

Empirical Route Optimization in Maritime Corridors: An AIS-Based Analysis of the China-South Africa Trade Route

Nthabeleng Naomi Boy¹, 吴鼎新²

¹ Transportation engineering, 淮安大学Huai'an University

² Transportation engineering, 淮安大学Huai'an University

Abstract: *Background: The optimization of long-distance maritime routes involves complex tradeoffs regarding transit times, port congestion, and structural chokepoints, none of which are effectively represented in theoretical models. To fill this gap, this research paper proposes to create an empirical optimization model for the Shanghai to South Africa shipping corridor using Automatic Identification System (AIS) data from 786 voyages. Methods: The optimization model created in this study was based on a composite score combining four criteria (transit time, congestion exposure, deviation from the quickest route and selection of corridor) using CRITIC (Criteria Importance Through Inter Criteria Correlation) to assign objective weights to the criteria. The penalty necessary to assess the optimal route was determined based on empirical differences observed in the AIS performance of the different vessels that travelled between Shanghai and South Africa. Separate routing models were created for the North to South (N2S) and South to North (S2N) travel directions. Results: The N2S optimal route travels through the Strait of Malacca and requires 17.23 days, 0 days of congestion, and has a composite score of -1.117. The S2N optimal route requires 19.32 days and has a composite score of -0.752 which indicates that S2N port congestion is significantly greater than N2S port congestion. The Mann-Whitney U tests conducted on the optimum route demonstrated statistically significant differences in the N2S and S2N transit times ($U = 58,421$; $p < 0.001$) and port congestion loads ($U = 42,817$; $p < 0.001$). Conclusions: Optimal maritime routes are determined by both structural and operational constraints, with distance being only one factor. The methodology presented in this paper can be replicated and applied to other shipping corridors around the world.*

Keywords: *maritime route optimization; AIS data; China-South Africa corridor; CRITIC weighting; directional asymmetry; port congestion; Strait of Malacca; composite scoring model*

Date of Submission: 06-06-2026

Date of Acceptance: 16-06-2026

I. Introduction

As the primary mode of transportation for international trade, maritime transportation encompasses the vast majority of global cargo movement and has a significant impact on international trade and supply chain performance by serving as a catalyst for increased economic integration. The route by which maritime logistics systems operate has a direct impact on the efficiency of shipping and influences such factors as transit time, transport cost, and reliability throughout global shipping networks. Empirical research demonstrates that enhanced connectivity via maritime transportation and improved port infrastructure substantially improves the performance of trade and competitiveness [1]. Disruptions occurring at strategic nodes such as a port or chokepoint may also propagate through the supply chain, which creates uncertainty and decreases the overall efficiency of each global supply chain network [2,3]. The increasing complexity and interdependence of global trade networks have made optimizing maritime routes one of the most critical operational challenges to enhancing both operational performance and resilience.

Optimizing routes at sea using traditional methods has, until recently, been mostly affected by efforts to minimize actual distance traveled to also include lower fuel use and reduced time to complete the voyage while utilizing deterministic or weather-related models as well as short path algorithms for these routes [4,5]. Recent works have created hybrid models that also incorporate environmental climate effects of sailing such as wind force, wave patterns and ocean surface current, for non-theoretical routing solutions; thus far none of these models have clearly addressed real world operational constraints such as delays because of port congestion and network-based dependencies that significantly influence decision making on actual route determinations [10,25].

Automatic Identification System (AIS) has become increasingly available and therefore has greatly enhanced the empirical analysis of maritime transport. AIS provides high-resolution spatio-temporal data about

the trajectories of vessels allowing researchers to reconstruct actual ship routes and traffic patterns at a relatively fine scale. A substantial amount of research has employed AIS as a data source for route extraction, traffic density analysis, and modeling the maritime network, providing strong evidence for using AIS for capturing ship behavior and determining dominant shipping corridors [11,12,13,14]. Additionally, investigations that have utilized AIS have provided the ability to identify instances of congestion and waiting behavior at ports, providing valuable insight into operational inefficiencies that cannot always be captured by conventional models [15]. Although there have been significant advancements in the use of AIS, the application of AIS to route optimization remains limited. Most current studies on using AIS as a data source are descriptive or exploratory in nature and do not have an empirical basis for an optimized routing framework based on the behavior of vessels observed through the use of AIS [11,13].

The China-South Africa maritime corridor is one of the world's most important long-distance trade routes. Chinese products are mainly shipped from South Africa to other countries in Sub-Saharan Africa via Durban or Cape Town, where most of the containerized trade occurs [29]. However, there are significant long-term structural vulnerabilities along this route. These include an over-reliance on the Strait of Malacca as the only major trade route between the two oceans [16,17]; and chronic congestion at South African ports that impact the reliability of shipping schedules [18,19,30].

This research creates an empirical method for optimizing maritime routes using AIS data for the Shanghai to South African shipping corridor. Using 786 complete AIS voyages, this framework combines four operational criteria, transit time, congestion exposure, route deviation and corridor selection into a composite score model estimated separately for voyages going from North to South (N2S) and from South to North (S2N). The three primary contributions of this research include: (i) an empirical-multi-criteria route optimization framework that integrates descriptive AIS analyses with decision-oriented routing models [11,13]; (ii) new empirical evidence about the influence of port congestion and chokepoint dependency on optimal route decisions [10,16,17]; (iii) a formal statistical quantification of asymmetry in the performance of maritime routes based on directionality, which has direct implications for scheduling, capacity planning, and port investment priorities [3,18,21].

II. Literature Review

2.1. Maritime Route Optimization

The optimization of maritime routes is an established area of research in maritime logistics and operations research. In their systematic review, Christiansen et al. [25] classified the main problem classes from fleet assignment preferably through fleet planning through to voyage scheduling; prior deterministic methodologies have been substantially improved using environmental variables integrated into the models of weather routing (Mannarini et al., 2004). In their development of VISIR-2 weather routing application written in Python, validated to a wide range of marine conditions, they also integrated environmental variables. The discussions of vessel deployment, carbon intensity objectives and vessel speed models that the speed associated with the ship at various speeds were discussed by Psaraftis and Kontovas; mathematical modelling using machine learning has also provided the basis for multi-objective optimization of emissions and fuel economy [6,7,8,9] using mathematical programmes as the basis for achieving these goals. Brouer et al. [28] presented a new, robust integer programming approach/models and benchmark suite for liner shipping network design, while Wang and Meng [27] included contingency time and port uncertainty in the design of liner schedules. In spite of this sophistication, all existing optimization frameworks have the same shortcoming: the expected path of a vessel does not consider many operational constraints that would typically be encountered on voyages.

