability and how to improve them.

2.3. Research gaps and theoretical framework

While previous studies have shown a potential direct correlation between BDAC and organisational performance, the mechanisms and conditions under which this correlation occurs have not been fully understood (Fosso Wamba et al., 2024; Orero-Blat et al., 2025). As seaports are vital links in the global supply chain, it is critical to understand whether big data analytics capabilities lead to competitive performance advantages and how these effects are generated (Raeesi et al., 2023). Meanwhile, although emerging digital technologies have been widely applied in sustainable shipping practices, academic research in operations management still lags significantly behind practical developments in exploring the relationship between big data analytics and maritime sustainability (Pang et al., 2026). Moreover, prior research on the relationship between BDAC and organisational performance has yielded conflicting conclusions because of a lack of empirical testing (Su et al., 2022). An organisation's BDAC includes a range of competencies. These factors may include management capabilities, organisational culture and technology capabilities. Although some companies recognise the potential of big data analytics in both tangible and intangible data assets, they still face challenges in integrating big data analytics into their strategic frameworks to improve performance (Raissi et al., 2025). Bejjani and Vanhaverbeke (2025) also call for further research on how organisational resources and capabilities can contribute to building big data-related capabilities and competitive advantages. As more organisations begin to use big data analytics, proposing a DSS can help organisations understand their BDAC and create implementation strategies. To address the above challenges and limitations, this paper collects data from experts at different ports through a questionnaire that includes three major dimensions: BDAC, PP, and PSP. The questionnaire enables the acquisition of data that are not explicitly available in firms' databases. The data obtained from different experts are then aggregated for each port through FST, as shown in Guo et al. (2021), and subsequently unsupervised machine learning is used to cluster data and classify ports into different categories for BDAC, PP, and PSP. Specifically, K-Means is used since it is a very popular tool for clustering and classification problems (Mengüç et al., 2023; Su & Wu, 2023). The clustering provides general results since many ports are considered in the analysis. Finally, the clustered and labelled data are processed through a Decision Tree (DT) and Bayesian Network (BN). The DT provides a visual and user-friendly DSS because of its interpretability (Jun & Lee, 2021). On the other hand, the BN is used to conduct an inference analysis of the relationships between BDAC, PSP, and PP. Bayesian analysis allows statistical inference based on real-world observations (Leoni & De Carlo, 2023). The preceding discussion is summarised in Figure 1, which represents the conceptual framework, while the subsequent section provides an in-depth description of each step.

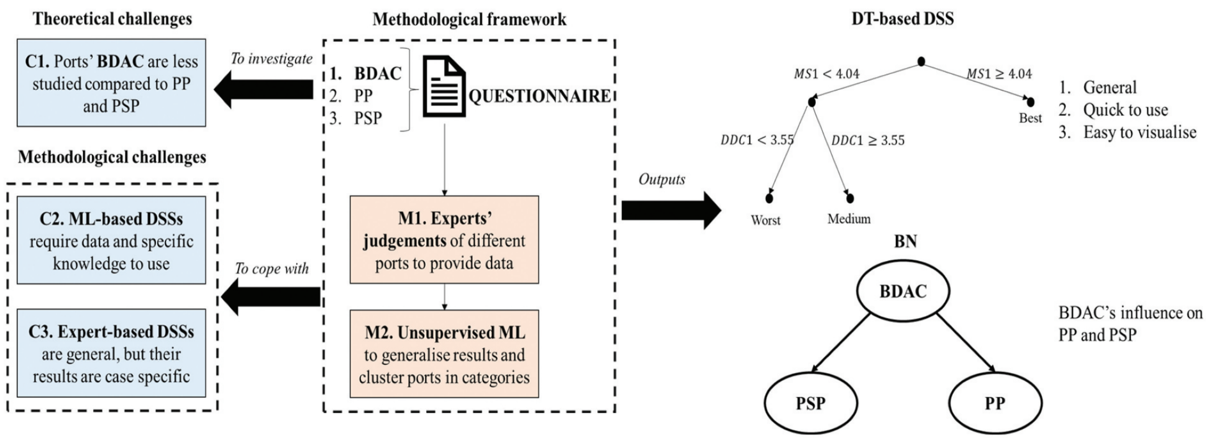


Figure 1. Conceptual framework reporting challenges, main methodological idea, and desired outputs.

3. Developed methodology

The flowchart of the proposed methodology is reported by Figure 2.

3.1. Stage 1: data collection

First, a questionnaire related to BDAC, PSP and PP is developed (step 1). The questionnaire is based on a fuzzy Likert scale, and answers are collected from different experts and different ports (step 2). The Likert scale is an approach that converts qualitative values, usually expressed through a linguistic scale (e.g. level of agreement), into quantitative ones. This allows more appropriate statistical analysis to be conducted (Santino et al., 2022). Human judgements involve uncertainty, which increases in cases of different or even conflicting answers (Pasman & Rogers, 2020). Within this context, fuzzy set theory can assist practitioners in dealing with the subjectivity of the answers (Lazim & Wahab, 2010). Accordingly, a fuzzy Likert scale can be employed to deal with the uncertainty arising from different judgements on the same matter.

3.2. Stage 2: clustering

The judgements from different experts at the same port are aggregated through fuzzy set theory (step 3.1), using the improved Similarity Aggregation Method (SAM) recently developed by Guo et al. (2021). The improved SAM can consider both respondents' credibility and the consensus among different respondents. A brief description of the improved SAM is reported in Guo et al. (2021) and Jianxing et al. (2021). Next, one sub-dimension is considered (step 4), and the observations are divided into three clusters identifying the worst, medium and best clusters. This choice was made to maintain uniformity across dimensions and sub-dimensions. Furthermore, the adoption of three clusters allows for ease of interpretation while avoiding overly small or large clusters. The process is performed with K-Means (step 5). K-Means has become one of the most popular clustering techniques in several fields (Huang et al., 2021), and it was chosen because of its efficiency and ease of implementation. K-Means partitions different observations into different clusters by minimising the squared error distance of each observation to the centre of the nearest cluster (Bai et al., 2023). The process is iterated for each sub-dimension. Then, the first macro-dimension is selected (step 6), and the observations are clustered into three clusters through K-Means (step 7). The process is iterated for each macro-dimension.

The clusters are analysed by considering the mean values and standard deviations to determine the best, medium, and worst clusters. Moreover, the differences are assessed through MANOVA using SPSS Statistics® and through Permutational MANOVA (PERMANOVA) using the vegan library and the in-built adonis function in R. Finally, post-hoc tests are conducted.

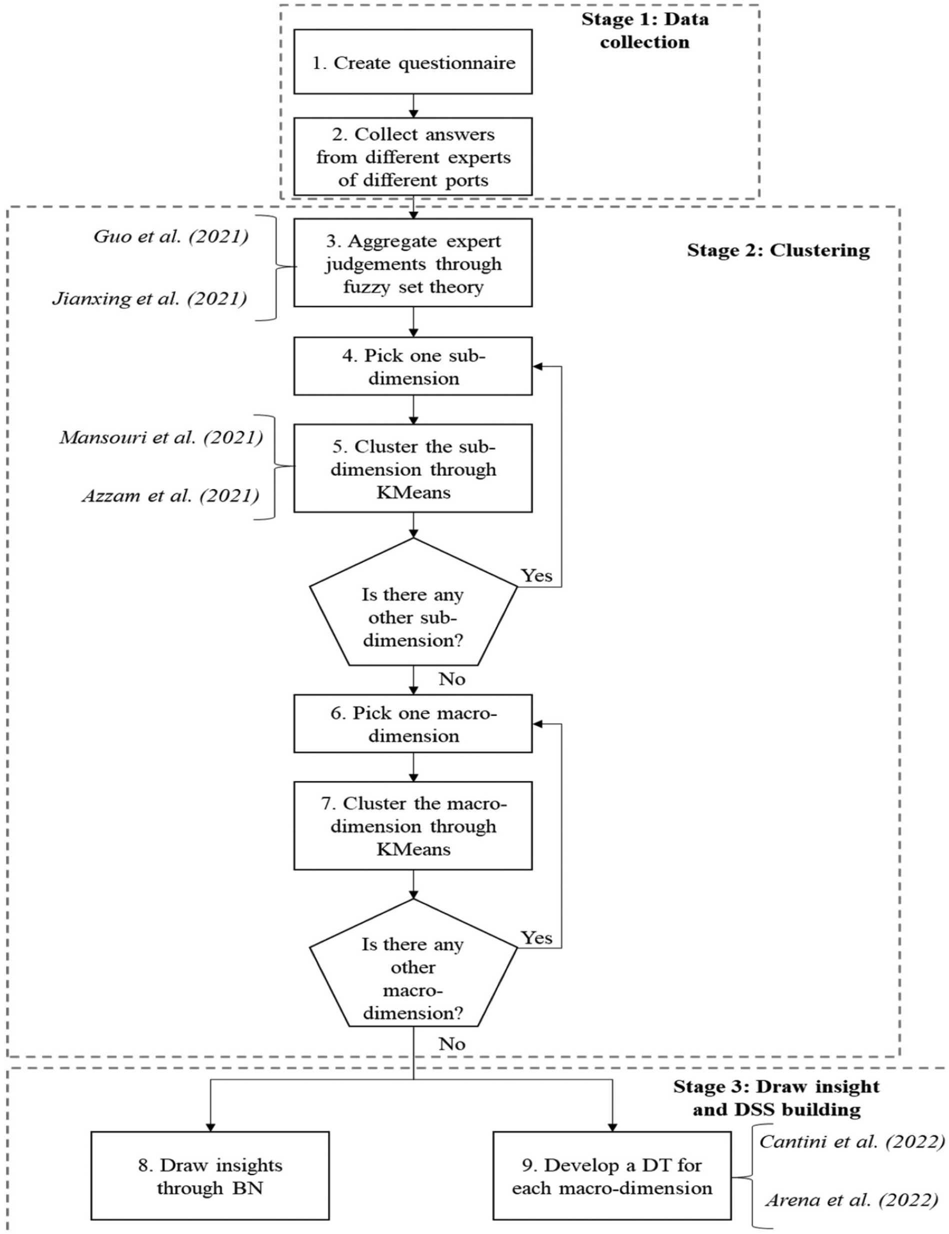


Figure 2. Flowchart and steps of the proposed approach with relevant references.

3.3. Stage 3: draw insights and decision support system building

After the worst, medium and best clusters are identified for each macro-category (e.g. BDAC, PSP and PP), the relationships among them are studied through a BN (step 8). The BN is a Directed Acyclic Graph (DAG) that enables statistical inference and supports the decision-making process (Matthias et al., 2017). For instance, given the best cluster for BDAC, the BN determines the conditional probability of observing the worst, medium and best clusters for both PSP and PP.

Considering the identified clusters and root variables, a DT is built for BDAC, PSP and PP (step 9). The DT is chosen due to its simplicity and interpretability (Zangaro et al., 2021). Indeed, thanks to its tree structure, it is possible to visualise DTs through different graphical representations (Mrva et al., 2019), allowing for easy practical use. This feature has attracted many researchers, who have used DTs to create user-friendly and easy-to-visualise DSSs. For instance, Cantini et al. (2022) developed a DT-based DSS for supply chain design, while a similar study was proposed by Arena et al. (2022) to guide predictive maintenance implementation. Furthermore, the DT can be used for descriptive purposes to highlight the differences among groups as shown in Sgarbossa et al. (2021). Based on the previous statements, the developed DTs will allow each port to assess its cluster for the three macro-dimensions, underlining its weaknesses and strengths. Thus, the DTs can be considered a DSS.

