Frontier Shifts Without Ranking Changes: Evaluating Service Level Agreements on the Efficiency and Productivity of Energy Tanker Fleets Using a DEA–Malmquist Approach and Mixed Methods

Muhammad Dzaky Al Tovic[1], Eko Budi Santoso[2]
[1,2] Program Studi Magister Manajemen Teknologi, Sekolah Interdisiplin Manajemen dan Teknologi, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
Email : [1]6032232034@student.its.ac.id; [2]eko_budi@urplan.its.ac.id

Maritime logistics plays a vital role in supporting the national energy supply chain, making operational cost efficiency and fleet reliability strategic concerns for energy transportation companies in Indonesia. This study evaluates the impact of implementing a Service Level Agreement (SLA) as a performance-based governance instrument on the Small Tanker I fleet of an energy maritime logistics company by comparing operational conditions before (2023) and after (2024) implementation. An explanatory sequential mixed-methods approach was employed involving ten vessels treated as Decision Making Units (DMUs). Relative technical efficiency was measured using input-oriented BCC Data Envelopment Analysis (DEA), total productivity change was assessed through the Malmquist Productivity Index (MPI), and six hypotheses were tested using the Wilcoxon signed-rank test, linear regression analysis, and the Baron and Kenny mediation procedure. Qualitative data obtained from interviews with five key informants were used to explain quantitative findings through a joint display analysis. The results indicate that SLA implementation has a positive and significant effect on fleet operational strategy (coefficient = 3.04; $p < 0.001$) and increases total productivity by 20.64%, with technological change emerging as the dominant component (average Malmquist index = 1.2064; $p < 0.001$). At the same time, relative technical efficiency among vessels did not change significantly ($p = 0.2754$), and operational strategy was not found to mediate the relationship between SLA implementation and efficiency. Benchmark vessels were consistently identified as references for continuous improvement. These findings suggest that SLA functions as a governance instrument that shifts the operational frontier of the fleet uniformly without altering relative efficiency rankings in the short term. The study also highlights the importance of combining static efficiency measurement and dynamic productivity assessment in evaluating operational governance performance.

Keywords: Data Envelopment Analysis, fleet operational efficiency, Malmquist Productivity Index, Service Level Agreement, performance-based governance.

Corresponding Author:
Muhammad Dzaky Al Tovic
Institut Teknologi Sepuluh Nopember
Kampus ITS Sukolilo, Surabaya 60111, Jawa Timur, Indonesia
6032232034@student.its.ac.id

1. Introduction

Maritime logistics supports approximately 80% of global trade volume and serves as the backbone of economic connectivity for archipelagic nations. Consequently, disruptions at critical maritime transport nodes can rapidly propagate into supply shortages and price pressures across global markets [1], [2]. For Indonesia, this role is even more fundamental because the inter-island distribution of goods and energy depends heavily on the reliability of maritime transportation, while national logistics costs remain relatively high and the country's performance in the Logistics Performance Index has yet to demonstrate consistent improvement [3], [4]. Therefore, the operational performance of energy transportation fleets extends beyond corporate profitability, as it plays a critical role in ensuring the stability of the national fuel supply chain.

Efficiency challenges in energy tanker operations are both distinctive and increasingly urgent. Major operating cost components including fuel consumption, vessel maintenance, crew wages, and port-related expenses are highly sensitive to technical disruptions and port congestion. As a result, unplanned downtime immediately affects operating costs and service reliability [5], [6]. Maintaining schedule reliability, which is a key expectation of customers, is often difficult due to uncertain weather conditions, fluctuating port productivity, and mechanical failures. Furthermore, fragmented information sharing among stakeholders slows operational decision-making and weakens coordination across the supply chain [7]–[9]. In response to these challenges, a national energy maritime logistics company, hereafter referred to as Company X, implemented a Service Level Agreement (SLA) for its Small Tanker I fleet. The initiative was introduced to address the ineffective implementation of the Planned Maintenance System, the high frequency of engine failures, and the increasing costs associated with reactive maintenance. Preliminary operational records indicate a consistent reduction in operating costs across all vessels following SLA implementation. However, the magnitude of cost reduction varied among vessels, and the underlying mechanisms responsible for these improvements have not yet been adequately explained.

This study focuses exclusively on the Small Tanker I fleet operated by Company X, comprising ten tanker vessels with capacities ranging from 3,001 to 6,000 deadweight tons (DWT). The observation period covers two consecutive years: 2023, representing the period before SLA implementation, and 2024, representing the period after implementation. Each vessel is treated as a Decision Making Unit (DMU) within the Data Envelopment Analysis (DEA) framework. Changes in productivity between the two periods are evaluated using the Malmquist Productivity Index (MPI), while the mechanisms underlying the quantitative findings are explored through interviews with five key informants using an explanatory sequential mixed-methods design. These research boundaries ensure the homogeneity of the comparison units while enabling the utilization of comprehensive internal operational data provided by the company.

This study pursues three primary objectives. First, it evaluates differences in fleet operational performance before and after the implementation of the Service Level Agreement, considering both absolute performance indicators and relative technical efficiency. Second, it examines changes in fleet operational management strategies resulting from SLA implementation and investigates whether these strategies mediate the relationship between SLA implementation and technical efficiency. Third, it measures changes in the fleet's total productivity over time using the Malmquist Productivity Index and identifies benchmark vessels that can serve as references for continuous operational improvement.

2. Literature Review and Problem Statement

A Service Level Agreement (SLA) is a performance-based contractual mechanism that defines service standards, key performance indicators (KPIs), monitoring procedures, and incentive–penalty schemes, thereby ensuring that operational objectives are enforced through formal contractual arrangements [10], [11]. Governance literature regards SLAs as mechanisms that translate abstract service expectations into measurable performance criteria, thereby promoting accountability and continuous improvement when implemented consistently [12], [13]. From a supply chain coordination perspective, linking service performance to financial consequences has been shown to align incentives among stakeholders and enhance overall system reliability [14], [15]. Similarly, within the field of maintenance management, performance-based contracts that compensate providers according to achieved outcomes such as fleet availability have demonstrated positive effects on asset uptime across various industries [16]–[18]. The theoretical foundation of this mechanism is rooted in Agency Theory, which considers formal control systems as instruments for aligning agents' behavior with the interests of principals while reducing moral hazard, including delayed reporting of operational issues and the concealment of minor equipment failures [19], [20].