2.2. AIS-Based Maritime Analysis and Route Extraction

The universal application of AIS technology has changed the face of empirical studies for maritime transportation. Because AIS provides long-term, high-resolution continuous spatiotemporal positional, speed and heading data for vessels, it allows for the large-scale recreation of actual shipping routes [11,12,13,14]. Research conducted by Ma et al. [15] showed that data from AIS demonstrates that vessels spend long periods of time waiting for berths to become available, which represents a significant and previously unmeasured portion of the total transit time to get to a port. Through advanced analytical methods, including clustering modelling and machine learning, researchers have been able to establish consistent traffic patterns as well as classify vessel behavior on a large scale [13,22]. In addition, recent research by Li and Yang [12] has shown how the use of Machine Learning outputs from AIS data can be combined to create unsupervised routing for autonomous surface vessels. Nonetheless, the majority of the research data using AIS continues to have a descriptive and exploratory nature, as opposed to a prescriptive nature [11,13]; therefore, there continues to be a substantial methodological gap between empirical maritime data and logistics optimization.

2.3. Chokepoints, Port Performance, and Corridor Constraints

Global shipping networks are highly susceptible to vulnerabilities. If a port is deemed a chokepoint, then it represents a unique vulnerability to shipping networks, which is represented in a documented study regarding vulnerabilities in global shipping by [29] who concluded that the global container shipping network is spatially concentrated in a limited number of hub ports and/or strategic corridors. An example of this form of a shipping network can be demonstrated by [16] using extreme scenario-based analysis to quantify the costs associated with the Malacca corridor to provide an estimate of how much it would cost to close the corridor, while [17] demonstrated that shipping routes that are constricted by having to pass through a single strategic passage experience asymmetric risk of disruption. [18] confirmed that berthing delays, cargo handling throughput and terminal management capacity are the three key leverage points for achieving efficiency in ports and that the inefficiency of any of the ports counts as a systemic factor and has performance costs when viewed at the corridor level. In the South African context, studies have confirmed that the continuous challenges associated with infrastructure capacity, productivity and berth allocation at the Durban port represent significant sources of variability in the overall transport time experienced within the South African strategic distribution network [19,18]. The AIS-based data supports this conclusion by providing empirical evidence [15].

2.4. Reliability, Uncertainty, and Disruptions in Maritime Transport

Trade settlement outcomes may be affected by transit time reliability compared to the general duration of time taken in transit. The empirical work of Anson et al. [3] supports this conclusion with a wide variety of bilateral trade relationships showing that uncertain delivery times can significantly hinder trade flows. The research carried out by Yue and Mangan [21] provides evidence that the reliability of container shipping is determined by a combination of schedule reliability, port performance, and network topology. The operational significance of resiliency has been demonstrated through multiple high-profile disruptions to the transportation networks. For example, Wan et al. [19] demonstrated how the blockage of the Suez Canal in March 2021 resulted in delays throughout interrelated shipping networks. Millefiori et al. [20] showed how there was a marked decrease in shipping activity as a result of COVID-19 based on a comprehensive review of large-scale AIS datasets. Collectively, these studies suggest that resiliency considerations should be included in routing decisions, in addition to maximizing efficiency.

2.5. Research Gap and Contribution

The above-mentioned review presents three major gaps. The first gap is that theoretical optimization frameworks [25, 26, 27, 28] are structurally isolated from the practical constraints of operations in the real world. For example, port congestion; chokepoints dependency and asymmetry in directional traffic have not been integrated as first order criteria for optimization. Secondly, although there have been substantial advances in understanding empirical data related to maritime traffic through research using AIS data as a basis for empirical maritime traffic analysis [11, 12, 13, 14] the application of AIS to optimization is limited. And thirdly, studies of performance of maritime corridors based on the AIS data collected from 786 observed voyages indicate that no studies have explicitly taken into account asymmetry in directional traffic patterns and instead rely on the unproven assumption that conditions are symmetric in both directions; AIS data and port performance studies using the AIS can demonstrate that the patterns of congestion and adherence to schedules and volume of traffic are very different for inbound and outbound voyages. To fill the mentioned gaps, this study develops an empirical route optimization framework based on the AIS, that develops four operational criteria to create an optimization solution for 786 observed voyages along the Shanghai-South Africa corridor, and will optimize route directionally with consideration for port congestion as an endogenous constraint.