4. Results: application of the methodology

4.1. Questionnaire development and data collection

This study uses a questionnaire-based survey as the primary data collection method. It develops item scales to assess causal conditions based on previous studies. Table A1 describes all the scales and items, as well as their respective sources. Respondents were asked to evaluate the questions based on a five-level fuzzy Likert scale, shown in Table 1 (Martínez-Mendoza et al., 2020). Before large-scale data collection, a pilot study was conducted to enhance the reliability and validity of the measurement scales (Kumar Sharma et al., 2024). Initially, a group of five port experts and five scholars was recruited to assess the accuracy and consistency of the survey. The group of port experts consisted of two high-level port executives and three information technology managers who work for the ports of Shanghai, Rotterdam, Shenzhen, and Piraeus. The five scholars specialise in the fields of Big Data and port management. They examined the definition of the constructs and assessed the significance of each item in relation to its theoretical construct. Based on the recommendations of port managers and scholars, we revised several questions and restructured the questionnaire to enhance clarity and facilitate respondents' ability to provide answers. After the pilot study, an email containing a link to the online survey was sent to port managers and IT department staff in the world's top 50 ports. The initially contacted port managers were encouraged to refer other port managers in the field in order to increase the sample size and obtain further insights (Emerson, 2015). A total of 207 questionnaires were collected. Of the 207 participants, 49 were excluded because they had insufficient experience with big data technologies or did not meet the criteria for the target ports. The final sample consisted of 158 respondents from 40 different ports, with each port having between 1 and 20 respondents. The variation in respondents per port was addressed during the data aggregation process through the SAM. The SAM accounts for both the level of agreement among respondents and their expertise. In other words, both the experience of the experts and the consistency across their judgements are considered. In this way, ports with a larger number of respondents did not disproportionately influence the results. A higher number of respondents reduces the influence of individual expert misjudgements. Nevertheless, since the experts were required to have sufficient experience with big data technologies, even ports with few respondents can be considered properly represented.

4.2. Cluster definition and development

First, a weight is assigned to each expert according to Table 2, based on age, service time, BDA working experience, and education level. Subsequently, the judgements from different experts are aggregated through the improved SAM described in Section 3.3 (step 3), adopting a relaxation factor equal to 0.5, which is the most common value (Jianxing et al., 2021).

After the data are aggregated, the first sub-dimension (i.e. Organisational Learning) is considered (step 4). Organisational Learning (OL) is composed of three items, which measure different aspects. Next, K-Means is conducted through Matlab® (step 5), considering 1,000 iterations, the in-built K-Means++ method to define the initial centroids, and the Euclidean distance, which is the most popular distance measure (Ghazal et al., 2021). As previously mentioned, three clusters were considered for each sub-dimension. Indeed, since only

Table 1. Linguistic scale, Likert scale and fuzzy scale adopted for the study.

Linguistic scale	Likert Scale	Fuzzy scale
Strongly disagree	1	(0,1,2)
Disagree	2	(1,2,3)
Neutral	3	(2,3,4)
Agree	4	(3,4,5)
Strongly Agree	5	(4,5,5)

Table 2. Considered categories and classifications for defining experts' weights.

Category	Classification	Score
Age	≥50	8
	40–49	6
	30–39	4
	<30	2
Service Time	≥30	10
	20–29	8
	10 19	6
	6 9	4
	≤5	2
Big Data Working Experience	≥5	8
	3–4	6
	2–3	4
	<2	2
Educational Level	Postgraduate	8
	Graduate	6
	High School Diploma	4
	Vocational Education	2

40 ports responded to the questionnaire, it would have been problematic for some dimensions to cluster the data using a higher number of clusters. Specifically, employing a high number of clusters could have resulted in very small clusters (e.g. one or two observations) for many sub-dimensions, possibly leading to overly similar clusters (Salvador & Chan, 2004). On the other hand, adopting too few clusters could lead to difficult-to-interpret results (Binkin et al., 1992). Furthermore, even though clustering the data into four or more clusters could have been possible for some sub-dimensions, using three clusters for all sub-dimensions allow to have uniformity. The results of the K-Means clustering associated with OL are shown in Figure 3. Moreover, the calculation produced a mean silhouette value of 0.399. In Figure 3, the blue dots represent the best ports for OL (i.e. those associated with the highest means for OL1, OL2, and OL3). The orange dots represent ports with medium performance, while the grey dots denote ports with poor performance (i.e. the worst for OL).

To assess the diversity of the identified clusters, both MANOVA and PERMANOVA are conducted. PERMANOVA was required since the assumptions of MANOVA (e.g. multivariate normality) were not always

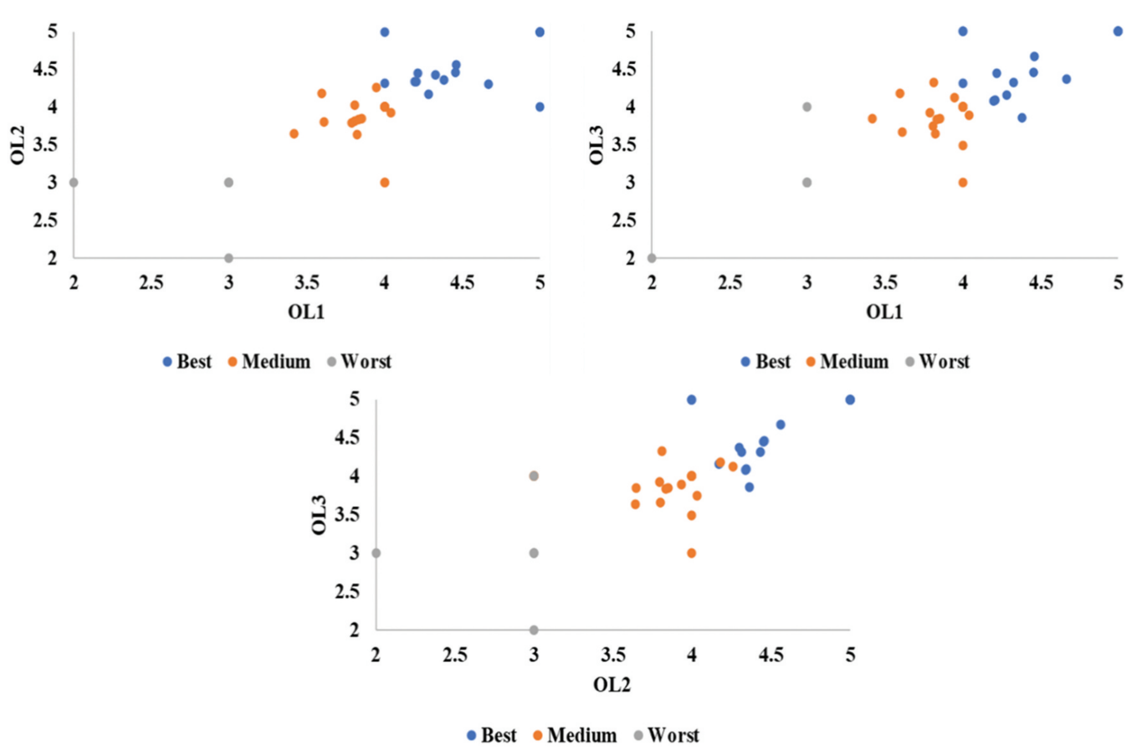


Figure 3. Graphical representation of the clusters identified for organisational learning.

met. PERMANOVA is a non-parametric multivariate test that is more robust than MANOVA because it is distribution-free and accounts for unbalanced designs and heterogeneous dispersion (Anderson, 2014). However, since differences in within-group dispersion could lead to misleading results (Anderson, 2001), the differences among within-group dispersions are tested using the betadisper function.

Considering the OL's clusters, the multivariate normality required by MANOVA was not met; however, the Pillai's trace is considered robust to the absence of this assumption. Moreover, the Box's test of equality of covariance matrices was significant with a *p*-value equal to 0.119. Finally, the Levene's test based on median for the homogeneity of variances had a *p*-value greater than 0.1, proving the equality of variances. Accordingly, MANOVA could be applied, and the calculation showed differences among the three clusters, with a Pillai's trace equal to 0.870 ($p - value = 0.000$), a Wilk's Lambda value of 0.157 ($p - value = 0.000$), and a Roy's largest root of 5.170 ($p - value = 0.000$). MANOVA assesses whether there is a difference between two of the identified clusters; thus, it does not provide any information on which clusters are and are not statistically different. To obtain this information, one-way ANOVAs with Scheffe's post-hoc test were performed, and the results were significant for each cluster and measurable item ($p - value \leq 0.001$).

Although the application of MANOVA could be considered valid for OL, this was not always verified for all sub-dimensions. As an example, the results related to the PERMANOVA application for OL are briefly illustrated. Specifically, F-ratio value of 42.876 ($p - value \leq 0.001$) emerged, highlighting the differences among the clusters. Furthermore, the test on the homogeneity of within-group dispersion gave an appropriate result with a $p - value = 0.165$. Finally, Bonferroni's post-hoc tests were conducted using the function pairwiseAdonis (Martinez Arbizu, 2020). The calculation showed significant differences among all groups with an $adjustedp - value \leq 0.003$.

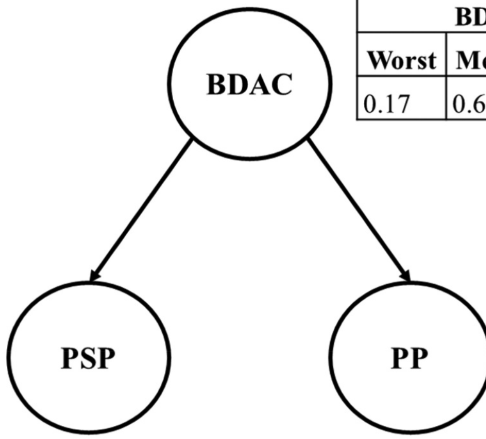
Following the flowchart presented in Figure 1, the process is iterated for all sub-dimensions of BDAC, PSP and PP. The results are shown in Table A2. The application of PERMANOVA showed that all the identified clusters were different with a $p - value \leq 0.001$. Furthermore, the analysis of the homogeneity of within-group dispersion provided significant results with $p - value \geq 0.05$. Considering the test, Customer Orientation (CO) is the only sub-dimension characterised by a *p*-value lower than 0.05 ($p - value = 0.029$). Finally, the Bonferroni's post-hoc test resulted in significant differences with $adjustedp - value \leq 0.003$ or $adjustedp - value \leq 0.006$ for all the sub-dimensions, except for Value-added Service (VAS) ($adjusted\ p - value < 0.05$).

After clustering all sub-dimensions, the first macro-dimension, i.e. BDAC, is considered (step 6). Next, the macro-dimension is clustered through K-Means with the same parameters presented in the previous paragraphs (step 7). Then, the process is iterated for the remaining macro-dimension (i.e. PSP and PP). To perform this phase, the clusters identified for the sub-dimensions are associated with a value from 1 to 3, with 1 and 3 denoting worst and best class respectively. The results are shown in Table A6. The results show that few ports have poor performance for the three macro-dimensions. Indeed, the numbers of ports in the worst clusters are 7, 9, and 6 for BDAC, PSP, and PP respectively. Furthermore, more ports belong to the best clusters of PSP and PP than to the BDAC best cluster. PERMANOVA was performed to evaluate the differences among the detected clusters of each macro-dimension. The results show that the F-Ratio of BDAC, PSP and PP are 21.67, 42.92 and 26.40 respectively, all characterised by $p - value \leq 0.001$. Furthermore, the homogeneity of within-group dispersion analysis gave positive results for both BDAC and PP, with $p - value \geq 0.1$. However, the homogeneity of within-group dispersion was not met for PSP ($p - value = 0.002$). Finally, the Bonferroni's post-hoc test resulted in $adjustedp - value \leq 0.006$, denoting the differences among the identified clusters.