Despite these theoretical developments, the application of SLA concepts within the maritime logistics sector remains relatively limited. The shipping industry has traditionally relied on conventional contractual arrangements, such as charter party agreements, which primarily emphasize freight volume and pricing rather than explicitly defined service-level commitments [21]. Although several maritime studies have proposed a transition toward performance-based ship management contracts supported by standardized performance indicators, their primary focus has been on contract design and indicator selection rather than empirically evaluating the impact of SLA implementation on fleet efficiency before and after adoption [22], [23]. This situation reveals a significant contextual research gap. Empirical evidence regarding how SLAs influence operational performance within company-owned energy tanker fleets, supported by detailed internal operational data, remains scarce. Unlike chartered fleets, company-owned fleets provide a more controlled research setting in which the effects of SLA implementation can be isolated from confounding factors such as charter rate fluctuations and contractual adjustments driven by global fuel price volatility.

A second gap is methodological in nature. Fleet performance evaluations frequently rely on a single analytical approach, making it difficult to distinguish between a vessel's relative efficiency compared with its peers and changes in overall fleet productivity over time. Data Envelopment Analysis (DEA), as a non-parametric efficiency measurement technique, is capable of evaluating the relative efficiency of multiple Decision Making Units (DMUs) using multiple inputs and outputs without assuming a predefined production function [24]–[26]. DEA has also been extended to accommodate undesirable factors such as operational downtime [27], [28]. Nevertheless, because DEA measures efficiency relative to peer units, it becomes less sensitive when operational improvements occur uniformly across all units, resulting in little or no change in relative efficiency rankings despite substantial improvements in overall operational standards. The Malmquist Productivity Index (MPI) addresses this limitation by measuring productivity changes across different time periods and decomposing them into efficiency change and technological change, thereby capturing dynamic shifts in the production frontier [29]–[31]. The integration of DEA and MPI within a governance evaluation framework, complemented by qualitative process evidence through methodological triangulation, remains uncommon in the existing literature. However, such an integrated approach provides a more comprehensive understanding of the actual impact of governance interventions than either method alone.

Based on these contextual and methodological gaps, this study addresses the following research question: How does the implementation of a Service Level Agreement affect the relative technical efficiency and total productivity of an energy tanker fleet, and through what mechanisms are these effects realized? To answer this question, six hypotheses are proposed and tested within an integrated analytical framework. H1 posits that SLA implementation has a positive effect on the fleet's relative technical efficiency. H2 proposes that SLA implementation positively influences fleet operational management strategy. H3 states that operational management strategy has a positive effect on technical efficiency. H4 hypothesizes that operational management strategy mediates the relationship between SLA implementation and technical efficiency. H5 proposes that the fleet's total factor productivity increases significantly following SLA implementation. Finally, H6 states that technically efficient vessels can serve as benchmark units for other vessels through DEA reference-set identification and slack analysis.

3. Methodology

This study employed a comparative and explanatory quantitative approach integrated with qualitative evidence using an explanatory sequential mixed-methods design. Quantitative analysis was conducted in the first stage to identify patterns of change in fleet performance before and after the implementation of the Service Level Agreement (SLA). Subsequently, qualitative data were collected and analyzed to explain the managerial mechanisms underlying the observed quantitative findings. This research design was selected because the study not only examines relationships among variables but also seeks to understand the operational processes through which SLA implementation influences fleet performance.

The research object was the Small Tanker I fleet operated by Company X, consisting of ten tanker vessels with capacities ranging from 3,001 to 6,000 deadweight tons (DWT). The vessels were selected purposively because they possess similar operational functions, comparable vessel sizes, and complete historical operational records, allowing them to be treated as relatively homogeneous comparison units. Within the Data Envelopment Analysis (DEA) framework, each vessel was considered a Decision Making Unit (DMU). Observations covered two annual periods: 2023, representing the pre-SLA condition, and 2024, representing the post-SLA implementation period. Each observation period covered 365 operational days, resulting in a balanced panel dataset comprising 20 observations. To preserve organizational confidentiality, vessel identities were anonymized and coded as Vessel_I through Vessel_X. Confirmation from company management verified that the SLA had been fully implemented throughout 2024, thereby validating its use as the post-implementation observation period.

The DEA model incorporated three input variables and three output variables. The input variables consisted of annual operating costs, fuel consumption, and off-hire days, while the output variables included on-time performance, fleet utilization, and vessel availability. Off-hire days were treated as an input because they represent unproductive operational time; therefore, higher values indicate greater operational inefficiency, consistent with the treatment of undesirable operational factors in the DEA literature [27], [28]. In addition, the study constructed an Operational Strategy Index (Z) as a mediating variable and employed an SLA dummy variable, coded as 0 for 2023 and 1 for 2024. The Operational Strategy Index was developed from three equally weighted dimensions: Planned Maintenance System (PMS) compliance, route optimization, and crew scheduling. Equal weighting was adopted because no sufficiently robust empirical evidence was available to justify differential weighting among these dimensions. Although the sample size of ten DMUs approaches the minimum practical threshold commonly recommended for DEA applications, it remains appropriate because the analysis focuses on a single, homogeneous fleet category.

Relative technical efficiency was estimated using an input-oriented BCC DEA model under the assumption of variable returns to scale (VRS) [24], [25]. This specification was selected because, although all vessels belong to the same fleet category, they differ in age, technical condition, and route characteristics. An input orientation was adopted because the company has greater managerial control over operating costs, fuel consumption, and downtime than over cargo demand. Technical efficiency scores (θ) ranged from 0 to 1, where a value of 1 indicates that a vessel operates on the efficiency frontier. Before comparing efficiency scores across periods, data normality was assessed using the Shapiro–Wilk test. Because the efficiency scores were concentrated near unity and violated the assumption of normality, differences between periods were evaluated using the Wilcoxon signed-rank test. Furthermore, slack analysis, reference-set analysis, and lambda weights were employed to identify sources of inefficiency and determine benchmark vessels for operational improvement.

The effect of SLA implementation on operational strategy and the effect of operational strategy on technical efficiency were examined using simple linear regression analysis. The mediating role of operational strategy was evaluated following the four-step mediation procedure proposed by Baron and Kenny [32]. Changes in productivity were measured using the Malmquist Productivity Index (MPI), which decomposes total factor productivity change into efficiency change (EC), technological change (TC), and scale efficiency change

(SEC) [29], [30]. DEA efficiency scores and productivity indices were computed using MaxDEA software, whereas all statistical analyses were performed in R.

Following completion of the quantitative analysis, qualitative data were collected through semi-structured interviews with five key informants representing different managerial functions: the Fleet Manager, Technical Manager, Operations Officer, Fleet Finance Manager, and Executive Management. Interview data were analyzed using thematic coding and subsequently integrated with the quantitative findings through a joint display matrix to facilitate comprehensive interpretation. Additional evidence from operational documents, SLA guidelines, and maintenance records was used to strengthen analytical triangulation. This study acknowledges the absence of an external control group as a limitation. Consequently, methodological and data triangulation were employed to enhance the credibility and robustness of the research findings.