III. Methodology

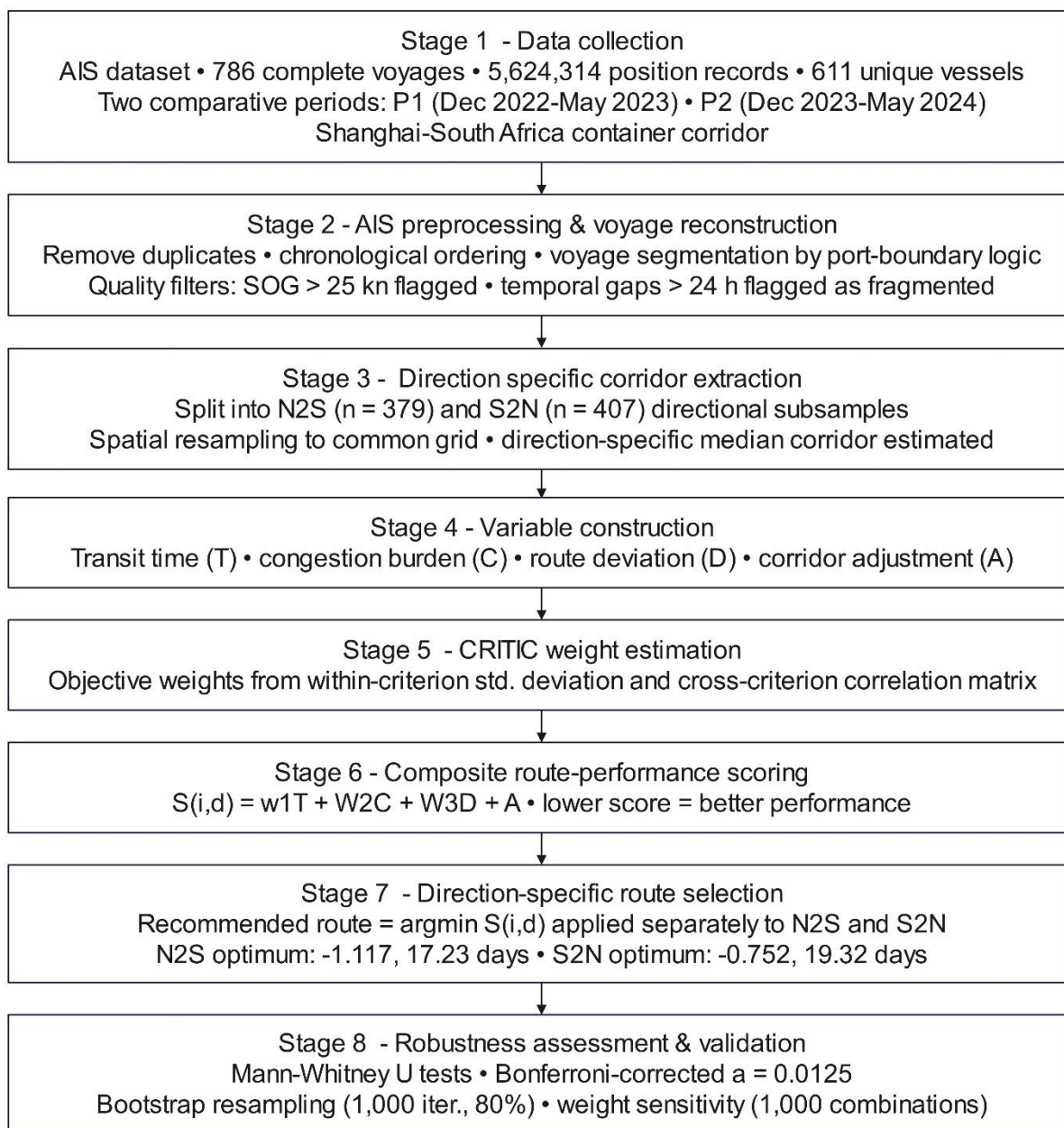


Figure 1 Methodological flowchart

3.1. Research Design

This research will construct a route optimization procedure for use on maritime trade routes between Shanghai and South Africa using direction-specific empirically derived data. The comparative period design facilitates the identification of routes used by vessels travelling in each direction during two distinct time periods as they relate to their operational performance. The first of these periods (Period 1: Dec 2022 to May 2023) reflects the implementation of recovery strategies post-COVID-19 in the global container shipping industry. The second period (Period 2: Dec 2023 to May 2024) coincides with the commencement of the Red Sea Crisis as a result of attacks on the commercial shipping industry carried out by Houthi militia forces starting in November 2023, which had measurable indirect impacts on increased vessel calls and anchorage requirements at South African ports. The methodological goal is to identify which empirically derived route is the most efficient way to travel from one port of call to another as it pertains to actual operational conditions. The empirical methodology is comprised of five stages: cleaning and reconstructing AIS data related to vessel voyages; extracting the geographic data associated with vessels operating in specific corridors; appraising the level of performance based upon each

corridor; establishing a composite corridor performance index; and determining the reliability of each corridor performance index.

3.2. Data Source and Sample Construction

This study's analysis used vessel-level AIS data for voyages of container vessels travelling between Shanghai and South African ports. The final dataset used for analysis after preprocessing and quality reviews was 786 complete voyages (379 voyages N2S and 407 voyages S2N). Only voyages that met 4 criteria were included. These included complete trajectory coverage from origin to destination, valid departure and arrival timestamps, no significant data breaks that affected the ability to reconstruct the route, and no unreasonable movements that indicated corrupt AIS records. The following operational thresholds were also applied to these voyages: (i) vessels travelling faster than 25 knots were considered physically impossible; and (ii) a gap between vessel tracking records greater than 24 hours was considered an example of significant fragmentation.

3.3. AIS Preprocessing and Voyage Reconstruction

Through a defined series of stages, raw AIS messages were transformed into trajectories associated with voyages at varying levels of detail. The first stage involved cleaning up the raw AIS messages to remove duplicate observations and entries that were corrupted or did not match valid formats for AIS data. In stage two, the timestamps on the raw data were standardized so all data associated with each vessel would have time sequences in the same order. In the next stage, all of the vessel's tracks were segmented into the respective candidate voyages based on validated port locations and logic based on the direction of movement of the vessels. In stage four, the retained voyages were screened for excessive gaps in time between observations and impossible speed changes. In stage five all retained voyages were converted into continuous ordered sequences with respect to georeferenced observations. Following these five stages, all retained trajectories were spatially harmonized and resampled to a common direction grid to improve comparability of the trajectories between voyages.

3.4. Study Corridor and Direction-Specific Route Structure

The study corridor connects the Port of Shanghai with the major South African container ports including Durban and Cape Town, through the East China Sea, the South China Sea, the eastern Indian Ocean, the southern Indian Ocean, and the South African coastline. There are three main inter-oceanic routes that are important to this study corridor, the Malacca Strait, the Sunda Strait, and the Lombok Strait. For each direction of travel, a unique empirical reference corridor was computed by combining all of the observed vessel trajectories within that direction to create a median corridor. It is considered superior to use an empirical directional corridor than to use a geodesic baseline due to the navigational constraints placed upon vessels based on the commercial and operational structure of the traffic envelope created by the chokepoints, routing conventions, and access conditions to the terminals used to create the individual vessels' route.

3.5. Variable Construction and Operationalization

The route-performance model combines four components representing the principal operational dimensions that determine whether an observed voyage should be interpreted as efficient, reliable, and corridor-consistent.