4.3. Decision support system development

A BN is built considering the clusters obtained for the macro-dimensions (step 8). Indeed, the BN, given a certain an observation belonging to a certain cluster for the BDAC (best, medium, or worst), points out the conditional probabilities for the other two macro-dimensions of being in one of the three identified clusters. For instance, given a port that falls in the best performance cluster for BDAC, the BN determines the probabilities of that port being in the worst, medium, and best clusters for PSP (or PP). The developed BN and the related CPTs are shown in Figure 4. The results show that if a port has poor performance for BDAC, it could still perform very well for PSP and PP. On the other hand, in the considered sample, no port

Probability Cluster BDAC		
Worst	Medium	Best
0.17	0.6	0.23



CPT ClusterPSP			
Parents	Probability		
ClusterBDAC	Worst	Medium	Best
Worst	0.71	0.14	0.14
Medium	0.17	0.58	0.25
Best	0	0	1

CPT ClusterPP			
Parents	Probability		
ClusterBDAC	Worst	Medium	Best
Worst	0.71	0.14	0.14
Medium	0.04	0.67	0.29
Best	0	0.11	0.89

Figure 4. Developed BN and CPTs displaying the relationships between BDAC, PSP, and PP.

has excellent performance for BDAC and poor performance for PSP and PP. Indeed, all ports in the excellent BDAC cluster are also in the excellent PSP and PP cluster. Only one port ends up in the medium cluster for PP. The aforementioned considerations are identified through the following conditional probabilities of the BN: P(ClusterPSP = Best|ClusterBDAC = Best) = 1, P(ClusterPP = Best|Cluster BDAC = Best) = 0.89, P(ClusterPP = Medium|ClusterBDAC = Best) = 0.11, P(ClusterPSP = Bad|Cluster BDAC = Best) = 0, P(ClusterPSP = Medium|ClusterBDAC = Best) = 0, P(ClusterPP = Bad|Cluster BDAC = Best) = 0.

A DT is built for each macro-dimension, considering the measurable observations as inputs and the cluster of the macro-dimension as the output (step 9). The training was conducted with 5-fold cross-validation, considering a maximum split equal to 2 for readability purposes and to make the DTs easier and faster to use. The calculation produced validation accuracies of 77.5%, 80%, and 70% for BDAC, PSP, and PP, respectively, showing decent generalisability performance. The DTs are illustrated in Figure 5, which also shows the accuracy of each leaf and the DT's overall accuracy estimated when the entire dataset is used as training data. The DTs can be used as a DSS, since any port can determine its cluster for BDAC, PSP, and PP and accordingly define the actions that should be undertaken to improve the weakest macro-dimension. It is worth mentioning that the DTs could change slightly based on the extracted training set and the initial random seed used for DT generation.

Considering the BDAC' DT, the overall accuracy was 92.5%, indeed only three ports were incorrectly classified. Specifically, one port belonging to the worst cluster was classified as best because of its high value of MS1. Moreover, two ports in the excellent cluster and one port in the worst category were classified as medium by the DT. Considering the DT related to PSP, the overall accuracy was estimated at 90%, since four ports were misclassified. To be more precise, two ports were assigned to the worst cluster by the DT, while their original class was medium. Furthermore, one medium port and one worst port were incorrectly classified as excellent. Finally, the overall accuracy of the PP's DT was 90% because four ports were assigned to the wrong cluster. Specifically, two distinct ports originally identified as excellent ended up in the worst and medium clusters respectively. Moreover, two different medium ports were incorrectly identified as worst and excellent respectively.

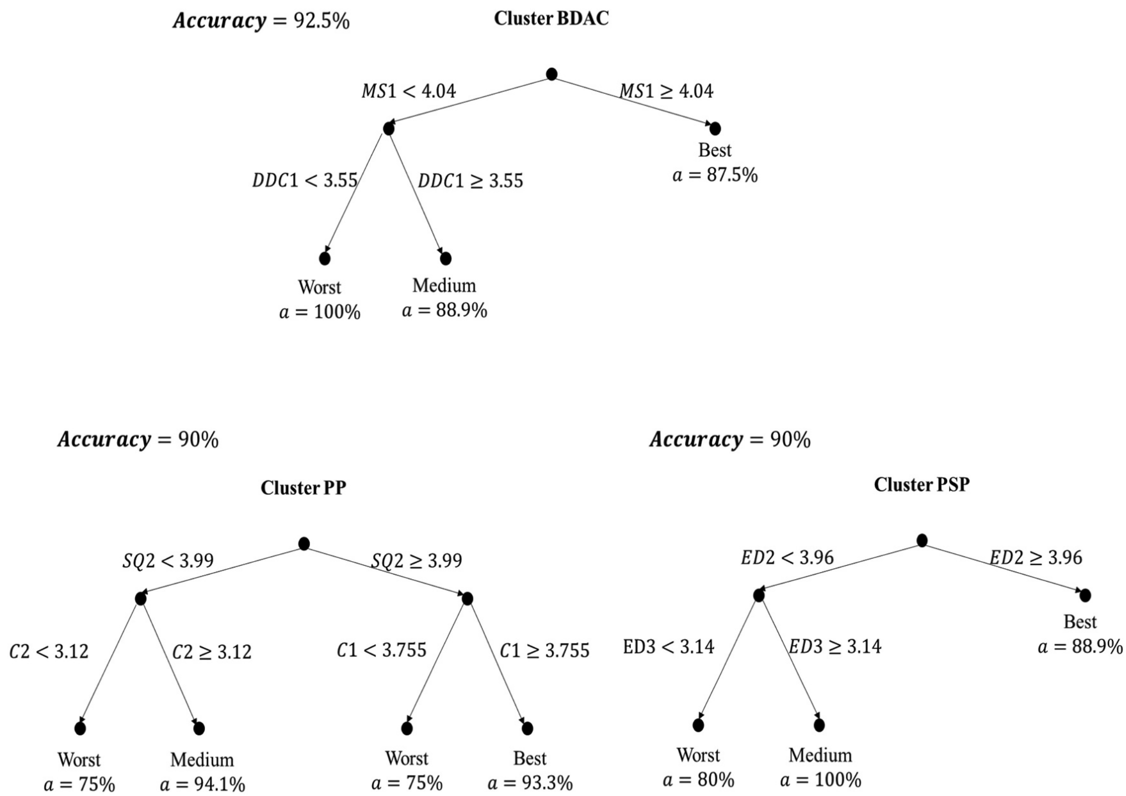


Figure 5. DT-based DSS to evaluate port performance and accordingly making decisions.

5. Discussion

Considering the Bayesian analysis, the results suggest a positive association between BDAC cluster membership and PSP/PP cluster membership in the sampled ports. Indeed, all the ports identified as the best ports for BDAC were also identified as the best ports for PSP, and for PP, their performance was either medium or best. This result is similar to the result of previous research (Akter et al., 2016; Dubey et al., 2019; Song et al., 2017; Wamba & Akter, 2019), which report that strong BDAC can assist ports in improving both PSP and PP. Being able to properly analyse available data is fundamental in many fields for pursuing greater performance. Furthermore, ports with lower BDAC are more likely to have low performance in the other two macro-dimensions. For instance, ports in the worst BDAC cluster have a 71% probability of also being classified as worst for PSP and PP. This study provides evidence consistent with the importance of BDAC for PSP and PP (Del Giudice et al., 2022; Jović et al., 2019), thereby extending the empirical work of Mikalef et al. (2020) and Jeble et al. (2018) to the port context. Nevertheless, some ports with poor BDAC still have decent or high PP and PSP. It is reasonable to assume that aspects other than BDAC influence PP and PSP. For instance, port size, national regulation, infrastructure maturity, digital capabilities, environmental compliance pressure, and operational excellence practices may compensate for lower BDAC. These aspects are still worthy of investigation.

Despite the previous considerations, a port with low BDAC performance could still perform well in terms of PSP or PP. Indeed, as depicted in Figure 5, a port in the worst BDAC cluster has a probability of 14% of being identified as best for PSP or PP. This finding agrees with Saunila et al. (2019), who showed that no significant correlations were obtained between smart technology and social sustainability or between smart technology and environmental sustainability. This means that BDAC alone is not sufficient to improve the PSP and PP of ports. Another possible reason for this result might be related to the method. Clustering is not a perfect or deterministic tool, and it strongly depends on the number of clusters selected for the analysis. As a matter of fact, one port could be far from every cluster (e.g. it could be an outlier), but it is still assigned to one of the clusters. Consequently, the port will share few similarities with the other ports in that cluster.

5.1. Decision support system and description of clusters

To fulfil the second objective of this work, a DT-based DSS is built by considering the clusters associated with each port for each macro-category. The DTs shown in Figure 5 can be used by any port to evaluate its performance compared with other ports and accordingly define which macro-category should be prioritised. For instance, if a port has medium performance for BDAC and PP but is classified as poor for PSP, the port should first work on PSP. To apply the decision trees (DTs), port managers must first evaluate their port's performance using the variables listed in Appendix. Based on the values assigned to each variable, they can then determine the class to which their port belongs for each macro-category defined by the DTs. Specifically, each variable value determines the path followed through the tree, with different branches taken until a final classification (leaf) is reached. For example, consider a port with the following values: MS1 = 5, SQ1 = 4, C1 = 4, ED2 = 3, and ED3 = 4. According to the DTs shown in Figure 5, this port is classified as 'best' for BDAC. This is because the right branch is selected at the first node, as the threshold condition MS1 ≥ 4.04 is satisfied. The port is also classified among the best for PP. At the first level of the DT, the right branch is taken because SQ2 ≥ 3.99. At the second level, the right branch is again selected because C1 ≥ 3.755. Following the same procedure for the DT for PSP, the port is classified as 'medium' in this category. Overall, this means that the port performs strongly in BDAC and PP, but there is room for improvement in PSP.

The DTs also provide an overview of the differences among the distinct clusters, highlighting the variables most strongly associated with the cluster classification of the macro-dimensions. Accordingly, the variables reported on the DT branches are those that should be considered when improving the performance of a given macro-category. Considering BDAC, the variables that ports should work on are MS1 and DDC1. MS1 refers to managers' ability to exploit Big Data-related activities to support other partners. Based on the results arising from the DT, managers working in the best ports are usually more skilled in supporting other partners through Big Data than managers in ports associated with medium BDAC. Strong management and data analytics skills can assist managers in better analysing big data and forecasting future demands of other supply chain participants, thereby improving service quality and organisational performance (Su et al., 2022). Strong management skills can also ensure a higher level of coordination among the internal divisions of the port, resulting in a better perception and reconfiguration of business processes (Li et al., 2022). Research obtained by Mikalef et al. (2020) and Majhi et al. (2026) are consistent with this finding. Hence, as firms use sophisticated data analysis methodologies and embark on big data initiatives, it is important for managers to have strong management skills and rich experience to understand their operations and potential. DDC1 expresses the extent to which the port considers its data a tangible asset. Accordingly, DDC1 can be regarded as the starting point of any data-driven organisation, since the more data are considered as an asset, the greater their importance and the greater the effort directed towards data analysis. In other words, it is a key aspect of organisational culture, which may positively or negatively influence other dimensions. Indeed, as depicted by the DT, a DDC1 value lower than 3.55 identifies only the worst ports. This finding also shows the importance of a data-driven culture in creating BDAC. The findings are in line with earlier studies (Ciampi et al., 2021), which show that fostering a data-driven culture can dismantle organisational barriers, facilitate the integration of knowledge across departments, and enhance the use of data within organisations Port managers are advised to make decisions based on data rather than relying on instinct when operating within a data-driven culture.