4. Results and Discussion

Descriptive Changes in Operational Performance Indicators

An initial overview of fleet performance changes was obtained by comparing the average values of six operational indicators between 2023 (pre-SLA implementation) and 2024 (post-SLA implementation). On the input side, all three variables declined following the implementation of the Service Level Agreement (SLA). Average operating expenditure (OPEX) decreased by 0.71%, fuel consumption declined by 3.01%, and off-hire days were reduced by 26.87%, equivalent to approximately nine fewer off-hire days per vessel per year. On the output side, all three performance indicators improved consistently, with on-time performance increasing by 4.82%, fleet utilization by 4.54%, and vessel availability by 4.56%. A summary of these descriptive statistics is presented in Table 1.

Table 1. Descriptive Statistics of Fleet Operational Indicators Before and After SLA Implementation

Variable	Mean (2023)	Mean (2024)	Change (%)	Direction
OPEX (USD)	1,162,245.52	1,153,966.50	–0.71	Decreased
Fuel consumption (tons)	3.99	3.87	–3.01	Decreased
Off-hire days (days)	33.50	24.50	–26.87	Decreased
On-time performance (%)	87.64	91.86	+4.82	Improved
Fleet utilization (%)	88.73	92.76	+4.54	Improved
Vessel availability (%)	90.30	94.42	+4.56	Improved

The pattern presented in Table 1 provides preliminary evidence that the primary impact of SLA implementation was concentrated on improvements in technical operational performance, particularly through reduced downtime and enhanced vessel availability, rather than substantial reductions in aggregate operating costs. The marked decline in off-hire days, accompanied by only a modest reduction in OPEX, suggests that improvements in technical reliability materialized more rapidly than measurable cost savings at the aggregate operational level. This interpretation is consistent with evidence obtained from the Fleet Manager, who indicated that the most noticeable improvements occurred in vessel downtime and operational readiness as ships became technically more reliable. In contrast, changes in overall operating expenditure remained influenced by external factors including spare-part prices, exchange rate fluctuations, and scheduled dry-docking activities which cannot be fully attributed to SLA implementation alone.

These descriptive findings demonstrate clear improvements in the fleet's absolute operational performance following SLA implementation. However, descriptive comparisons alone do not determine whether these operational improvements were sufficient to alter the relative technical efficiency of individual vessels within the fleet. Consequently, a formal efficiency analysis using Data Envelopment Analysis (DEA) is required to

ISSN 2828-9463

examine whether the observed operational gains translated into changes in relative efficiency rankings among vessels.

Relative Technical Efficiency and Hypothesis Testing (H1)

The results of the input-oriented BCC Data Envelopment Analysis (DEA) indicate that the fleet's average technical efficiency improved only marginally, increasing from 0.9965 in 2023 to 0.9987 in 2024. A notable characteristic of both observation periods is the concentration of efficiency scores very close to the efficiency frontier. In 2023, six out of ten vessels were fully efficient with a technical efficiency score (θ) equal to 1.0000, while this number increased to eight vessels following the implementation of the Service Level Agreement (SLA) in 2024. The efficiency scores for each vessel are presented in Table 2.

Table 2. DEA Technical Efficiency Scores by Vessel in 2023 and 2024

Vessel	θ (2023)	θ (2024)	Change	Interpretation
Vessel_I	0.9783	1.0000	+0.0217	Improved
Vessel_II	0.9984	0.9929	–0.0055	Declined
Vessel_III	0.9898	0.9996	+0.0098	Improved
Vessel_IV	1.0000	0.9946	–0.0054	Declined
Vessel_V	1.0000	1.0000	0.0000	Unchanged
Vessel_VI	1.0000	1.0000	0.0000	Unchanged
Vessel_VII	0.9980	1.0000	+0.0020	Improved
Vessel_VIII	1.0000	1.0000	0.0000	Unchanged
Vessel_IX	1.0000	1.0000	0.0000	Unchanged
Vessel_X	1.0000	1.0000	0.0000	Unchanged
Average	0.9965	0.9987	+0.0022	Marginal improvement

The Shapiro–Wilk normality test rejected the assumption of normality for both observation periods, yielding extremely small *p*-values. This result is expected because most technical efficiency scores were concentrated at or near the maximum value of one, producing a highly skewed distribution. Consequently, differences between the two periods were evaluated using the Wilcoxon signed-rank test, which produced a p-value of 0.2754, exceeding the 0.05 significance threshold. Therefore, no statistically significant difference was found in the relative technical efficiency scores between 2023 and 2024. Accordingly, Hypothesis 1 (H1), which proposed that SLA implementation would significantly improve relative technical efficiency, was not supported.

The absence of a statistically significant improvement should not be interpreted as evidence that the SLA failed to enhance operational performance. Instead, it reflects the inherent characteristics of DEA as a relative efficiency measurement method. DEA evaluates each vessel relative to an efficiency frontier constructed from the best-performing units within the sample. Consequently, when operational improvements occur relatively uniformly across most vessels, the efficiency frontier itself shifts upward, while the relative positions of individual vessels remain largely unchanged. This phenomenon accurately describes the operational changes observed in the Small Tanker I fleet. Although nearly all vessels experienced improvements in absolute operational indicators including lower downtime, higher availability, and better service performance the improvements were insufficiently heterogeneous to alter their relative efficiency rankings.

This interpretation is further supported by three additional considerations. First, the concentration of technical efficiency scores near the upper boundary of one left only limited room for further relative improvement. Second, the number of Decision Making Units (DMUs) approached the lower practical limit commonly recommended for DEA applications, thereby reducing the discriminatory power of the efficiency frontier. Third, structural differences among vessels including age, technical condition, and operational deployment continued to influence relative performance despite overall operational improvements. According to the Fleet Manager, vessels that remained relatively less efficient generally faced more challenging initial conditions, such as advanced vessel age, recurring mechanical failures, or deployment on routes characterized by prolonged port waiting times. As a result, although SLA implementation improved internal operational processes, these structural constraints prevented certain vessels from achieving the efficiency frontier within the relatively short post-implementation observation period.

Slack Analysis, Benchmark Identification, and Hypothesis Testing (H6)

Slack analysis provides a more operationally meaningful diagnosis of inefficiency than aggregate DEA efficiency scores alone. In 2023, the primary sources of inefficiency were concentrated in operating expenditure (OPEX). The largest OPEX slack was observed for Vessel_I, amounting to approximately USD 37,787, followed by Vessel_II with approximately USD 21,032, and Vessel_III with approximately USD 5,639. Following the implementation of the Service Level Agreement (SLA) in 2024, the structure of inefficiency changed considerably. OPEX slack declined substantially, with Vessel_I no longer exhibiting excess operating costs and the OPEX slack of Vessel_II decreasing to approximately USD 9,013. At the same time, the remaining sources of inefficiency shifted toward fuel consumption and off-hire days, particularly for Vessel_III and Vessel_IV. Concurrently, the number of vessels operating on the DEA efficiency frontier increased from six to eight.