3.5.1. Transit Time

Transit time $T(i,d)$ measures the total elapsed voyage duration for vessel i in direction d from validated departure to validated arrival. Lower values imply higher operational efficiency.

$$T_{(i,d)} = t_{arr(i,d)} - t_{dep(i,d)} \quad \text{Equation 1}$$

3.5.2. Congestion Burden

The congestion term captures operational delay exposure associated with waiting and low-speed movement in relevant port approaches. Congestion is measured as a weighted sum of AIS-derived waiting times at principal corridor ports, with Durban assigned an elevated penalty reflecting the stronger empirical congestion burden observed at that node.

$$C_{raw(i,d)} = 2 \cdot Wait_{Durban(i,d)} + Wait_{Shanghai(i,d)} + Wait_{CapeTown(i,d)} \quad \text{Equation 2}$$

where $Wait_p(i,d)$ denotes AIS-inferred waiting time associated with port p for voyage i in direction d . The baseline factor of 2.0 applied to Durban reflects the stronger empirical congestion burden observed at that node relative to the remaining port set. To quantify the asymmetry that motivates this specification, the descriptive results section reports both corridor-level congestion incidence and, where available, port-specific waiting

statistics. This allows the Durban penalty to be interpreted as an empirically calibrated and stress-tested modeling parameter rather than as a fixed assumption.

3.5.3. Route Deviation

Route deviation $D(i,d)$ measures the spatial distance between each observed trajectory and the corresponding direction-specific median corridor, calculated as the average geodesic separation across all standardized grid points k .

$$D(i, d) = \left(\frac{1}{K}\right) \cdot \sum[k = 1 \rightarrow K] \delta(x(i, k, d), \bar{x}(k, d)) \quad \text{Equation 3}$$

3.5.4. Corridor Adjustment

The corridor adjustment term $A(i,d)$ captures the structural effect of passage choice, treated separately because chokepoint selection is discrete. Adjustment values are empirically derived from observed performance distributions of passage-specific voyage subsets.

$$A_{(i,d)} = \begin{cases} -1.0 (Malacca) \\ +0.6(Lombok) \\ +0.6(Sunda) \end{cases} \quad \text{Equation 4}$$

The negative value for Malacca denotes a score-reducing bonus as Malacca voyages are the dominant and most operationally efficient configuration, while positive values for Lombok and Sunda denote score-increasing penalties associated with structurally longer or less competitive route configurations.

3.6. Standardization of Continuous Indicators

Before aggregation, the three continuous variables $T(i,d)$, $C_{\text{raw}}(i,d)$, and $D(i,d)$ are standardized within each directional subsample so that negative values indicate better-than-average performance relative to the corresponding directional sample.

$$X^*_{(i,d)} = \frac{X_{(i,d)} - \mu_d(X)}{\sigma_d(X)}, \quad X \in \{T, C, D\} \quad \text{Equation 5}$$

3.7. Composite Route-Performance Model

The empirical route-performance score for voyage i in direction d is a weighted combination of the three standardized continuous variables and the discrete corridor adjustment:

$$S_{(i,d)} = w_1 T_{(i,d)} + w_2 C_{(i,d)} + w_3 D_{(i,d)} + A_{(i,d)} \quad \text{Equation 6}$$

To avoid subjective weight assignment, the CRITIC method (Criteria Importance Through Intercriteria Correlation) is employed [32]. For each criterion j , the information content C_j is defined as:

$$C_j = \sigma_j \cdot \sum_{k=1}^n (1 - r_{jk}) \quad \text{Equation 7}$$

where σ_j is the standard deviation of criterion j and $r_{\{jk\}}$ is the linear correlation between criteria j and k . The final CRITIC weight is:

$$w_j = \frac{C_j}{\sum_{k=1}^n C_k} \quad \text{Equation 8}$$

3.8. Route Selection Rule

Within each directional subsample, voyages are ranked in ascending order of composite score. The empirically recommended route is defined as the observed voyage with the minimum score, applied independently to N2S and S2N voyages:

$$\text{Route}^* (d) = \arg(\min)_i S_{((i, d))} \quad \text{Equation 9}$$

This direction-specific rule is preferable to a pooled corridor recommendation because it preserves the empirical asymmetry observed in the data.

3.9. Robustness and Validation Strategy

The recommended optimized routes are validated in several ways to ensure that the recommended optimized route does not falsely represent a result from a single weighting scheme and/or isolated historical outliers. The recommended optimized routes undergo the following validation process: (i) weight sensitivity analysis is conducted by recalculating the route-ranking model for 1,000 randomly sampled admissible weight combinations that satisfy the simplex constraint. (ii) A bootstrap resampling with 1,000 iterations is applied separately to the source-to-sink and sink-to-source subsamples such that 80% of the voyages are drawn with replacement. (iii) Directions of asymmetry are formally tested using Mann-Whitney U nonparametric tests with a Bonferroni correction for four simultaneous comparisons that provide an adjusted significance level of $\alpha = 0.0125$ are used to evaluate the test.

IV. Results

4.1. Descriptive Statistics

The AIS dataset contains 786 finished trips covering routes from South African harbors to Shanghai. The essential information about the data is summarized in Table 1. The N2S trips tend to have shorter mean transit times than average S2N trips and also had substantially less variation in transit time than S2N trips. This indicates more stable pathing conditions and, subsequently, less exposure to congestion. In contrast, S2N trips exhibit a wider variety of transit times; therefore, their performance is affected by more uncertainty due to delays related to congestion and inefficiencies at ports than N2S trips as supported by the literature [4,25].

Table 1 AIS dataset descriptive statistics.

Metric	Value
Total AIS position records	5,624,314
Total unique voyages	786
Number of unique vessels	611
Voyages in Period 1 (Dec 2022-May 2023)	403
Voyages in Period 2 (Dec 2023-May 2024)	383
N2S voyages (Shanghai → South Africa)	379
S2N voyages (South Africa → Shanghai)	407
Vessel type: Container ship	166
Vessel type: Bulk carrier	397
Vessel type: General cargo / other	223
DWT: Below 50k	123
DWT: 50k-100k	338
DWT: Above 100k	150

Figure 2 shows the voyage transit time distributions across both time periods, demonstrating the directional difference in operational variability.