Regarding PSP, the most relevant variable was ED2, which is related to the adoption of data analytics for wastewater reduction. ED2 values higher than 3.96 are mainly associated with the best ports, which show better performance than the worst and medium ports with regard to wastewater treatment. The finding of Mahmud et al. (2024) is consistent with this result, pointing out that pollution control measures have the greatest impact on green port management practices. One possible interpretation of this finding is that seaports are often situated near residential communities or environmentally sensitive locations, and pollution from wastewater affects both the environment and the public health (Molavi et al., 2020). The other relevant variable is ED3, which represents the adoption of data analytics technology to reduce noise in port areas. This factor is considered more in ports belonging to the medium cluster than in those associated with the worst cluster (see the corresponding DT in Figure 5). This finding agrees with Puig and Darbra (2024), who indicated that utilising innovative methods such as big data analytics to monitor noise pollution is crucial for mitigating its impacts on surrounding ecosystems and communities. Noise produced by ports and their surrounding activities affects the environment, the health of port workers and communities, and the image of the port.

Monitoring and reducing noise levels could contribute to the growth and development of ports together with cities and regions (Zain et al., 2022). Considering PP, the first and most relevant variable is SQ2, which refers to the port's ability to exploit data analytics to handle shipments in a timely manner. As shown by the DT in Figure 5, the best ports generally have a greater ability to exploit Big Data to avoid shipment delays than medium-performing ports. Attanasio et al. (2023) argued that an important reason for the digitisation of ports is to reduce resource consumption, non-value-added operations, and idle time in order to increase transhipment productivity and competitiveness. Ports need to respond quickly to customer requirements and avoid delays. By leveraging Big Data technology, ports can capture and integrate a diverse range of customer data. By analysing collected customers data, ports can gain a deeper understanding of customers' expectations and complaints, enabling them to make informed decisions and fulfil customer requirements promptly. Thus, SQ2 is important and should be among the variables that a port works on to improve PP. The findings are in line with earlier studies (Hsu et al., 2023; Melnyk et al., 2024), which show that adopting smart port technologies such as big data analytics can improve port operations and customer satisfaction. The other relevant variable is C2, which evaluates how much the port charges for intermodal transport compared with its competitors. The results show that the worst ports usually charge more than the better categories, as denoted by the low values of C2. With the integration of ports into supply chains, intermodal transport has become the backbone of maritime containerised transport. Prior research has shown the significance of intermodal transport service costs, emphasising that costs and time as crucial considerations in the selection of intermodal transport services (Vural et al., 2020). Customers tend to choose the port that offers economically efficient intermodal transport services (Lun et al., 2016). Ports can enhance communication with different transportation service providers and improve service efficiency by building BDAC to deliver cargo to consumers using the most economical routes (Vural et al., 2020). Thus, C2 is among the variables that a port should work on to improve PP. The last relevant variable is C1, which assesses how much the port charges for cargo handling compared with its competitors. In the face of increasing competition between ports, it has become imperative for them to operate with minimal delay, optimal efficiency, and equitable pricing in order to attract customers' attention (Fahim et al., 2022). Thus, C1 is significant, and ports need to provide high-quality service and keep costs low to enhance PP. Considering the PP's DT, there are three ports that have high performance for SQ2 (higher than the medium cluster) but perform poorly for C1. A possible justification for this classification could be related to the nature of the cost variables. Indeed, the cost variables are characterised by a high correlation (>0.8). Accordingly, a low value of C1 usually corresponds to low values of C2 and C3. Based on this consideration, it can be assumed that the ports that end up in the worst cluster may perform better than the medium ports with regard to SQ2, but they usually have low performance related to the cost sub-dimension. That said, the third leaf of the PP's DT could be regarded as less reliable.

5.2. Comparison with linear regression

To evaluate the robustness of the proposed framework, the results are compared with those obtained from linear regression.

For the linear regression analysis, eight different regression models (one for each variable of PSP and PP) with seven independent variables (the sub-dimensions of BDAC) were implemented. The mean values of each sub-dimension are considered for this approach as well. The results obtained from each regression model are shown in Table 3, which reports the slopes of the regression models.

The result show that each sub-dimension of PSP and PP has at least one statistically significant positive relationship (i.e. positive slope) with a sub-dimension of BDAC. Only ECD is characterised by weaker relationships since the only influential sub-dimensions of BDAC are DDC and OL, both with a low significance level ($p - value < 0.15$). Accordingly, it can be concluded that there is a certain degree of association between BDAC and PSP/PP. This is similar to what is depicted by the other approaches (including the one proposed in this paper). In addition, the regression models underline the importance of DDC since there are positive associations with one and four sub-dimensions of PSP and PP, respectively. Compared to the proposed framework, linear regression is easier and more straightforward to implement. Moreover, it defines which sub-dimensions of BDAC are positively associated with the sub-dimensions of PSP and PP. However, it does not highlight which sub-dimensions are most important for BDAC, PSP, and PP. Thus, the regression models and the proposed framework could be employed as complementary tools to acquire more in-depth knowledge.

Table 3. Relationships among BDAC's dimensions, PSP, and PP's dimensions.

	Dependent variable							
	PSP dimension			PP dimension				
BDAC independent variable	ED	SD	ECD	C	SQ	OE	VAS	CO
D	−0.061	0.097	0.216	0.282	−0.118	0.365*	0.129	−0.230
T	0.056	−0.045	0.057	0.193	0.505**	−0.070	−0.013	0.051
BR	0.502**	0.082	−0.107	−0.236	0.441*	−0.082	0.211*	0.020
TS	0.191	0.181	−0.036	0.026	−0.150	−0.077	0.404***	0.084
MS	0.019	0.012	0.114	0.058	0.223	0.079	0.080	0.156
DDC	−0.198	0.469**	0.319°	0.767***	0.683**	0.361**	−0.098	0.512***
OL	0.422*	−0.093	0.331°	−0.006	0.032	0.076	0.189	0.203
F for the regression	7.89***	4.36**	6.46***	11.71***	2.47**	7.77***	22.57***	8.12***
R^2(%)	63.32	48.8	58.57	71.93	35.07	62.96	79.47	63.97

****p*-value < 0.001, ***p*-value < 0.05, **p*-value < 0.1, °p-value < 0.15.

6. Conclusion

6.1. Theoretical contributions

Within the context of port operations, this paper develops a framework for creating a DSS capable of evaluating port performance based on the integration of expert judgements and ML. Second, this work investigates the impact of BDAC on PSP and PP. To accomplish the first task, an extensive questionnaire based on a fuzzy Likert scale was proposed to different employees of 40 ports. The collected answers were aggregated through an improved SAM, and subsequently the ports were clustered through K-Means. The diversity among the clusters was evaluated through MANOVA and PERMANOVA. Next, three DTs were developed based on the cluster associated with the macro-dimensions of each port. The trained DTs serve two distinct purposes, since they are both a DSS for guiding choices and descriptive tools that provide an overview of the identified clusters. To fulfil the second task, a BN was built to study the possible association between BDAC and PSP/PP.

The main theoretical finding of this study is that there appears to be a relationship between the BDAC level and port performance from both operational and sustainability perspectives. Indeed, the ports that fall into the best cluster for BDAC are also associated with the best clusters for PSP and PP. This finding echoes recent studies that call for investigating the impact of BDAC on ports (de la Peña Zarzuelo et al., 2020; Munim et al., 2020). Furthermore, the results show ports classified as worst for BDAC are more likely to show low levels of PSP and PP. Hence, the study extends existing research (Ciampi et al., 2021; Jha et al., 2020; Wamba & Akter, 2019; Yasmin et al., 2020) by using a K-means approach to examine the association between BDAC and PP/PSP. Finally, the developed DTs provide insights into which variables are associated with the cluster classification of ports with regard to BDAC, PSP, and PP. Although recent studies suggest the importance of human resources and intangible resources for developing BDAC (Jeble et al., 2018; Lozada et al., 2019; Mikalef et al., 2020), few researchers have empirically investigated how to allocate organisational resources to develop greater BDAC. This research fills this gap by developing a DT to identify the resources that a port should work on to improve BDAC.

6.2. Managerial implications

Regarding managerial implications, the main outcome of this work is represented by the developed DTs, which can aid ports in performing a preliminary analysis to identify their weakest macro-dimension. Indeed, the DTs are built based on the clusters associated with each port. The DTs provide an interpretable approximation of the cluster assignments and can serve as a preliminary diagnostic tool. It is worth mentioning that the DTs should be used only as an initial tool to provide a broad overview of port performance. In fact, a more in-depth analysis is required to understand where the real weaknesses lie. However, the developed framework is general and could be employed to study different sectors or fields.

6.3. Limitations and future research

Among the limitations of the present work, the initial sample size of only 40 ports is worth mentioning. Accordingly, it would be useful and interesting to extend the study to other ports. This could allow more significant results and insights to be obtained. However, the process of sharing and explaining the questionnaire, along with receiving the answers, is very time-consuming. Furthermore, both K-Means and DTs depend on the random seed set; accordingly, changing it could lead to slightly different results. Thus, it would be useful to perform a sensitivity analysis on both the clustering and the DSS. Moreover, only one clustering algorithm was considered, and no feature selection was performed. Therefore, it would be interesting to perform a sensitivity analysis by varying the clustering and classification algorithms, along with adopting a feature selection technique, such as Neighbourhood Component Analysis (Goldberger et al., 2004), to evaluate whether higher accuracy could be obtained. Another relevant limitation is that only the association between BDAC and PP/PSP was investigated, and it was found that some ports perform poorly for BDAC but still have good PSP and PP. Consequently, it is reasonable to assume that other factors (e.g. infrastructure maturity and environmental pressure) can positively impact PSP and PP. These factors may be investigated in future studies. Finally, another limitation is that the variables of PSP and PP partly capture analytics-enabled perceived performance. This creates a conceptual overlap with BDAC, potentially inflating the positive association between BDAC and PP/PSP.

Author contributions

CRediT: **Leonardo Leoni:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing; **Xiaotian Xie:** Writing – original draft; **Guoqing Zhao:** Investigation, Writing – review & editing; **Yi Wang:** Writing – review & editing; **Filippo De Carlo:** Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Data availability statement

The data used to support the findings of this study are available from the author upon request.

References

AbuAlhaol, I., Falcon, R., Abielmona, R., & Petriu, E. (2018). Mining port congestion indicators from big AIS data. In *2018 International Joint Conference on Neural Networks (IJCNN)* (pp. 1–8).

Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*, *182*, 113–131. https://doi.org/10.1016/j.ijpe.2016.08.018

Alamoush, A.S., Ballini, F., & Ölçer, A.I. (2022). Ports, maritime transport, and industry: The immediate impact of COVID-19 and the way forward. *Maritime Technology and Research*, *4*(1), 250092–250092. https://doi.org/10.33175/mtr.2022.250092

Alyami, H., Yang, Z., Riahi, R., Bonsall, S., & Wang, J. (2019). Advanced uncertainty modelling for container port risk analysis. *Accident Analysis and Prevention*, *123*, 411–421. https://doi.org/10.1016/j.aap.2016.08.007

Anderson, M.J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, *26*(1), 32–46. https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x

Anderson, M.J. (2014). Permutational multivariate analysis of variance (PERMANOVA). *Wiley StatsRef: Statistics Reference Online*, 1–15.

Arena, S., Florian, E., Zennaro, I., Orrù, P.F., & Sgarbossa, F. (2022). A novel decision support system for managing predictive maintenance strategies based on machine learning approaches. *Safety Science*, *146*, 105529. https://doi.org/10.1016/j.ssci.2021.105529

Attanasio, G., Battistella, C., & Chizzolini, E. (2023). The future of energy management: Results of a Delphi panel applied in the case of ports. *Journal of Cleaner Production*, *417*, 137947. https://doi.org/10.1016/j.jclepro.2023.137947

Attardi, R., De Rosa, F., & DiPalma, M. (2015). From visual features to shared future visions for Naples 2050. *Applied Spatial Analysis and Policy*, *8*(3), 249–271. https://doi.org/10.1007/s12061-015-9143-5

Azzam, I.A., Al-Khatib, S.F., & Albataineh, W.M. (2021). Strategic port classification: International clustering-based approach for decision-making optimization. *Journal of Public Affairs*, *21*(1), e1963. https://doi.org/10.1002/pa.1963

Bai, C., Kusi-Sarpong, S., Badri Ahmadi, H., & Sarkis, J. (2019). Social sustainable supplier evaluation and selection: A group decision-support approach. *International Journal of Production Research*, *57*(22), 7046–7067. https://doi.org/10.1080/00207543.2019.1574042

Bai, C., Sarkis, J., & Ibrahim, S. (2023). An analytical method for investigating multi-tier sustainable supply chains: Simplifying the complex. *International Journal of Production Research*, *61*(1), 1–18. https://doi.org/10.1080/00207543.2022.2131930

Banan-Dallalian, M., Shokatian-Beiragh, M., Golshani, A., & Abdi, A. (2023). Use of a Bayesian network for storm-induced flood risk assessment and effectiveness of ecosystem-based risk reduction measures in coastal areas (Port of Sur, sultanate of Oman). *Ocean Engineering*, *270*, 113662. https://doi.org/10.1016/j.oceaneng.2023.113662

Bejjani, Y., & Vanhaverbeke, W. (2025). Advancing big data decision-making capabilities: Findings from a systematic literature review and a qualitative inquiry. *Journal of Decision Systems*, *34*(1), 2581374. https://doi.org/10.1080/12460125.2025.2581374

Bhattacharya, S., Govindan, K., Dastidar, S.G., & Sharma, P. (2024). Applications of artificial intelligence in closed-loop supply chains: Systematic literature review and future research agenda. *Transportation Research Part E: Logistics and Transportation Review*, *184*, 103455. https://doi.org/10.1016/j.tre.2024.103455

Binkin, N., Sullivan, K., Staehling, N., & Nieburg, P. (1992). Rapid nutrition surveys: How many clusters are enough? *Disasters*, *16*(2), 97–103. https://doi.org/10.1111/j.1467-7717.1992.tb00383.x

Bray, S., Caggiani, L., & Ottomanelli, M. (2015). Measuring transport systems efficiency under uncertainty by fuzzy sets theory based data envelopment analysis: Theoretical and practical comparison with traditional DEA model. *Transportation Research Procedia*, *5*, 186–200. https://doi.org/10.1016/j.trpro.2015.01.005

Bregar, A. (2024). Decision process for industry 5.0. *Journal of Decision Systems*, *33*(sup1), 433–447. https://doi.org/10.1080/12460125.2024.2359158

Cabral, A.M.R., & de Sousa Ramos, F. (2014). Cluster analysis of the competitiveness of container ports in Brazil. *Transportation Research Part A: Policy and Practice*, *69*, 423–431. https://doi.org/10.1016/j.tra.2014.09.005

Cankaya, B., Eren Tokgoz, B., Dag, A., & Santosh, K.C. (2021). Development of a machine-learning-based decision support mechanism for predicting chemical tanker cleaning activity. *Journal of Modelling in Management*, *16*(4), 1138–1165. https://doi.org/10.1108/JM2-12-2019-0284

Cantini, A., Peron, M., De Carlo, F., & Sgarbossa, F. (2022). A decision support system for configuring spare parts supply chains considering different manufacturing technologies. *International Journal of Production Research*, *62*(8), 1–21. https://doi.org/10.1080/00207543.2022.2041757

Castellano, R., Fiore, U., Musella, G., Perla, F., Punzo, G., Risitano, M., Sorrentino, A., & Zanetti, P. (2019). Do digital and communication technologies improve smart ports? A fuzzy DEA approach. *IEEE Transactions on Industrial Informatics*, *15*(10), 5674–5681. https://doi.org/10.1109/TII.2019.2927749

Celik, E., & Akyuz, E. (2018). An interval type-2 fuzzy AHP and TOPSIS methods for decision-making problems in maritime transportation engineering: The case of ship loader. *Ocean Engineering*, *155*, 371–381. https://doi.org/10.1016/j.oceaneng.2018.01.039

Chen, M., Pu, X., Zhang, M., Cai, Z., Chong, A.Y.-L., & Tan, K.H. (2022). Data analytics capability and servitization: The moderated mediation role of bricolage and innovation orientation. *International Journal of Operations and Production Management*, *42*(4), 440–470. https://doi.org/10.1108/IJOPM-10-2021-0663

Ciampi, F., Demi, S., Magrini, A., Marzi, G., & Papa, A. (2021). Exploring the impact of big data analytics capabilities on business model innovation: The mediating role of entrepreneurial orientation. *Journal of Business Research*, *123*, 1–13. https://doi.org/10.1016/j.jbusres.2020.09.023

Cinelli, M., Kadziński, M., Miebs, G., Gonzalez, M., & Słowiński, R. (2022). Recommending multiple criteria decision analysis methods with a new taxonomy-based decision support system. *European Journal of Operational Research*, *302*(2), 633–651. https://doi.org/10.1016/j.ejor.2022.01.011

Coşkun, S.S., Kumru, M., & Kan, N.M. (2022). An integrated framework for sustainable supplier development through supplier evaluation based on sustainability indicators. *Journal of Cleaner Production*, *335*, 130287. https://doi.org/10.1016/j.jclepro.2021.130287

Cremen, G., Bozzoni, F., Pistorio, S., & Galasso, C. (2022). Developing a risk-informed decision-support system for earthquake early warning at a critical seaport. *Reliability Engineering & System Safety*, *218*, 108035. https://doi.org/10.1016/j.ress.2021.108035

DeBrusk, C. (2018). *The risk of machine-learning bias (and how to prevent it)*. MIT Sloan Manageemnt Review. Retrieved June 01, 2026, from https://sloanreview.mit.edu/article/the-risk-of-machine-learning-bias-and-how-to-prevent-it/

de la Peña Zarzuelo, I., Soeane, M.J.F., & Bermúdez, B.L. (2020). Industry 4.0 in the port and maritime industry: A literature review. *Journal of Industrial Information Integration*, *20*, 100173. https://doi.org/10.1016/j.jii.2020.100173

Del Giudice, M., DiVaio, A., Hassan, R., & Palladino, R. (2022). Digitalization and new technologies for sustainable business models at the ship-port interface: A bibliometric analysis. *Maritime Policy & Management*, *49*(3), 410–446. https://doi.org/10.1080/03088839.2021.1903600

Di Vaio, A., & Varriale, L. (2020). Digitalization in the sea-land supply chain: Experiences from Italy in rethinking the port operations within inter-organizational relationships. *Production Planning & Control*, *31*(2–3), 220–232. https://doi.org/10.1080/09537287.2019.1631464

Dubey, R., Gunasekaran, A., Childe, S.J., Papadopoulos, T., Luo, Z., Wamba, S.F., & Roubaud, D. (2019). Can big data and predictive analytics improve social and environmental sustainability? *Technological Forecasting and Social Change*, *144*, 534–545. https://doi.org/10.1016/j.techfore.2017.06.020

Ebrahim, R.A., Singh, S., Li, Y., & Ji, W. (2022). Discrete event simulation for port berth maintenance planning. *2022 Winter Simulation Conference*. https://doi.org/10.1109/WSC57314.2022.10015296

Emerson, R.W. (2015). Convenience sampling, random sampling, and snowball sampling: How does sampling affect the validity of research? *Journal of Visual Impairment & Blindness*, *109*(2), 164–168. https://doi.org/10.1177/0145482X1510900215

Fahim, P.B., Rezaei, J., Montreuil, B., & Tavasszy, L. (2022). Port performance evaluation and selection in the physical internet. *Transport Policy*, *124*, 83–94. https://doi.org/10.1016/j.tranpol.2021.07.013

Filom, S., Amiri, A.M., & Razavi, S. (2022). Applications of machine learning methods in port operations-a systematic literature review. *Transportation Research Part E: Logistics and Transportation Review*, *161*, 102722. https://doi.org/10.1016/j.tre.2022.102722

Fosso Wamba, S., Queiroz, M.M., Wu, L., & Sivarajah, U. (2024). Big data analytics-enabled sensing capability and organizational outcomes: Assessing the mediating effects of business analytics culture. *Annals of Operations Research*, *333*(2), 559–578. https://doi.org/10.1007/s10479-020-03812-4

Fuentes, G. (2021). Generating bunkering statistics from AIS data: A machine learning approach. *Transportation Research Part E: Logistics and Transportation Review*, *155*, 102495. https://doi.org/10.1016/j.tre.2021.102495

Gao, Y., Chang, D., Fang, T., & Fan, Y. (2019). The daily container volumes prediction of storage yard in port with long short-term memory recurrent neural network. *Journal of Advanced Transportation*, *2019*, 1–11. https://doi.org/10.1155/2019/5764602

Garg, C.P., Kashav, V., & Wang, X. (2022). Evaluating sustainability factors of green ports in China under fuzzy environment. *Environment, Development, and Sustainability*, *25*(8), 1–27. https://doi.org/10.1007/s10668-022-02375-7

Ghazal, T.M., Hussain, M.Z., Said, R.A., Nadeem, A., Hasan, M.K., Ahmad, M., Khan, M.A., & Naseem, M.T. (2021). Performances of k-means clustering algorithm with different distance metrics. *Intelligent Automation & Soft Computing*, *29*(3), 735–742. https://doi.org/10.32604/iasc.2021.019067

Gohari, A., Ahmad, A.B., Balasbaneh, A.T., Gohari, A., Hasan, R., & Sholagberu, A.T. (2022). Significance of intermodal freight modal choice criteria: MCDM-based decision support models and SP-based modal shift policies. *Transport Policy*, *121*, 46–60. https://doi.org/10.1016/j.tranpol.2022.03.015

Goldberger, J., Hinton, G.E., Roweis, S., & Salakhutdinov, R.R. (2004). Neighbourhood components analysis. *Advances in Neural Information Processing Systems*, 17.