These changes in the slack structure provide substantive evidence that the initial priority of SLA implementation strengthening operational cost control produced measurable improvements. As cost inefficiencies were progressively reduced, subsequent improvement priorities shifted toward optimizing fuel efficiency and minimizing operational downtime. This interpretation is supported by information obtained from the Fleet Finance Manager, who explained that a portion of expenditures previously incurred as emergency repair costs had been converted into planned maintenance expenditures. Consequently, although total OPEX declined only modestly, the composition of operating costs changed substantially, reflecting more proactive maintenance practices and improved asset management.

The reference-set analysis and lambda weights further identified the vessels serving as performance benchmarks for relatively inefficient units. Vessel_VIII consistently emerged as the principal benchmark in both observation periods, serving as the reference vessel for five vessels in 2023 and four vessels in 2024, while receiving the largest lambda weights. Vessel_VI and Vessel_X also functioned as secondary benchmark vessels. The persistence of Vessel_VIII as the dominant benchmark across both years suggests that its superior performance reflects stable and consistently effective operational practices rather than temporary operational advantages.

Based on the acceptance criterion that at least one technically efficient vessel should serve as a benchmark for at least two other vessels, Hypothesis 6 (H6) is supported. Qualitative evidence obtained from the Technical Manager further enriches this finding by indicating that benchmark performance results from the combined effects of sound technical condition, disciplined crew performance, and effective supervision by fleet superintendents. Interestingly, the benchmark vessel was not necessarily the newest vessel in the fleet. Instead, its sustained reliability was attributed to systematic defect management, well-organized maintenance planning, and rapid corrective actions. These findings suggest that benchmarking programs within the fleet should prioritize the transfer of managerial best practices, maintenance discipline, and operational routines rather than attempting to standardize physical vessel conditions through capital-intensive investments.

ISSN 2828-9463

Changes in Operational Strategy and Hypothesis Testing (H2)

The Operational Strategy Index (Z) was developed as a composite indicator consisting of three equally weighted dimensions: Planned Maintenance System (PMS) compliance, route optimization, and crew scheduling. Equal weighting was adopted because qualitative triangulation indicated that these three dimensions were considered equally important by company management. Descriptive analysis shows that the average Operational Strategy Index increased from 75.92 in 2023 to 78.96 in 2024, representing an average improvement of approximately 3.04 points. Improvements were relatively uniform across all vessels, ranging from 2.9 to 3.6 points. This homogeneous pattern suggests that operational strategy improvements occurred as a system-wide organizational transformation rather than as isolated improvements among individual vessels. This interpretation is further supported by the Malmquist Productivity Index (MPI) results, which demonstrate productivity improvements across every vessel in the fleet.

Hypothesis 2 (H2) was evaluated using simple linear regression, with the SLA dummy variable serving as the independent variable. The regression produced a coefficient of 3.04, a p-value of 4.23×10^{-15}, and an R^2 of 0.9696. The estimated coefficient indicates that the implementation of the SLA increased the Operational Strategy Index by an average of 3.04 points compared with the pre-implementation period. The exceptionally high coefficient of determination is reasonable because the Operational Strategy Index was specifically designed to capture changes in maintenance management, route planning, and crew scheduling directly associated with SLA implementation. Therefore, Hypothesis 2 is strongly supported.

Table 3. Regression Results for Hypotheses H2 and H3

Model	Variable	Coefficient	*p*-value	R^2	Decision
H2: Operational Strategy (Z) on SLA	SLA (dummy)	3.04	4.23×10^{-15}	0.9696	Supported
H3: Technical Efficiency (θ) on Operational Strategy	Operational Strategy Index (Z)	0.0009	0.2553	0.0713	Not Supported

The findings related to H2 are best interpreted through the perspectives of performance-based governance and Agency Theory. The SLA functioned as a governance instrument that translated broad strategic objectives such as improving fleet reliability into measurable operational obligations supported by clearly defined performance indicators, structured escalation procedures, and explicit accountability mechanisms. According to the Fleet Manager, this transformation represented a fundamental shift in managerial thinking. Decisions that had previously been made reactively after operational problems emerged increasingly became proactive, supported by earlier performance monitoring, more structured review meetings, and clearer follow-up procedures.

From the perspective of Agency Theory, these findings are consistent with the proposition that formal control mechanisms reduce the tendency of agents to pursue actions that diverge from the interests of principals [19]. One common manifestation of such agency problems is moral hazard, including delayed reporting of operational issues and the concealment of minor technical failures [20]. The SLA's requirements for documentation, photographic evidence, timestamp verification, and formal review procedures reduced information asymmetry by making operational deviations more transparent and traceable. Interview participants acknowledged that while delayed reporting had not been eliminated entirely, it had become substantially less frequent and considerably easier to detect. Consequently, the SLA can be viewed as an effective mechanism for reducing information asymmetry, although its long-term effectiveness remains dependent upon consistent enforcement and managerial commitment.

A methodological consideration should accompany the interpretation of these findings. Because the SLA dummy variable also represents the distinction between two observation periods, the H2 regression should be interpreted as indicating a strong pre–post association rather than definitive causal proof. Other temporal factors including personnel rotation, fluctuations in fuel prices, and changes in port operating conditions may also have influenced the Operational Strategy Index during the study period. Nevertheless, according to Executive Management, although external conditions undoubtedly contributed to operational outcomes, no single external policy or environmental factor sufficiently explains the improvements observed in 2024. Accordingly, the SLA may reasonably be regarded as one of the principal internal drivers of operational improvement without claiming that it represents the sole causal explanation.

The Limited Effect of Operational Strategy on Technical Efficiency and Hypothesis Testing (H3 and H4)

Hypothesis 3 (H3) was evaluated by regressing the DEA technical efficiency score (θ) on the Operational Strategy Index (Z). The regression yielded a coefficient of 0.0009, a p-value of 0.2553, and an R^2 of 0.0713, indicating that the relationship between operational strategy and relative technical efficiency was not statistically significant. Furthermore, the Pearson correlation coefficient between the two variables was 0.267, suggesting only a weak positive association.

At first glance, these findings appear counterintuitive because operational strategy improved substantially following the implementation of the Service Level Agreement (SLA), while fleet productivity also increased, yet no significant direct relationship emerged between operational strategy and relative technical efficiency. Several factors help explain this result. First, the Operational Strategy Index measures improvements in managerial and operational processes, whereas the DEA efficiency score (θ) measures relative technical efficiency at the output level. Improvements in operational processes do not necessarily translate immediately into measurable improvements in technical efficiency, particularly within the same observation period, because operational interventions often require time before their full effects become evident.