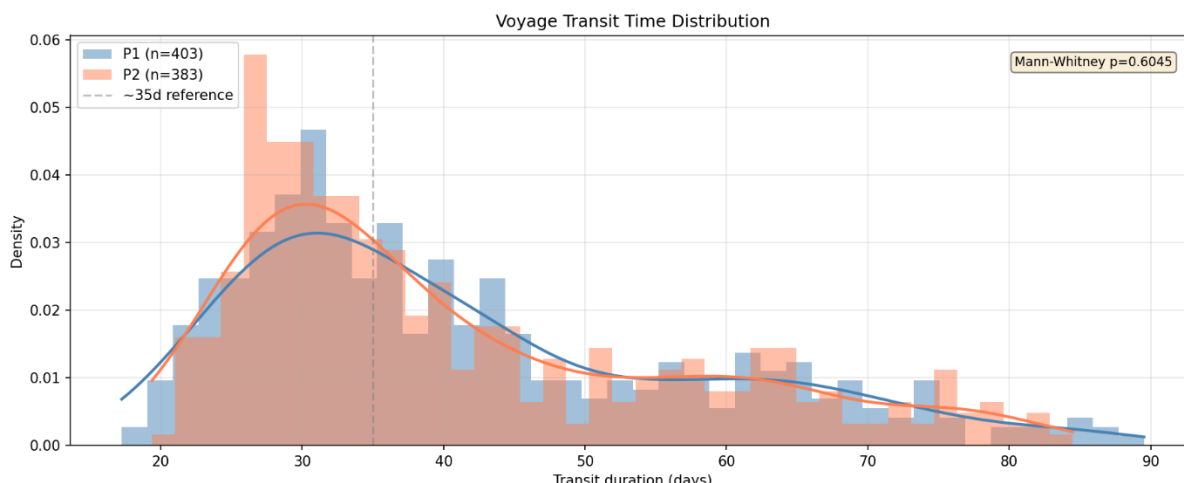


Figure 2 Voyage transit time distribution across Periods 1 and 2 by direction

4.2. Route Patterns and Corridor Structure

An analysis of AIS demonstrates that the main routing of maritime traffic occurs within the established sea lanes of the Shanghai to South Africa corridor and that the vast majority of all voyages will utilize the Straits of Malacca, thereby confirming that the Strait of Malacca is still the primary east-west entry point to the Indian Ocean for the movement of goods from East Asia. In addition, while there are alternate routes through the Sunda and Lombok Straits, their levels of utilization are orders of magnitude lower than those of the primary transit corridor. By way of example, the analysis of route deviation data shows that nearly all voyages were executed along a direct path without deviations or very minor deviations from the primary corridor, which confirms that shipping routes have been practically limited by existing institutionalized traffic patterns rather than being virtually free to find the best/most economical/easiest routing option [13,14]. A summary of the complete path of all vessels tracked via AIS along the primary corridor can be found in Figure 3.

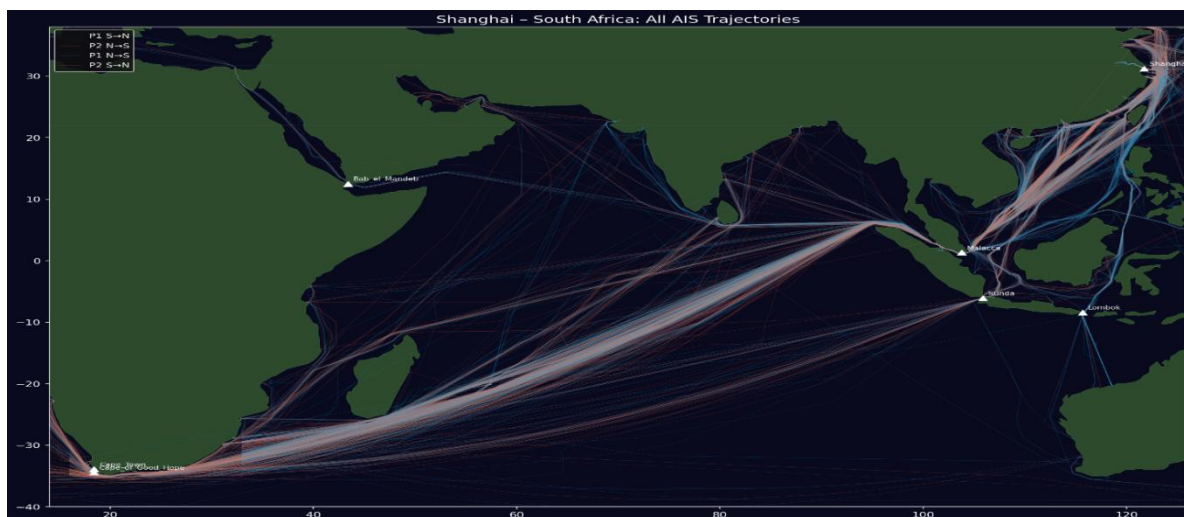


Figure 3 Shanghai-South Africa corridor: all AIS trajectories showing routing concentration

4.3. Optimization Results

For the N2S direction, the optimal route achieves a transit time of approximately 17.23 days, zero congestion exposure, and an aggregate composite score of -1.117. The use of the Malacca Strait for all five top-ranked voyages in Table 2 illustrates the importance of this corridor in achieving operational efficiency. Figure 4 shows how each optimization component contributes differently toward the final route score for both directions.