Guo, X., Ji, J., Khan, F., Ding, L., & Yang, Y. (2021). Fuzzy Bayesian network based on an improved similarity aggregation method for risk assessment of storage tank accident. *Process Safety and Environmental Protection*, *149*, 817–830. https://doi.org/10.1016/j.psep.2021.03.017

Gupta, M., & George, J.F. (2016). Toward the development of a big data analytics capability. *Information & Management*, *53*(8), 1049–1064. https://doi.org/10.1016/j.im.2016.07.004

Hsu, C.-T., Chou, M.-T., & Ding, J.-F. (2023). Key factors for the success of smart ports during the post-pandemic era. *Ocean & Coastal Management*, *233*, 106455. https://doi.org/10.1016/j.ocecoaman.2022.106455

Huang, S., Kang, Z., Xu, Z., & Liu, Q. (2021). Robust deep k-means: An effective and simple method for data clustering. *Pattern Recognition*, *117*, 107996. https://doi.org/10.1016/j.patcog.2021.107996

Jaidee, R., Sumrit, D., & Vanichchinchai, A. (2022). Identifying sustainable roll-on/roll-off seaport assessment criteria using the fuzzy Delphi method: A case study of Thailand. In *2022 10th International Conference on Traffic and Logistic Engineering (ICTLE)* (pp. 101–106).

Jain, N., Singh, A.R., & Upadhyay, R.K. (2020). Sustainable supplier selection under attractive criteria through FIS and integrated fuzzy MCDM techniques. *International Journal of Sustainable Engineering*, *13*(6), 441–462. https://doi.org/10.1080/19397038.2020.1737751

Jeble, S., Dubey, R., Childe, S.J., Papadopoulos, T., Roubaud, D., & Prakash, A. (2018). Impact of big data and predictive analytics capability on supply chain sustainability. *The International Journal of Logistics Management*, *29*(2), 513–538. https://doi.org/10.1108/IJLM-05-2017-0134

Jha, A.K., Agi, M.A., & Ngai, E.W. (2020). A note on big data analytics capability development in supply chain. *Decision Support Systems*, *138*, 113382. https://doi.org/10.1016/j.dss.2020.113382

Jianxing, Y., Shibo, W., Yang, Y., Haicheng, C., Haizhao, F., Jiahao, L., & Shenwei, G. (2021). Process system failure evaluation method based on a noisy-OR gate intuitionistic fuzzy Bayesian network in an uncertain environment. *Process Safety and Environmental Protection*, *150*, 281–297. https://doi.org/10.1016/j.psep.2021.04.024

Jović, M., Tijan, E., Marx, R., & Gebhard, B. (2019). Big data management in maritime transport. *Pomorski Zbornik*, *57*(1), 123–141. https://doi.org/10.18048/2019.57.09

Jraisat, L., Upadhyay, A., Ghalia, T., Jresseit, M., Kumar, V., & Sarpong, D. (2021). Triads in sustainable supply-chain perspective: Why is a collaboration mechanism needed? *International Journal of Production Research*, *61*(14), 1–17. https://doi.org/10.1080/00207543.2021.1936263

Jun, S., & Lee, S. (2021). Learning dispatching rules for single machine scheduling with dynamic arrivals based on decision trees and feature construction. *International Journal of Production Research*, *59*(9), 2838–2856. https://doi.org/10.1080/00207543.2020.1741716

Kim, S., Kang, D., & Dinwoodie, J. (2016). Competitiveness in a multipolar port system: Striving for regional gateway status in Northeast Asia. *The Asian Journal of Shipping and Logistics*, *32*(2), 119–125. https://doi.org/10.1016/j.ajsl.2016.05.005

Kong, Y., Liu, J., & Chen, J. (2022). Exploring the carbon abatement measures in maritime supply chain: A scenario-based system dynamics approach. *International Journal of Production Research*, *61*(18), 1–22. https://doi.org/10.1080/00207543.2022.2088427

Kumar Sharma, R., Khan, S., Singh, R., & Birari, A. (2024). Speculative investment decisions in cryptocurrency: A structural equation modelling approach. *Journal of Decision Systems*, 1–22. https://doi.org/10.1080/12460125.2024.2341194

Lazim, A., & Wahab, N. (2010). A fuzzy decision making approach in evaluating ferry service quality. *Management Research and Practice*, *2*(1), 94–107.

Leoni, L., & De Carlo, F. (2023). Integration of fuzzy reliability analysis and consequence simulation to conduct risk assessment. *Journal of Loss Prevention in the Process Industries*, *83*, 105081. https://doi.org/10.1016/j.jlp.2023.105081

Li, H., Jiao, H., & Yang, Z. (2023). Ais data-driven ship trajectory prediction modelling and analysis based on machine learning and deep learning methods. *Transportation Research Part E: Logistics and Transportation Review*, *175*, 103152. https://doi.org/10.1016/j.tre.2023.103152

Li, L., Lin, J., Ouyang, Y., & Luo, X.R. (2022). Evaluating the impact of big data analytics usage on the decision-making quality of organizations. *Technological Forecasting and Social Change*, *175*, 121355. https://doi.org/10.1016/j.techfore.2021.121355

Lim, S., Pettit, S., Abouarghoub, W., & Beresford, A. (2019). Port sustainability and performance: A systematic literature review. *Transportation Research Part D: Transport and Environment*, *72*, 47–64. https://doi.org/10.1016/j.trd.2019.04.009

Lirn, T., Jim Wu, Y.-C., & Chen, Y.J. (2013). Green performance criteria for sustainable ports in Asia. *International Journal of Physical Distribution & Logistics Management*, *43*(5/6), 427–451. https://doi.org/10.1108/IJPDLM-04-2012-0134

Liu, C., Liu, J., Zhou, X., Zhao, Z., Wan, C., & Liu, Z. (2020). Ais data-driven approach to estimate navigable capacity of busy waterways focusing on ships entering and leaving port. *Ocean Engineering*, *218*, 108215. https://doi.org/10.1016/j.oceaneng.2020.108215

Lorenz, R., Kraus, M., Wolf, H., Feuerriegel, S., & Netland, T.H. (2022). Selecting advanced analytics in manufacturing: A decision support model. *Production Planning & Control*, 1–14. https://doi.org/10.1080/09537287.2022.2126951

Lozada, N., Arias-Pérez, J., & Perdomo-Charry, G. (2019). Big data analytics capability and co-innovation: An empirical study. *Heliyon*, *5*(10), e02541. https://doi.org/10.1016/j.heliyon.2019.e02541

Lun, Y.V., Shang, K., Lai, K., & Cheng, T.C.E. (2016). Examining the influence of organizational capability in innovative business operations and the mediation of profitability on customer satisfaction: An application in intermodal transport operators in Taiwan. *International Journal of Production Economics*, *171*, 179–188. https://doi.org/10.1016/j.ijpe.2015.02.019

Mahmud, K.K., Chowdhury, M.M.H., & Shaheen, M.M.A. (2024). Green port management practices for sustainable port operations: A multi method study of Asian ports. *Maritime Policy & Management*, *51*(8), 1902–1937. https://doi.org/10.1080/03088839.2023.2258125

Majhi, S.G., Chowdhury, P., Mukherjee, A., & Anand, A. (2026). Building a cognitive analytics capability: A resource-based perspective. *Journal of Decision Systems*, *35*(1), 2599919. https://doi.org/10.1080/12460125.2025.2599919

Mansouri, F.S., Allal, A., Mansouri, K., & Qbadou, M. (2021). A model for analyzing capacity of ports to accommodate autonomous ships using k-means cluster analysis: A case study. *International Journal on Technical and Physical Problems of Engineering*, *13*, 68–75.

Martinez Arbizu, P. (2020). *PairwiseAdonis: Pairwise multilevel comparison using adonis. R package version 0.4* https://github.com/pmartinezarbizu/pairwiseAdonis

Martínez-Mendoza, E., Rivas-Tovar, L.A., Fernández-Echeverría, E., & Fernández-Lambert, G. (2020). Social impact of wind energy in the Isthmus of Tehuantepec, Mexico, using Likert-fuzzy. *Energy Strategy Reviews*, *32*, 100567. https://doi.org/10.1016/j.esr.2020.100567

Matthias, O., Fouweather, I., Gregory, I., & Vernon, A. (2017). Making sense of big data-can it transform operations management? *International Journal of Operations and Production Management*, *37*(1), 37–55. https://doi.org/10.1108/IJOPM-02-2015-0084

Melnyk, O., Onyshchenko, S., Rudenko, S., Pavlova, N., & Muradian, A. (2024). Embracing integrated optimization strategies for cost-effective port operations in marine logistics. *Systems, Decision and Control in Energy VI: Volume I: Energy Informatics and Transport: Springer*, 491–508. https://doi.org/10.1007/978-3-031-68372-5_27

Mengüç, K., Aydin, N., & Ulu, M. (2023). Optimisation of COVID-19 vaccination process using GIS, machine learning, and the multi-layered transportation model. *International Journal of Production Research*, *63*(2), 1–14. https://doi.org/10.1080/00207543.2023.2182151

Meski, O., Belkadi, F., Laroche, F., Ritou, M., & Furet, B. (2021). A generic knowledge management approach towards the development of a decision support system. *International Journal of Production Research*, *59*(22), 6659–6676. https://doi.org/10.1080/00207543.2020.1821930

Mikalef, P., Krogstie, J., Pappas, I.O., & Pavlou, P. (2020). Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. *Information & Management*, *57*(2), 103169. https://doi.org/10.1016/j.im.2019.05.004

Molavi, A., Lim, G.J., & Race, B. (2020). A framework for building a smart port and smart port index. *International Journal of Sustainable Transportation*, *14*(9), 686–700. https://doi.org/10.1080/15568318.2019.1610919

Moscoso-López, J.A., Urda, D., Ruiz-Aguilar, J.J., Gonzalez-Enrique, J., & Turias, I.J. (2021). A machine learning-based forecasting system of perishable cargo flow in maritime transport. *Neurocomputing*, *452*, 487–497. https://doi.org/10.1016/j.neucom.2019.10.121

Mrva, J., Neupauer, Š., Hudec, L., Ševcech, J., & Kapec, P. (2019). Decision support in medical data using 3D decision tree visualisation. In *2019 E-Health and Bioengineering Conference (EHB)* (pp. 1–4).