Second, because a large proportion of vessels were already operating on or very close to the DEA efficiency frontier, the sensitivity of technical efficiency scores to further improvements became inherently limited. Once a vessel reaches an efficiency score of 1.0000, additional operational improvements cannot be reflected through higher DEA scores, thereby creating a ceiling effect that reduces the discriminatory capability of the efficiency measure. According to the Technical Manager, technical efficiency is also strongly influenced by structural characteristics of individual vessels, including vessel age, spare-part procurement lead times, and limited maintenance windows. Consequently, the performance targets introduced through the SLA cannot immediately eliminate structural differences between older and newer vessels without substantial capital investment. Similarly, the Fleet Finance Manager explained that financial benefits from preventive maintenance generally become visible only several months after implementation, once preventive maintenance activities begin reducing equipment failures. During the initial implementation phase, operating costs may remain relatively high because accumulated maintenance backlogs must first be addressed.

The mediating role proposed in Hypothesis 4 (H4) was evaluated using the four-step mediation procedure of Baron and Kenny [32]. The analysis revealed that only the second condition the significant effect of SLA implementation on the Operational Strategy Index was satisfied. The first condition, requiring a significant direct effect of SLA implementation on DEA technical efficiency, was not met, consistent with the findings for H1. Likewise, the third condition, requiring a significant relationship between operational strategy and technical efficiency, was also not satisfied, consistent with the results of H3. Consequently, the fourth condition, demonstrating a valid mediation pattern, could not be established. Therefore, Hypothesis 4 is not supported.

These findings suggest that the proposed mediation mechanism cannot be statistically confirmed within the context of relative DEA technical efficiency because the Baron and Kenny framework requires a significant direct relationship between the independent variable and the dependent variable, which was absent in this study. Nevertheless, qualitative evidence indicates that the underlying operational mechanism remains plausible. SLA implementation clearly improved operational management practices, and interview findings consistently suggest that these strategic improvements contributed to better operational performance. Accordingly, a more appropriate interpretation is that the effects of improved operational strategy are reflected more clearly in dynamic productivity growth than in static measures of relative technical efficiency. This interpretation provides a methodological rationale for future research to examine mediation effects using the Malmquist Productivity Index (MPI) as the outcome variable and employing multi-year panel data, thereby increasing statistical power and allowing sufficient time for process improvements to be translated into measurable efficiency gains.

Dynamic Productivity Analysis and Hypothesis Testing (H5)

The Malmquist Productivity Index (MPI) provides a perspective that complements Data Envelopment Analysis (DEA) by measuring changes in fleet productivity between 2023 and 2024, rather than evaluating the relative efficiency of vessels at a single point in time. The results indicate that all ten vessels experienced productivity improvements, with MPI values ranging from 1.1048 to 1.3764. The fleet achieved an average MPI of 1.2064, indicating an overall increase in total factor productivity of 20.64%, and no vessel exhibited a decline in productivity during the observation period. The vessel-level results and the decomposition of the Malmquist index are presented in Table 4.

Table 4. Malmquist Productivity Index Results by Vessel

Vessel	Efficiency Change	Technological Change	Scale Efficiency Change	MPI
Vessel_I	1.0222	1.3454	1.2008	1.3764
Vessel_II	0.9945	1.2066	1.2004	1.2004
Vessel_III	1.1999	1.2718	0.9965	1.2799
Vessel_IV	0.9946	1.1503	1.1934	1.1481
Vessel_V	1.0000	1.1356	1.0000	1.1356
Vessel_VI	1.0000	1.3092	1.0000	1.3092
Vessel_VII	1.0020	1.1535	1.0010	1.1569
Vessel_VIII	1.0000	1.2320	1.0000	1.2320
Vessel_IX	1.0000	1.1486	1.0000	1.1486
Vessel_X	1.0000	1.1048	1.0000	1.1048
Average	1.0213	1.2058	1.0592	1.2064

The Shapiro–Wilk normality test produced a p-value of 0.3209, indicating that the distribution of MPI values did not significantly deviate from normality. Consequently, a one-sample *t*-test was conducted using the benchmark value of 1.0000, representing no productivity change. The analysis yielded a t-statistic of 7.50 with a p-value of 3.69×10^{-5}, demonstrating that the average MPI was significantly greater than one. Therefore, fleet productivity increased significantly following SLA implementation, providing strong support for Hypothesis 5 (H5).

The decomposition of the Malmquist index provides further insight into the sources of productivity improvement. The average Technological Change (TC) component reached 1.2058, substantially exceeding the average Efficiency Change (EC) component of 1.0213. This pattern indicates that productivity growth was driven primarily by a system-wide shift of the production frontier, whereby nearly all vessels improved simultaneously, rather than by substantial improvements in the relative efficiency rankings of individual vessels. In other words, the fleet achieved broad operational advancement through collective improvements rather than through isolated gains by a small number of high-performing vessels.

The interpretation of the Technological Change component requires careful consideration to avoid overstating its meaning. Within the DEA literature, technological change is frequently associated with the adoption of new physical technologies. However, such an interpretation would be inappropriate in the present study because no major technological upgrades or advanced predictive maintenance systems were introduced during the observation period. According to the Technical Manager, the fleet did not implement full-scale predictive maintenance supported by real-time sensor technology throughout 2024. Instead, operational improvements were achieved primarily through strengthened process management, including more active utilization of the Computerized Maintenance Management System (CMMS), standardized digital reporting templates, faster compilation of operational performance indicators, and more disciplined monitoring of fuel consumption and engine performance.

Accordingly, the observed 20.58% technological frontier shift within the Small Tanker I fleet should be interpreted as reflecting process digitalization, organizational learning, and improved data governance, rather than investment in new physical technologies. This interpretation is consistent with the broader understanding of the Malmquist Productivity Index, which conceptualizes technological change as encompassing not only physical innovations but also improvements in organizational practices, operational processes, and information systems [30], [31]. Consequently, the observed frontier shift is more appropriately viewed as an improvement in operational governance standards resulting from more effective management systems and greater procedural discipline.

This interpretation is also consistent with the maritime digitalization literature, which emphasizes that information system integration, standardized reporting, and data governance frequently contribute more to operational performance improvements during the early stages of digital transformation than investments in advanced hardware or physical technologies [21]. Within the Small Tanker I fleet, the strengthened implementation of the CMMS and the standardization of reporting procedures functioned as a form of organizational digital infrastructure, enabling a significant upward shift in the production frontier without requiring substantial capital expenditure on new technologies. From a managerial perspective, this finding suggests that sustaining the observed productivity gains depends largely on maintaining consistent data governance, disciplined reporting practices, and organizational commitment. Any deterioration in these governance mechanisms could gradually erode the operational frontier, even if the underlying technological infrastructure remains available.