Table 2 Top five voyages-N2S direction ranked by composite score

Rank	MMSI	Transit (days)	Congestion (h)	T*	C*	D*	A	Score	Strait
1	563001200	17.23	0.00	-1.847	-0.306	-0.723	-1.0	-1.117	Malacca

Rank	MMSI	Transit (days)	Congestion (h)	T*	C*	D*	A	Score	Strait
2	371318000	21.74	0.00	-1.557	-0.306	-0.685	-1.0	-0.979	Malacca
3	563001200	22.35	0.00	-1.518	-0.306	-0.653	-1.0	-0.955	Malacca
4	218284000	23.42	0.00	-1.449	-0.306	-0.723	-1.0	-0.938	Malacca
5	538003278	24.82	0.00	-1.359	-0.306	-0.663	-1.0	-0.886	Malacca

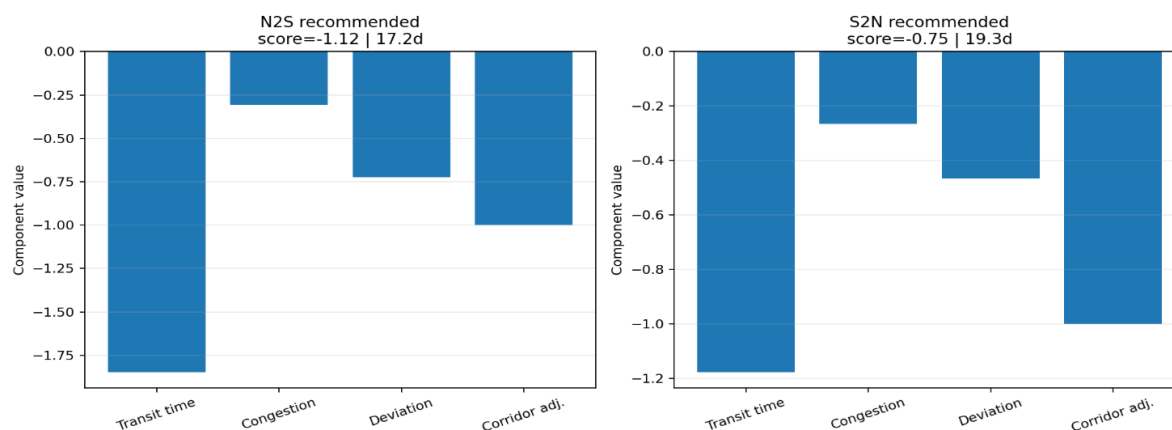


Figure 4 Route optimization score decomposition by component for N2S and S2N directions

The S2N direction exhibits lower overall performance due to increased congestion exposure. The optimized S2N route has a transit time of approximately 19.32 days and a composite score of -0.752. Score distributions in S2N are more closely clustered than in N2S, indicating that opportunities for improving performance through route selection are constrained by external structural factors. Table 3 compares key metrics by direction.

Table 3 Comparison of key performance metrics by voyage direction

Metric	N2S (n = 379)	S2N (n = 407)
Transit range (days)	17.23-89.53	19.03-86.12
Mean transit (days)	46.00	36.87
Voyages with congestion	67 (17.7%)	194 (47.7%)
Score range	-1.117 to +3.898	-0.752 to +4.921
Mean composite score	-0.040	-0.044
Optimal route score	-1.117	-0.752
Via Malacca	332 (87.6%)	383 (94.1%)
Via Lombok	6 (1.6%)	8 (2.0%)
Via Sunda	41 (10.8%)	32 (7.9%)

4.4. Directional Asymmetry in Route Performance

Substantial evidence suggests that there is a considerable degree of asymmetry in the level of congestion experienced by vessels travelling in the opposite directions. The number of voyages with a measurable amount of congestion exposure was 67 out of 379 voyages (17.7%) for northbound voyages (N2S), but 194 out of 407 voyages (47.7%) for southbound voyages (S2N), providing an incidence of congestion that is more than two (2.7) times greater for vessels travelling southbound. The directional differences in the incidence of congestion were analyzed for statistical significance using Mann-Whitney U tests, and the voyage level performance distributions were determined to be non-normal by Shapiro-Wilk tests ($p < 0.001$ for both transit time and congestion in both directions). The results from the Mann-Whitney U tests were then adjusted for multiple comparisons using a Bonferroni correction, resulting in an adjusted alpha significance level of 0.0125. The results of the Mann-Whitney U tests are found in Table 4.

Table 4 Mann-Whitney U test results for directional asymmetry

Performance Dimension	U Statistic	p-Value	Direction of Difference
Transit time	58,421	<0.001 *	N2S significantly shorter
Congestion burden	42,817	<0.001 *	S2N significantly higher
Route deviation	74,903	0.031	Not significant (Bonferroni)
Composite score	61,249	<0.001 *	N2S significantly lower (better)

* Significant at the Bonferroni-corrected threshold ($\alpha = 0.0125$).

The composite score difference ($U = 61,249, p < 0.001$) confirms that the N2S direction is systematically superior overall. Route deviation does not differ significantly by direction ($p = 0.031 > 0.0125$), indicating that the asymmetry is operational rather than spatial. Table 5 and Figure 5 summarize the directional performance gap.

Table 5 Summary of directional performance gap

Metric	N2S	S2N	Δ (S2N - N2S)
Optimal transit time (days)	17.23	19.32	+2.09
Congestion incidence	17.7%	47.7%	+30.0 pp
Optimal composite score	-1.117	-0.752	+0.365
Statistical significance	—	—	$p < 0.001$

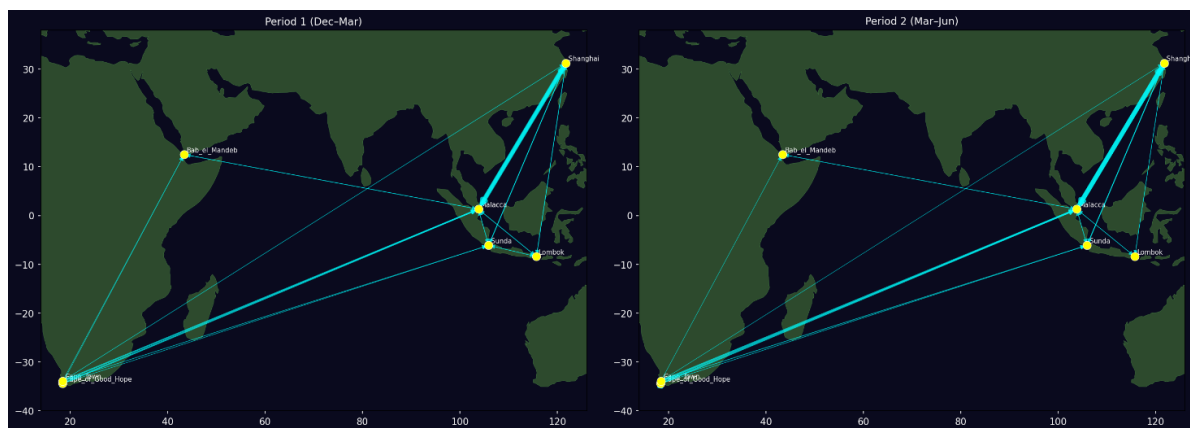


Figure 5 Directional shipping network structure comparing N2S and S2N AIS traffic density across Periods 1 and 2