Munim, Z.H., Dushenko, M., Jimenez, V.J., Shakil, M.H., & Imset, M. (2020). Big data and artificial intelligence in the maritime industry: A bibliometric review and future research directions. *Maritime Policy & Management*, *47*(5), 577–597. https://doi.org/10.1080/03088839.2020.1788731

Nguyen, T.Q., Ngo, L.T.T., Huynh, N.T., Quoc, T.L., & Hoang, L.V. (2022). Assessing port service quality: An application of the extension fuzzy AHP and importance-performance analysis. *PLOS ONE*, *17*(2), e0264590. https://doi.org/10.1371/journal.pone.0264590

Oh, H., Lee, S.-W., & Seo, Y.-J. (2018). The evaluation of seaport sustainability: The case of South Korea. *Ocean & Coastal Management*, *161*, 50–56. https://doi.org/10.1016/j.ocecoaman.2018.04.028

Olan, F., Spanaki, K., Ahmed, W., & Zhao, G. (2024). Enabling explainable artificial intelligence capabilities in supply chain decision support making. *Production Planning & Control*, *36*(6), 1–12. https://doi.org/10.1080/09537287.2024.2313514

Orero-Blat, M., Palacios-Marqués, D., Leal-Rodriguez, A.L., & Ferraris, A. (2025). Beyond digital transformation: A multi-mixed methods study on big data analytics capabilities and innovation in enhancing organizational performance. *Review of Managerial Science*, *19*(2), 649–685. https://doi.org/10.1007/s11846-024-00768-8

Pahuja, H., Delir Haghighi, P., & Jayaraman, P.P. (2024). Data-driven decision-support system for expansions in retail stores: Dexter. *Journal of Decision Systems*, *33*(sup1), 222–233. https://doi.org/10.1080/12460125.2024.2349428

Pamucar, D., & Görçün, Ö.F. (2022). Evaluation of the European container ports using a new hybrid fuzzy LBWA-CoCoSo'B techniques. *Expert Systems With Applications*, *203*, 117463. https://doi.org/10.1016/j.eswa.2022.117463

Pan, I., Mason, L.R., & Matar, O.K. (2022). Data-centric engineering: Integrating simulation, machine learning and statistics. Challenges and opportunities. *Chemical Engineering Science*, *249*, 117271. https://doi.org/10.1016/j.ces.2021.117271

Pang, Q., Du, J., Fang, M., & Wang, L. (2026). Strategic mechanism for enhanced sustainable practice performance in shipping organizations through big data analytics powered by artificial intelligence. *Journal of Enterprise Information Management*, *39*(1), 188–213. https://doi.org/10.1108/JEIM-05-2024-0233

Pasman, H.J., & Rogers, W.J. (2020). How to treat expert judgment? With certainty it contains uncertainty! *Journal of Loss Prevention in the Process Industries*, *66*, 104200. https://doi.org/10.1016/j.jlp.2020.104200

Peng, W., Bai, X., Yang, D., Yuen, K.F., & Wu, J. (2022). A deep learning approach for port congestion estimation and prediction. *Maritime Policy & Management*, *50*(7), 1–26. https://doi.org/10.1080/03088839.2022.2057608

Puig, M., & Darbra, R.M. (2024). Innovations and insights in environmental monitoring and assessment in port areas. *Current Opinion in Environmental Sustainability*, *70*, 101472. https://doi.org/10.1016/j.cosust.2024.101472

Raad, N.G., Rajendran, S., & Salimi, S. (2022). A novel three-stage fuzzy GIS-MCDA approach to the dry port site selection problem: A case study of Shahid Rajaei Port in Iran. *Computers & Industrial Engineering*, *168*, 108112. https://doi.org/10.1016/j.cie.2022.108112

Raeesi, R., Sahebjamnia, N., & Mansouri, S.A. (2022). The synergistic effect of operational research and big data analytics in greening container terminal operations: A review and future directions. *European Journal of Operational Research*, *310*(3), 943–973. https://doi.org/10.1016/j.ejor.2022.11.054

Raeesi, R., Sahebjamnia, N., & Mansouri, S.A. (2023). The synergistic effect of operational research and big data analytics in greening container terminal operations: A review and future directions. *European Journal of Operational Research*, *310*(3), 943–973. https://doi.org/10.1016/j.ejor.2022.11.054

Raissi, N., Hakeem, A., & Khallouli, W. (2025). Big data analytics capability and strategic metrics for enhancing firm innovation: A PLS-SEM modelling approach. *Journal of Decision Systems*, *34*(1), 2486407. https://doi.org/10.1080/12460125.2025.2486407

Rapaccini, M., Cadonna, V.L., Leoni, L., & De Carlo, F. (2022). Application of machine learning techniques for cost estimation of engineer to order products. *International Journal of Production Research*, *61*(20), 1–23. https://doi.org/10.1080/00207543.2022.2141907

Raut, R.D., Mangla, S.K., Narwane, V.S., Dora, M., & Liu, M. (2021). Big data analytics as a mediator in lean, agile, resilient, and green (LARG) practices effects on sustainable supply chains. *Transportation Research Part E: Logistics and Transportation Review*, *145*, 102170. https://doi.org/10.1016/j.tre.2020.102170

Rotunno, G., Lo Zupone, G., Carnimeo, L., & Fanti, M.P. (2023). Discrete event simulation as a decision tool: A cost benefit analysis case study. *Journal of Simulation*, 1–17. https://doi.org/10.1080/17477778.2023.2167618

Sabouhi, F., Jabalameli, M.S., & Jabbarzadeh, A. (2021). An optimization approach for sustainable and resilient supply chain design with regional considerations. *Computers & Industrial Engineering*, *159*, 107510. https://doi.org/10.1016/j.cie.2021.107510

Salleh, N.H.M., Riahi, R., Yang, Z., & Wang, J. (2014). Risk assessment of liner shipping from a business environment perspective. *Vulnerability, Uncertainty, and Risk: Quantification, Mitigation, and Management*, 2320–2329. https://doi.org/10.1061/9780784413609.233

Salvador, S., & Chan, P. (2004). Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms. In *16th IEEE International Conference on Tools with Artificial Intelligence* (pp. 576–584). https://doi.org/10.1109/ICTAI.2004.50

Santino, C.N., de Oliveira Fontes, C.H., & Rodríguez, J.L.M. (2022). Model for analyzing the critical variables of an industrial process using fuzzy inference system and Likert scale. *Arabian Journal for Science & Engineering*, *47*(7), 8677–8690. https://doi.org/10.1007/s13369-021-06364-2

Sarkar, B.D., Shankar, R., & Kar, A.K. (2022). Severity analysis and risk profiling of port logistics barriers in the industry 4.0 era. *Benchmarking: An International Journal*, *30*(9), ahead-of–print. https://doi.org/10.1108/BIJ-03-2022-0153

Saunila, M., Nasiri, M., Ukko, J., & Rantala, T. (2019). Smart technologies and corporate sustainability: The mediation effect of corporate sustainability strategy. *Computers in Industry*, *108*, 178–185. https://doi.org/10.1016/j.compind.2019.03.003

Seo, Y.-J., Dinwoodie, J., & Roe, M. (2016). The influence of supply chain collaboration on collaborative advantage and port performance in maritime logistics. *International Journal of Logistics Research and Applications*, *19*(6), 562–582. https://doi.org/10.1080/13675567.2015.1135237

Sgarbossa, F., Peron, M., Lolli, F., & Balugani, E. (2021). Conventional or additive manufacturing for spare parts management: An extensive comparison for Poisson demand. *International Journal of Production Economics*, *233*, 107993. https://doi.org/10.1016/j.ijpe.2020.107993

Simoni, M., Schiavone, F., Risitano, M., Leone, D., & Chen, J. (2022). Group-specific business process improvements via a port community system: The case of Rotterdam. *Production Planning & Control*, *33*(4), 371–385. https://doi.org/10.1080/09537287.2020.1824029

Song, M., Cen, L., Zheng, Z., Fisher, R., Liang, X., Wang, Y., & Huisingh, D. (2017). How would big data support societal development and environmental sustainability? Insights and practices. *Journal of Cleaner Production*, *142*, 489–500. https://doi.org/10.1016/j.jclepro.2016.10.091

Su, C., & Wu, L. (2023). Opportunistic maintenance optimisation for offshore wind farm with considering random wind speed. *International Journal of Production Research*, *62*(5), 1–17. https://doi.org/10.1080/00207543.2023.2202280

Su, X., Zeng, W., Zheng, M., Jiang, X., Lin, W., & Xu, A. (2022). Big data analytics capabilities and organizational performance: The mediating effect of dual innovations. *European Journal of Innovation Management*, *25*(4), 1142–1160. https://doi.org/10.1108/EJIM-10-2020-0431

Tirkolaee, E.B., & Torkayesh, A.E. (2022). A cluster-based stratified hybrid decision support model under uncertainty: Sustainable healthcare landfill location selection. *Applied Intelligence*, *52*(12), 13614–13633. https://doi.org/10.1007/s10489-022-03335-4

Tiwari, S., Sharma, P., Choi, T.-M., & Lim, A. (2023). Blockchain and third-party logistics for global supply chain operations: Stakeholders' perspectives and decision roadmap. *Transportation Research Part E: Logistics and Transportation Review*, *170*, 103012. https://doi.org/10.1016/j.tre.2022.103012

Tseng, P.-H., & Liao, C.-H. (2015). Supply chain integration, information technology, market orientation and firm performance in container shipping firms. *The International Journal of Logistics Management*, *26*(1), 82–106. https://doi.org/10.1108/IJLM-09-2012-0088

Tseng, P.-H., & Pilcher, N. (2019). Evaluating the key factors of green port policies in Taiwan through quantitative and qualitative approaches. *Transport Policy*, *82*, 127–137. https://doi.org/10.1016/j.tranpol.2018.12.014

Tseng, P.-H., & Yip, T.L. (2020). An evaluation model of cruise ports using fuzzy analytic hierarchy process. *Maritime Business Review*, *6*(1), 22–48. https://doi.org/10.1108/MABR-01-2020-0004

Tsolakis, N., Zissis, D., Papaefthimiou, S., & Korfiatis, N. (2022). Towards AI driven environmental sustainability: An application of automated logistics in container port terminals. *International Journal of Production Research*, *60*(14), 4508–4528. https://doi.org/10.1080/00207543.2021.1914355

Uddin, S., Yan, S., & Lu, H. (2024). Machine learning and deep learning in project analytics: Methods, applications and research trends. *Production Planning & Control*, 1–20. https://doi.org/10.1080/09537287.2024.2320790

Udokporo, C., Anosike, A., & Lim, M. (2021). A decision-support framework for lean, agile and green practices in product life cycle stages. *Production Planning & Control*, *32*(10), 789–810. https://doi.org/10.1080/09537287.2020.1764124

Vural, C.A., Roso, V., Halldórsson, Á., Ståhle, G., & Yaruta, M. (2020). Can digitalization mitigate barriers to intermodal transport? An exploratory study. *Research in Transportation Business & Management*, *37*, 100525. https://doi.org/10.1016/j.rtbm.2020.100525

Wamba, S.F., & Akter, S. (2019). Understanding supply chain analytics capabilities and agility for data-rich environments. *International Journal of Operations and Production Management*, *39*(6/7/8), 887–912. https://doi.org/10.1108/IJOPM-01-2019-0025

Wang, C., Li, G., Han, P., Osen, O., & Zhang, H. (2022). Impacts of COVID-19 on ship behaviours in port area: An AIS data-based pattern recognition approach. In *IEEE Transactions on Intelligent Transportation Systems*, *23*(12), 25127–25138. https://doi.org/10.1109/TITS.2022.3147377

Wang, T., Cheng, P., & Zhen, L. (2023). Green development of the maritime industry: Overview, perspectives, and future research opportunities. *Transportation Research Part E: Logistics and Transportation Review*, *179*, 103322. https://doi.org/10.1016/j.tre.2023.103322

Wen-Xin, Z., & Guo-Ping, G. (2021). Research and countermeasures on LNG ship port area navigation risk assessment based on fuzzy comprehensive evaluation method. *IOP Conference Series Earth and Environmental Science*, *809*(1), 012008. https://doi.org/10.1088/1755-1315/809/1/012008

Wong, E.Y.-C., Ling, K.K.-T., & Zhang, X. (2021). Yield and port performance shipping allocation model for revamp service deployments under a dynamic trading landscape. *Transportation Research Part C: Emerging Technologies*, *130*, 103279. https://doi.org/10.1016/j.trc.2021.103279

Xin, X., Liu, K., Loughney, S., Wang, J., & Yang, Z. (2022). Maritime traffic clustering to capture high-risk multi-ship encounters in complex waters. *Reliability Engineering & System Safety*, *230*, 108936. https://doi.org/10.1016/j.ress.2022.108936