Frontier Shift Without Changes in Relative Rankings

The joint display integrating the quantitative and qualitative findings reveals three distinct integration patterns. For Hypotheses H2, H5, and H6, the numerical evidence and qualitative narratives demonstrate strong convergence, with interview findings reinforcing the statistical results. In contrast, for Hypotheses H1, H3, and H4, an apparent discrepancy exists between the operational improvements perceived by the interview participants and the statistical evidence obtained from the quantitative analyses. This discrepancy originates from the different dimensions measured by the two analytical approaches static relative efficiency versus dynamic productivity rather than from conflicting empirical evidence. Consequently, the findings remain methodologically consistent despite leading to different statistical conclusions.

The Malmquist Productivity Index (MPI) demonstrates a significant improvement in fleet productivity, whereas the Data Envelopment Analysis (DEA) results indicate no statistically significant change in relative technical efficiency. These findings are not contradictory because the two methods evaluate different aspects of operational performance. DEA assesses the relative efficiency position of each vessel compared with other vessels within the same observation period, while MPI measures changes in productivity over time. Therefore, although the relative ranking of vessels changed very little following SLA implementation, the overall productivity of the fleet improved substantially between 2023 and 2024.

The central contribution of this study lies in demonstrating that the Service Level Agreement (SLA) functions primarily as an instrument of system-wide operational transformation rather than as a mechanism that immediately reshapes relative efficiency rankings. Most vessels had already achieved very high DEA efficiency scores in 2023, leaving only limited scope for further improvements in relative technical efficiency and making statistically significant changes in DEA scores inherently difficult to detect. At the same time, SLA implementation substantially strengthened operational discipline, compliance with the Planned Maintenance System (PMS), monitoring effectiveness, and the overall consistency of fleet management practices. These organizational improvements are clearly reflected in the significant increases observed in the Malmquist Productivity Index.

Accordingly, the findings indicate that the operational production frontier shifted upward across the fleet as a whole, while the relative positions of individual vessels remained largely unchanged. This interpretation provides a coherent explanation for the complete set of empirical results. Specifically, the average DEA technical efficiency score (θ) increased only marginally and the Wilcoxon signed-rank test detected no statistically significant difference between the pre- and post-implementation periods, whereas the MPI analysis revealed a substantial and statistically significant increase in total factor productivity. These outcomes are complementary rather than contradictory because they describe different dimensions of organizational performance.

This interpretation carries important implications for the evaluation of performance-based governance systems. If organizational performance were assessed solely through DEA, evaluators might incorrectly conclude that SLA implementation had little or no meaningful impact, because DEA is inherently less sensitive to improvements that occur uniformly across all decision-making units. The inclusion of the Malmquist Productivity Index as a complementary analytical tool enables the evaluation framework to capture temporal productivity improvements that remain invisible under static efficiency analysis. Furthermore, qualitative triangulation provides valuable insight into the organizational mechanisms responsible for these quantitative outcomes. Together, the integration of DEA, MPI, and qualitative evidence offers a more comprehensive and balanced evaluation of SLA implementation, demonstrating that the agreement effectively strengthened operational strategy and increased total fleet productivity, while improvements in relative technical efficiency rankings may require a longer observation period before becoming statistically observable.

Theoretical and Managerial Implications

From a theoretical perspective, this study contributes to three major streams of literature. First, regarding Agency Theory, the findings from Hypothesis 2 (H2) and the qualitative evidence concerning the reduction of moral hazard confirm that the Service Level Agreement (SLA) functions as a formal control mechanism that aligns the behavior of agents with the interests of principals. However, the findings also suggest a more nuanced interpretation: rather than eliminating moral hazard entirely, SLA implementation reduces its occurrence and makes deviations more transparent and easier to detect through structured monitoring and documentation.

Second, this study contributes to the literature on performance-based governance by demonstrating that the effects of governance instruments are more readily observed through system-wide operational improvements, reflected in an upward shift of the production frontier, than through changes in the relative efficiency rankings of individual operational units. Consequently, governance evaluations should incorporate analytical approaches capable of capturing systemic improvements rather than relying exclusively on measures of relative efficiency.

Third, this study contributes to the efficiency measurement literature by demonstrating the complementary roles of Data Envelopment Analysis (DEA) and the Malmquist Productivity Index (MPI). When operational improvements occur uniformly across multiple decision-making units, DEA may fail to identify statistically significant changes because relative efficiency rankings remain stable. In contrast, MPI successfully captures substantial shifts in the production frontier over time. Furthermore, the decomposition of the Malmquist index enables researchers to distinguish whether productivity improvements originate primarily from changes in relative efficiency or from broader technological and organizational transformations affecting the entire system.

From a managerial perspective, the findings suggest several priorities for future operational improvement. First, the definition of performance indicators should be refined. In particular, fuel consumption should be normalized according to route characteristics because fuel usage is influenced by cargo load, weather conditions, sailing speed, and hull condition. Similarly, vessel availability should be separated into technical availability and commercial availability to provide a more accurate and equitable assessment of operational performance.

Second, strengthening fuel governance should become the next operational priority. The 2024 slack analysis indicates that the primary sources of inefficiency have shifted from operating costs toward fuel consumption and operational downtime. Accordingly, voyage-based fuel monitoring, disciplined speed management, and standardized reporting procedures should receive greater managerial attention.
Third, the Planned Maintenance System (PMS) and the Computerized Maintenance Management System (CMMS) should be further strengthened through automatic integration with the SLA performance dashboard and automated escalation mechanisms for overdue maintenance tasks. Such integration would reduce dependence on the initiative of individual fleet superintendents and enhance the consistency of operational discipline across the organization.

Fourth, the company's benchmark vessel program should be formalized by systematically documenting best operational practices and establishing cross-vessel mentoring initiatives. Because the superior performance of benchmark vessels primarily reflects transferable managerial discipline rather than superior physical assets, substantial performance improvements can be achieved without requiring significant capital investment.

Fifth, the balance between performance incentives and sanctions should be improved. Qualitative triangulation indicates that the current governance system places greater emphasis on sanctions than on rewards. Such an imbalance may encourage defensive organizational behavior over the long term, potentially reducing employee motivation and organizational learning.

Finally, standardized operational data recording across observation periods should become a managerial priority. The higher quality of operational records observed in 2024 compared with 2023 may partially influence the perceived magnitude of operational improvements. Maintaining consistent documentation standards is therefore essential to ensure that future performance evaluations accurately reflect genuine operational changes rather than improvements in data quality alone.

Research Limitations

Several limitations should be acknowledged to ensure an appropriate interpretation of the findings. First, the study included only ten Decision Making Units (DMUs), which falls below the commonly recommended practical threshold for DEA applications involving three input variables and three output variables.

ISSN 2828-9463

Consequently, the discriminatory power of the efficiency frontier was limited, resulting in a large proportion of vessels receiving perfect efficiency scores.