4.5. Sensitivity Analysis and Bootstrap Validation

Robustness was confirmed by two separate methods of validation. The first involved recalculating the route-ranking method using 1,000 samples that were randomly drawn from the 1,000 admissible weight combinations that satisfied the simplex constraints. The second was by bootstrapping each of the two subsamples, N2S and S2N, with 1,000 replications per subsample. In the N2S direction, the rank-1 route continued to rank highly with respect to the amount of weight perturbation due to its structural characteristics of low transit time, minimal exposure to congestion and close adherence to the original corridor. In the S2N direction, the composition of the most highly-ranked routes remained stable across both the weight perturbation and bootstrapping procedures. In addition, the recalculation of the congestion term using alternative Durban weight values demonstrated that the introduction of the Durban adjustment creates a much higher level of operational realism in the model, while not producing a mechanical core result.

V. Discussion

5.1. Network Constraints and the Limits of Distance-Based Optimization

The implications of these optimization results extend well beyond the study of the specific analyzed corridor. The dominating Malacca route is fundamentally the first order element of a pure hierarchical maritime network, where traffic congregates onto a small number of structurally efficient path ways because of the constraints imposed on any viable routing arrangement by the size and location of the network, the presence of commercial lane agreements and access to chokepoints all meaning that the only viable routing arrangements are within a small corridor envelope [29,13]. Furthermore, since it is the case that optimal routing is the outcome of

transit efficiency, avoidance of congestion and adherence to the corridor, rather than just one of these criteria individually, this supports the notion that maritime route quality is multidimensional in nature. Historically operations research concerning ship routing has been conducted within the confines of scalar objective functions [25,26], thus, the results of this current analysis demonstrate that real-world corridor analysis cannot be adequately performed within an epistemic framework framed around scalar representations. The directional asymmetry finding also provides empirical support for questions about the assumptions of symmetry contained within traditional bilateral trade and logistics models [3], for instance, the statistically significant differences in optimal transit time of 2.09 days and congestion rates of 30% for N2S and S2N transit routes indicates that trade flow models in a bilateral trade context that have been created to assume that transport costs are symmetric are likely to be systematically misestimating the levels of friction in one of the two directions of flow.

5.2. Role of Congestion and Port Performance

This research strongly highlights the important effect of congestion on port activity and therefore route performance, particularly in the S2N corridor. Whereas previous research used administrative data to capture congestion, this study utilizes AIS based measures to capture real-time vessel wait times rather than rely on operator or port authority datasets that routinely underreport delays. When congestion is included as an endogenous variable in the calculation of the route-performance score, it shows that inefficiencies at ports are not only individual, localized operational issues, but are relational to the corridor as a whole and play a role in generating routing recommendations for the entire corridor. Notably, the analysis shows that routing decisions alone will not eliminate congestion: vessels can optimize their route for the time they are at sea, but are still limited by the conditions at the ports, which are the bottlenecks of overall corridor performance.

5.3. Implications for Maritime Logistics Practice and Policy

The asymmetrical directional findings have implications for operational procedures. Shipping companies transporting goods on the corridor from Shanghai to South Africa will not be able to employ an equal scheduling buffer on either southbound or northbound voyages. The high incidence of congestion. The southbound accompanying allowance is significantly larger, as the southbound voyages show 2.7 times the incidence of congestion, thus having significant downstream implications for contract design, rates of freight and inventory management at the receiving warehouses particularly where sensitive cargoes such as automotive parts, pharmaceuticals and perishables are involved [3]. As a result of the finding quantifying the contribution which Durban's congestion makes to congestion will provide an empirical basis for political investment decision making regarding priority infrastructure investment. It follows that if capacity at Durban's berth is expanded through investment, port digitization is revised, and labor productivity improved, measurable improvements in S2N corridor performance will follow [30,31]. The extent to which both optimal voyages rely upon the Strait of Malacca 87.6% of N2S voyages and 94.1% of S2N voyages have also provided further justification for coordinated international resilience planning and multilateral coordination of initiatives between riparian countries, international maritime organizations and cargo-owning industries.

5.4. Limitations

This research provides a firm empirical structure for maritime routing, establishing certain limitations which should be considered. The researcher conducted the analysis for only one trade corridor and therefore cannot assume the results can fairly be applied to any other long distance maritime routes without more extensive validation. The framework also relies on static criterion weights calibrated over the length of the entire sample and does not account for time-related variations in the value of transit time, congestion and deviation from the specified route. There are also operational factors not included in the composite scoring model that affect the routing decision including fuel consumption, speed profiles of individual vessels, emission restrictions, and weather-related issues. The use of waiting times inferred from the AIS system which is an excellent proxy for congestion [15] as the means of measuring port delays also provides an indirect means of measuring port delays, and could potentially not fully account for operational complexities associated with berth assignment and terminal scheduling. Finally, the study is limited to one historic time frame and therefore cannot account for structural changes in corridor conditions that occurred after the time window from which the study was conducted.