Yadav, P., Yadav, S., Singh, D., Kapoor, R.M., & Giri, B.S. (2022). An analytical hierarchy process based decision support system for the selection of biogas up-gradation technologies. *Chemosphere*, *302*, 134741. https://doi.org/10.1016/j.chemosphere.2022.134741

Yasmin, M., Tatoglu, E., Kilic, H.S., Zaim, S., & Delen, D. (2020). Big data analytics capabilities and firm performance: An integrated MCDM approach. *Journal of Business Research*, *114*, 1–15. https://doi.org/10.1016/j.jbusres.2020.03.028

Zain, R.M., Salleh, N.H.M., Zaideen, I.M.M., Menhat, M.N.S., & Jeevan, J. (2022). Dry ports: Redefining the concept of seaport-city integrations. *Transportation Engineering*, *8*, 100112. https://doi.org/10.1016/j.treng.2022.100112

Zangaro, F., Minner, S., & Battini, D. (2021). A supervised machine learning approach for the optimisation of the assembly line feeding mode selection. *International Journal of Production Research*, *59*(16), 4881–4902. https://doi.org/10.1080/00207543.2020.1851793

Zhang, X., Liang, N., & Chen, C. (2022). Dangerous goods container location allocation strategy based on improved NSGA-II algorithm. *International Journal of Advanced Computer Science & Applications*, *13*(8). https://doi.org/10.14569/IJACSA.2022.0130830

Zhang, Z., & Li, Z. (2022). Consensus-based topsis-sort-b for multi-criteria sorting in the context of group decision-making. *Annals of Operations Research*, *325*(2), 1–28. https://doi.org/10.1007/s10479-022-04985-w

Zhao, D., Bao, T., Cullinane, K., & Wang, Y. (2021). Green port performance evaluation under uncertainty: A multiple attribute group decision analysis. *International Journal of Shipping and Transport Logistics*, *13*(1–2), 130–155. https://doi.org/10.1504/IJSTL.2021.112918

Zhen, L., Lv, W., Tan, Z., & Dong, B. (2022). Container transportation scheduling between port yards and the hinterland in Yunfeng. *INFORMS Journal on Applied Analytics*, *52*(3), 250–266. https://doi.org/10.1287/inte.2021.1102

Zhou, C., Xu, J., Miller-Hooks, E., Zhou, W., Chen, C.-H., Lee, L.H., Chew, E.P., & Li, H. (2021). Analytics with digital-twinning: A decision support system for maintaining a resilient port. *Decision Support Systems*, *143*, 113496. https://doi.org/10.1016/j.dss.2021.113496

Appendix

Table A1. Variable description

Construct		Item	Description
Big Data analytics capability (Ciampi et al., 2021; Gupta & George, 2016)	Data	D1	We have access to huge, unstructured (or fast-moving) data for analysis.
		D2	We integrate data from multiple internal sources into a data warehouse or mart for easy access.
		D3	We integrate external with internal data to facilitate high-value analysis of our business environment.
	Technology	T1	We have adopted parallel computing approaches (e.g., Hadoop, Storm, and Spark) to Big Data processing.
		T2	We have adopted different data visualization software (e.g., Sisense, Periscope Data, Tableau, Microsoft Power BI, and IBM Watson Analytics).
		T3	We have adopted cloud services (e.g., IBM Cloud, Amazon Web Services, Microsoft Azure, and Google Cloud) for processing data performing analytics.
	Basic Resources	BR1	We have allocated large funds for Big Data projects.
		BR2	We have enough time to achieve desired results from Big Data analytics projects.
	Technical Skills	TS1	We provide data analytics training to our own employees.
		TS2	We hire new employees with data analytics skills.
		TS3	Our analysts have appropriate data analytics skills to accomplish their jobs successfully.
		TS4	Our analysts have suitable education to fulfill their jobs.
	Managerial Skills	MS1	Our analytics managers can coordinate Big Data-related activities to support other partners.
		MS2	Our analytics managers can anticipate the future business needs of other managers, suppliers, and customers.
		MS3	Our analytics managers have a good sense of where to use Big Data.
		MS4	Our analytics managers can interpret the resources obtained using complex analyses and offer inputs that are useful for swift decision-making.
	Data-driven Culture	DDC1	We consider data a tangible asset.
		DDC2	We base our decisions on data only, not on instinct.
		DDC3	We are willing to override our intuition when data contradict our viewpoints.
		DDC4	We continuously assess and improve business activities in response to insights extracted from data.
	Organisational Learning	OL1	We actively search for new and relevant knowledge.
		OL2	We assimilate new and relevant knowledge.
		OL3	We make concerted efforts to exploit existing competencies and explore new knowledge.
Port sustainability (Dubey et al., 2019; Jeble et al., 2018; Lim et al., 2019; Oh et al., 2018)	Environmental Dimension	ED1	Our port has adopted data analytics technology to reduce air emissions.
		ED2	Our port has adopted data analytics technology to reduce wasted water.
		ED3	Noise in our port areas has been reduced after using data analytics technology.
		ED4	Our port has adopted data analytics technology to reduce oil consumption.
	Social Dimension	SD1	Our port has adopted data analytics technology to improve service quality.
		SD2	The relationship between neighboring residents and our port authorities is improving after building a smart port.
		SD3	Staff security and safety have improved after building a smart port.
		SD4	Our port provides support for employee training and education.
	Economic Dimension	ECD1	Our port offers more employment opportunities.
		ECD2	Our port authorities make multifunctional and efficient use of port areas by data analytics technology.
		ECD3	Our port authorities actively cooperate with industrial and economic development by building a smart port.
		ECD4	Our port is driving the economic development of the area surrounding the port by developing data analytics technology.

(*Continued*)

Table A1. (Continued).

Construct		Item	Description
Port performance (Kim et al., 2016; Seo et al., 2016; Tseng & Liao, 2015)	Cost	C1	By using data analytics technology, our port cargo handling charge is lower than our major competitor.
		C2	By using data analytics technology, our port charges for intermodal transport are lower than our major competitor.
		C3	By using data analytics technology, our port auxiliary service (pilotage, towage, and customers) charge is lower than our major competitor.
	Service Quality	SQ1	By using data analytics technology, our port handles cargo at quoted or anticipated times.
		SQ2	By using data analytics technology, our port handles cargo promptly per customer requirements.
		SQ3	By using data analytics technology, our port's service lead time is shorter than our major competitors.
		SQ4	By using data analytics technology, our port provides shipment information accurately.
	Operational Efficiency	OE1	By using data analytics technology, our terminal productivity is higher than our major competitor.
		OE2	By using data analytics technology, port turn-around time is less (Ship waiting time given congestion) than our major competitor.
		OE3	By using data analytics technology, our time for transportation mode transit is shorter than our major competitor.
	Value-added Service	VAS1	By using data analytics technology, our port can handle different types of cargo.
		VAS2	By using data analytics technology, our port has various services to transfer cargo from one mode to another.
		VAS3	By using data analytics technology, our port can convey cargo through diversified routes or modes in the least possible time to the receiver.
		VAS4	By using data analytics technology, our port can launch new tailored services when necessary.
	Customer Orientation	CO1	By using data analytics technology, our port quickly makes decisions regarding altering schedules, amending orders, and changing design processes to meet customer demand.
		CO2	By using data analytics technology, our port can provide individual port services to our customers.
		CO3	By using data analytics technology, our port's response time for customer complaints is faster than that of our major competitors.
		CO4	By using data analytics technology, our port has smooth operational processes for port users.

Table A2. Mean value for each cluster of the considered sub-dimension.

Macro-dimension	Sub-dimension	Measurable dimension	Best	Medium	Worst
BDAC	Organisational Learning (OL)	Number of ports	16	19	5
		OL1	4.45	3.87	2.80
		OL2	4.48	3.83	2.80
		OL3	4.55	3.87	3.00
	Technology (T)	Number of ports	6	25	9
		T1	4.62	3.81	3.13
		T2	4.62	3.97	2.97
		T3	4.65	3.96	2.67
	Basic Resources (BR)	Number of ports	13	17	10
		BR1	4.39	3.70	2.60
		BR2	4.48	3.40	2.61
	Technical Skills (TS)	Number of ports	14	21	5
		TS1	4.27	3.50	2.10
		TS2	4.30	3.47	2.10
		TS3	4.41	3.45	2.10
		TS4	4.41	3.57	2.50
	Managerial Skills (MS)	Number of ports	14	22	4
		MS1	4.18	3.48	2.00
		MS2	4.21	3.40	2.25
		MS3	4.34	3.56	2.50
		MS4	4.23	3.53	2.00
	Data (D)	Number of ports	6	25	9
		D1	4.94	4.07	2.67
		D2	4.75	3.93	3.00
		D3	4.91	4.02	3.33
	Data-driven Culture (DDC)	Number of ports	17	18	5
		DDC1	4.50	3.96	2.60
		DDC2	4.52	3.71	3.40
		DDC3	4.19	3.67	3.20
		DDC4	4.71	3.80	3.00
PSP	Environmental Dimension (ED)	Number of ports	15	21	4
		ED1	4.27	3.39	1.50
		ED2	4.13	3.46	2.25
		ED3	4.51	3.39	2.50
		ED4	4.55	3.61	2.50
	Social Dimension (SD)	Number of ports	6	25	9
		SD1	4.73	4.12	3.28
		SD2	4.92	3.99	3.00
		SD3	4.94	4.04	3.33
		ED4	4.92	4.14	3.17
	Economic Dimension (ECD)	Number of ports	15	20	5
		ECD1	4.58	3.79	2.80
		ECD2	4.68	3.69	3.00
		ECD3	4.43	3.94	2.80
		ECD4	4.56	3.85	3.20
PP	Cost (C)	Number of ports	7	23	10
		C1	4.87	3.92	3.11
		C2	4.73	3.94	3.06
		C3	4.73	3.98	3.05
	Service Quality (SQ)	Number of ports	17	15	8
		SQ1	4.35	3.72	3.00
		SQ2	4.34	3.65	3.31
		SQ3	4.34	3.71	3.06
		SQ4	4.55	3.94	3.00
	Operational Efficiency (OE)	Number of ports	20	15	5
		OE1	4.24	3.70	2.80
		OE2	4.27	3.65	3.00
		OE3	4.09	3.57	3.00
	Value-added Service (VAS)	Number of ports	9	29	2
		VAS1	4.42	3.45	2.00
		VAS2	4.16	3.42	1.50
		VAS3	4.36	3.45	2.00
		VAS4	4.50	3.59	2.50
	Customer Orientation (CO)	Number of ports	15	18	7
		CO1	4.44	3.80	3.00
		CO2	4.31	3.69	3.14
		CO3	4.15	3.83	3.00
		CO4	4.48	3.85	3.14

Table A3. Mean value for each cluster of the considered macro-dimension.

Macro-dimension	Sub-dimension	Best	Medium	Worst
BDAC	umber of ports	9	24	7
	OL	3.00	2.25	1.43
	T	2.56	1.88	1.29
	BR	3.00	2.00	1.14
	TS	2.89	2.21	1.43
	MS	3.00	2.17	1.57
	D	2.56	1.92	1.14
	DDC	2.89	2.38	1.29
PSP	Number of ports	16	15	9
	ED	2.94	2.00	1.56
	SD	2.38	2.00	1.00
	ECD	2.75	2.20	1.44
PP	Number of ports	16	18	6
	C	2.44	1.78	1.00
	SQ	2.88	1.83	1.67
	OE	3.00	2.22	1.17
	VAS	2.50	2.06	1.67
	CO	2.69	2.17	1.00