Second, the strong concentration of DEA efficiency scores near the value of one produced a highly non-normal distribution and left relatively little room for additional improvements in relative efficiency. This limitation is inherent in studies involving relatively homogeneous fleets in which many operational units already perform close to the efficiency frontier.

Third, the two-year observation period was sufficient to capture the initial effects of SLA implementation but remains inadequate for evaluating medium- and long-term organizational impacts. In particular, the mediation mechanism proposed in this study may require a substantially longer period before its effects become observable through quantitative efficiency measures.

Fourth, the Operational Strategy Index was constructed using equal weighting across its three constituent dimensions. Alternative weighting approaches based on expert judgment or statistical weighting techniques could potentially produce different results. In addition, some interview responses addressed organizational issues at a general managerial level, thereby limiting qualitative triangulation at the individual vessel level.

Fifth, several important external factors including weather conditions, port congestion, dry-docking schedules, and fluctuations in fuel prices were not explicitly incorporated into the analytical models. Furthermore, because the DEA model employed in this study is deterministic, efficiency estimates were not accompanied by bootstrap confidence intervals, limiting statistical inference regarding the precision of efficiency scores.

Finally, customer satisfaction was not available as a consistently measured quantitative variable across the two observation periods. Consequently, this important dimension of service performance could only be incorporated qualitatively rather than as an output variable within the DEA framework.

5. Conclusion

This study evaluated the impact of Service Level Agreement (SLA) implementation on the Small Tanker I fleet of a national energy maritime logistics company by integrating static efficiency analysis, dynamic productivity measurement, and qualitative process evidence within an explanatory sequential mixed-methods framework.

Regarding the first research objective, the findings indicate that operational performance improved across all absolute performance indicators following SLA implementation. The most substantial improvement occurred in the reduction of off-hire days, which declined by 26.87%. Nevertheless, these operational improvements did not result in a statistically significant change in relative technical efficiency, as indicated by the Wilcoxon signed-rank test ($p = 0.2754$).

With respect to the second research objective, SLA implementation produced consistent improvements in operational performance indicators and significantly increased dynamic total factor productivity by 20.64%, while its direct effect on relative technical efficiency remained statistically insignificant. Furthermore, SLA implementation successfully transformed fleet management practices by shifting maintenance activities from a reactive toward a preventive approach, strengthening discipline in route planning and fuel management, and improving cross-functional coordination. These organizational improvements were reflected in a significant increase of 3.04 points in the Operational Strategy Index.

The evaluation of the six research hypotheses revealed a coherent pattern of findings. Hypothesis 2 (H2), proposing that SLA implementation positively influences operational strategy, was strongly supported.

Hypothesis 5 (H5), predicting significant improvements in total factor productivity, was also supported. Likewise, Hypothesis 6 (H6) was supported through the successful identification of benchmark vessels capable of guiding continuous operational improvement. Conversely, Hypothesis 1 (H1) was not supported because SLA implementation did not significantly alter the relative technical efficiency of individual vessels. Hypothesis 3 (H3) was also unsupported, indicating that improvements in operational strategy did not translate directly into statistically significant improvements in relative DEA efficiency. Similarly, Hypothesis 4 (H4) was not supported because the proposed mediating role of operational strategy could not be statistically confirmed within the context of relative technical efficiency. These findings primarily reflect differences between static and dynamic performance measurement frameworks, together with the limitations associated with the number of DMUs and the relatively short observation period. Consequently, they should not be interpreted as evidence of policy failure.

Instead, the findings suggest that the Service Level Agreement should be understood as a governance instrument that systematically shifts the operational production frontier upward across the entire fleet, enabling all vessels to improve simultaneously while leaving their relative efficiency rankings largely unchanged in the short term.

This study makes three principal contributions. Empirically, it extends the application of Service Level Agreements beyond the field of information technology services into the operational management of maritime energy logistics using detailed internal company data. Methodologically, it demonstrates the importance of integrating Data Envelopment Analysis (DEA) and the Malmquist Productivity Index (MPI) as complementary analytical tools, thereby preventing governance evaluations from underestimating policy effectiveness when improvements occur uniformly across operational units. Practically, the study proposes a comprehensive SLA evaluation framework that can be replicated by other maritime logistics organizations. The framework encompasses the systematic collection of operational input and output data, DEA efficiency analysis, Malmquist productivity decomposition, slack analysis for identifying operational improvement priorities, benchmark vessel identification, and integration with qualitative evidence to provide a comprehensive understanding of organizational performance.

Future research is recommended to expand the number of Decision Making Units (DMUs) by incorporating additional fleet categories, extend the observation period to three years or longer, employ bootstrapped DEA to generate confidence intervals for efficiency estimates, investigate mediation mechanisms using the Malmquist Productivity Index within multi-year panel datasets, and integrate standardized customer satisfaction surveys so that customer service quality can be incorporated as an output variable in future efficiency evaluation models.

6. Referensi

[1] T. Denamiel and E. Brown, "The state of maritime supply-chain threats," Center for Strategic and International Studies, Washington, DC, USA, 2024. [Online]. Available: https://www.csis.org/analysis/state-maritime-supply-chain-threats

[2] L. F. Pratson, "Assessing impacts to maritime shipping from marine chokepoint closures," Communications in Transportation Research, vol. 3, p. 100083, 2023, doi: 10.1016/j.commtr.2022.100083.

[3] World Bank, "Improving Indonesia's freight logistics system: A plan of action," World Bank Group, Washington, DC, USA, 2015. [Online]. Available: https://documents1.worldbank.org/curated/en/617051540584814484/pdf/WP-PUBLIC-2015-WBG-Improving-Freight-Logistics.pdf

[4] Bisnis Indonesia, "Logistics Performance Index (LPI) Indonesia berada di peringkat ke-63 dari 139

negara," Bisnis Indonesia, Feb. 13, 2024. [Online]. Available: https://foto.bisnis.com/view/20240213/1740489/logistics-performance-index-lpi-indonesia-berada-di-peringkat-ke-63-dari-139-negara
[5] N. D. Belesis, C. G. Kampouris, and A. E. Fousteris, "Vessels operating expenses: A key variable on companies strategic decisions," Corporate and Business Strategy Review, vol. 5, no. 2, pp. 102-117, 2024, doi: 10.22495/cbsrv5i2art9.
[6] D.-W. Song and P. M. Panayides, Maritime Logistics: A Guide to Contemporary Shipping and Port Management. London, U.K.: Kogan Page, 2017.
[7] S. Senses and M. Kumral, "An optimization-based approach to fleet reliability and allocation in open-pit mining," Decision Analytics Journal, vol. 15, p. 100583, 2025, doi: 10.1016/j.dajour.2025.100583.
[8] S. Nikghadam, K. F. Molkenboer, L. Tavasszy, and J. Rezaei, "Information sharing to mitigate delays in port: The case of the Port of Rotterdam," Maritime Economics and Logistics, vol. 25, pp. 576-601, 2021, doi: 10.1057/s41278-021-00203-9.
[9] B. Al-Kurdi, M. Alshurideh, I. A. Akour, and A. AlHamad, "The impact of information technology competencies and fleet management practices on effective service delivery in the construction industry," International Journal of Business Analytics and Security, vol. 3, no. 2, pp. 118-129, 2023, doi: 10.54489/ijbas.v3i2.262.
[10] C. C. Marquezan, A. Metzger, R. Franklin, and K. Pohl, "Runtime management of multi-level SLAs for transport and logistics services," in Service-Oriented Computing (ICSOC 2014), Lecture Notes in Computer Science, vol. 8831. Berlin, Germany: Springer, 2014, pp. 560-574, doi: 10.1007/978-3-662-45391-9_49.
[11] J. J. M. Trienekens, J. J. Bouman, and M. van der Zwan, "Specification of service level agreements: Problems, principles and practices," Software Quality Journal, vol. 12, no. 1, pp. 43-57, 2004, doi: 10.1023/B:SQJO.0000013358.61395.96.
[12] J. Goo and C. D. Huang, "Facilitating relational governance through service level agreements in IT outsourcing: An application of the commitment-trust theory," Decision Support Systems, vol. 46, no. 1, pp. 216-232, 2008, doi: 10.1016/j.dss.2008.06.005.
[13] A. K. Swain and V. R. Garza, "Key factors in achieving service level agreements (SLA) for information technology (IT) incident resolution," Information Systems Frontiers, vol. 25, pp. 819-834, 2022, doi: 10.1007/s10796-022-10266-5.
[14] M. A. Sieke, R. W. Seifert, and U. W. Thonemann, "Designing service level contracts for supply chain coordination," Production and Operations Management, vol. 21, no. 4, pp. 698-714, 2012, doi: 10.1111/j.1937-5956.2011.01301.x.
[15] Z. Yin and S. Ma, "Incentives to improve the service level in a random yield supply chain: The role of bonus contracts," European Journal of Operational Research, vol. 244, no. 3, pp. 778-791, 2015, doi: 10.1016/j.ejor.2015.02.006.
[16] K. Selviaridis and F. Wynstra, "Performance-based contracting: A literature review and future research directions," International Journal of Production Research, vol. 53, no. 12, pp. 3505-3540, 2015, doi: 10.1080/00207543.2014.978031.
[17] J. A. Guajardo, M. A. Cohen, S.-H. Kim, and S. Netessine, "Impact of performance-based contracting on product reliability: An empirical analysis," Management Science, vol. 58, no. 5, pp. 961-979, 2012, doi: 10.1287/mnsc.1110.1465.
[18] H. Husniah, N. A. Andriana, and B. P. Iskandar, "Game theoretic models in fleet performance-based maintenance contracts," International Journal of Business and Globalisation, vol. 26, no. 1/2, pp. 41-55, 2020, doi: 10.1504/ijbg.2020.109830.
[19] R. Beal Partyka, "Supply chain management: An integrative review from the agency theory

perspective," Revista de Gestao, vol. 29, no. 2, pp. 175-198, 2021, doi: 10.1108/rege-04-2021-0058.

[20] Y. Lee, S. H. Song, and T. Cheong, "The value of supply chain coordination under moral hazard: A case study of the consumer product supply chain," PLOS ONE, vol. 13, no. 3, p. e0194043, 2018, doi: 10.1371/journal.pone.0194043.

[21] F. Zeng, A. Chen, S. Xu, H. K. Chan, and Y. Li, "Digitalization in the maritime logistics industry: A systematic literature review of enablers and barriers," Journal of Marine Science and Engineering, vol. 13, no. 4, p. 797, 2025, doi: 10.3390/jmse13040797.

[22] A. Rialland, D. A. Nesheim, J. A. Norbeck, and O. J. Rodseth, "Performance-based ship management contracts using the Shipping KPI standard," WMU Journal of Maritime Affairs, vol. 13, no. 2, pp. 191-206, 2014, doi: 10.1007/s13437-014-0058-9.

[23] K. Konsta and E. Plomaritou, "Key performance indicators (KPIs) and shipping companies performance evaluation: The case of Greek tanker shipping companies," International Journal of Business and Management, vol. 7, no. 10, pp. 142-155, 2012, doi: 10.5539/ijbm.v7n10p142.

[24] A. Charnes, W. W. Cooper, and E. Rhodes, "Measuring the efficiency of decision making units," European Journal of Operational Research, vol. 2, no. 6, pp. 429-444, 1978, doi: 10.1016/0377-2217(78)90138-8.

[25] R. D. Banker, A. Charnes, and W. W. Cooper, "Some models for estimating technical and scale inefficiencies in data envelopment analysis," Management Science, vol. 30, no. 9, pp. 1078-1092, 1984, doi: 10.1287/mnsc.30.9.1078.

[26] A. Emrouznejad and G.-L. Yang, "A survey and analysis of the first 40 years of scholarly literature in DEA: 1978-2016," Socio-Economic Planning Sciences, vol. 61, pp. 4-8, 2018, doi: 10.1016/j.seps.2017.01.008.

[27] H. Scheel, "Undesirable outputs in efficiency valuations," European Journal of Operational Research, vol. 132, no. 2, pp. 400-410, 2001, doi: 10.1016/S0377-2217(00)00160-0.

[28] L. M. Seiford and J. Zhu, "Modeling undesirable factors in efficiency evaluation," European Journal of Operational Research, vol. 142, no. 1, pp. 16-20, 2002, doi: 10.1016/S0377-2217(01)00293-4.

[29] R. Färe, S. Grosskopf, M. Norris, and Z. Zhang, "Productivity growth, technical progress, and efficiency change in industrialized countries," The American Economic Review, vol. 84, no. 1, pp. 66-83, 1994. [Online]. Available: https://www.jstor.org/stable/2117971

[30] T. J. Coelli, D. S. P. Rao, C. J. O'Donnell, and G. E. Battese, An Introduction to Efficiency and Productivity Analysis, 2nd ed. New York, NY, USA: Springer, 2005, doi: 10.1007/b136381.

[31] W. W. Cooper, L. M. Seiford, and J. Zhu, Eds., Handbook on Data Envelopment Analysis, 2nd ed. New York, NY, USA: Springer, 2011, doi: 10.1007/978-1-4419-6151-8.

[32] R. M. Baron and D. A. Kenny, "The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations," Journal of Personality and Social Psychology, vol. 51, no. 6, pp. 1173-1182, 1986, doi: 10.1037/0022-3514.51.6.1173.