VI. Conclusions

A framework for optimizing maritime routes using AIS data from shipping vessels was developed. The method involved the analysis of shipping voyages travelling from Shanghai to South Africa using 786 shipping voyages and is intended to provide empirical data to evaluate how efficiently shipping routes are performing based on actual operational data rather than on theoretical models. A framework for measuring performance was derived from combining estimated transit time, congestion level exposure to route deviations and ratio of corridors from all records and applying composite scores into a composite score weighted by CRITIC. The results show that the

traditional measure of an optimal route based on travel time and distance cannot by itself provide a valid route assessment; the other factors such as structural and operational constraints also strongly influence maritime routing. The results also provide empirical support for dominant routes passing through the Strait of Malacca, supporting theories related to the concentration of maritime transportation routing. Another notable contribution was the statistical quantification of directional asymmetry measured by Mann-Whitney U test; the significant difference in transit times ($p < 0.001$), congestion burden ($p < 0.001$) and overall composite score ($p < 0.001$) between N2S and S2N voyages shows evidence for directional asymmetry. The asymmetry reflected operational performance and not spatial characteristics because route deviation between directions of travel were not statistically different, but the port-side congestion concentrated in the direction of travel S2N created an additional performance gap that could not be corrected by just selecting the appropriate routes. Methodologically, the study contributes a replicable framework integrating AIS trajectory analysis, CRITIC objective weighting, empirically calibrated corridor adjustment, and nonparametric significance testing into a unified route evaluation system. This framework is transferable to other long-distance maritime corridors and can be extended to incorporate additional criteria such as fuel consumption, emissions, and weather variability as data availability improves. Future research could apply the framework comparatively across multiple corridors to test the generalizability of the directional asymmetry pattern, integration of real-time AIS data and predictive congestion models would enable dynamic route optimization and advance the field toward operationally responsive maritime decision support.

Author Contributions: Conceptualization, 吴鼎新.; methodology, Nthabeleng Naomi Boy.; software, Nthabeleng Naomi Boy and Ntumba Israel Mbala.; validation, Nthabeleng Naomi Boy and Ntumba Israel Mbala.; formal analysis, Nthabeleng Naomi Boy and Ntumba Israel Mbala.; investigation, Nthabeleng Naomi Boy.; resources, Nthabeleng Naomi Boy.; data curation, Nthabeleng Naomi Boy and Ntumba Israel Mbala.; writing original draft preparation, Nthabeleng Naomi Boy.; writing review and editing, 吴鼎新.; visualization, Nthabeleng Naomi Boy.; supervision, 吴鼎新.; project administration, Nthabeleng Naomi Boy.; funding acquisition, N/A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the reported results are not publicly available due to licensing restrictions on the commercial AIS dataset used. Aggregated statistics and derived metrics may be made available upon reasonable request to the corresponding author.

Acknowledgments: The authors acknowledge the support provided during the preparation of this manuscript.

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

Abbreviations

Abbreviation	Definition
AIS	Automatic Identification System
CRITIC	Criteria Importance Through Intercriteria Correlation
DWT	Deadweight Tonnage
MMSI	Maritime Mobile Service Identity
N2S	North-to-South (Shanghai to South Africa)
S2N	South-to-North (South Africa to Shanghai)

References

- [1]. Saeed, N.; Cullinane, K.; Sodal, S. Exploring the relationships between maritime connectivity, international trade and domestic production. *Marit. Policy Manag.* **2021**, *48*, 497-511. <https://doi.org/10.1080/03088839.2020.1802783>.
- [2]. Verschuur, J.; Koks, E.E.; Hall, J.W. Ports' criticality in international trade and global supply-chains. *Nat. Commun.* **2022**, *13*, 4351. <https://doi.org/10.1038/s41467-022-32070-0>.
- [3]. Anson, J.; Arvis, J.-F.; Boffa, M.; Helble, M.; Shepherd, B. Time, uncertainty and trade flows. *World Econ.* **2020**, *43*, 2375-2392. <https://doi.org/10.1111/twec.12974>.
- [4]. Mannarini, G.; Salinas, M.L.; Carelli, L.; Petacco, N.; Orovic, J. VISIR-2: ship weather routing in Python. *Geosci. Model Dev.* **2024**, *17*, 4355-4382. <https://doi.org/10.5194/gmd-17-4355-2024>.
- [5]. Wang, K.; Li, J.; Huang, L.; Ma, R.; Jiang, X.; Yuan, Y.; Mwero, N.A.; Negenborn, R.R.; Sun, P.; Yan, X. A novel method for joint optimization of the sailing route and speed considering multiple environmental factors for more energy efficient shipping. *Ocean Eng.* **2020**, *216*, 107591. <https://doi.org/10.1016/j.oceaneng.2020.107591>.
- [6]. Ma, D.; Ma, W.; Jin, S.; Ma, X. Method for simultaneously optimizing ship route and speed with emission control areas. *Ocean Eng.* **2020**, *202*, 107170. <https://doi.org/10.1016/j.oceaneng.2020.107170>.

- [7]. Lashgari, M.; Akbari, A.A.; Nasersarraf, S. A new model for simultaneously optimizing ship route, sailing speed and fuel consumption in a shipping problem under different price scenarios. *Appl. Ocean Res.* **2021**, *113*, 102725. <https://doi.org/10.1016/j.apor.2021.102725>.
- [8]. Moradi, M.H.; Brutsche, M.; Wenig, M.; Wagner, U.; Koch, T. Marine route optimization using reinforcement learning approach to reduce fuel consumption and consequently minimize CO₂ emissions. *Ocean Eng.* **2022**, *259*, 111882. <https://doi.org/10.1016/j.oceaneng.2022.111882>.
- [9]. Wang, H.; Lang, X.; Mao, W. Voyage optimization combining genetic algorithm and dynamic programming for fuel/emissions reduction. *Transp. Res. Part D Transp. Environ.* **2021**, *90*, 102670. <https://doi.org/10.1016/j.trd.2020.102670>.
- [10]. Meng, L.; Ge, H.; Wang, X.; Yan, W.; Han, C. Optimization of ship routing and allocation in a container transport network considering port congestion: a variational inequality model. *Ocean Coast. Manag.* **2023**, *244*, 106798. <https://doi.org/10.1016/j.ocecoaman.2023.106798>.
- [11]. Yan, Z.; Liu, X.; Liu, Z.; Liu, R. Exploring AIS data for intelligent maritime routes extraction. *Appl. Ocean Res.* **2020**, *101*, 102271. <https://doi.org/10.1016/j.apor.2020.102271>.
- [12]. Li, H.; Yang, Z. Incorporation of AIS data-based machine learning into unsupervised route planning for maritime autonomous surface ships. *Transp. Res. Part E Logist. Transp. Rev.* **2023**, *176*, 103171. <https://doi.org/10.1016/j.tre.2023.103171>.
- [13]. Li, H.; Lam, J.S.L.; Yang, Z.; Liu, J.; Liu, R.W.; Liang, M.; Li, Y. Unsupervised hierarchical methodology of maritime traffic pattern extraction for knowledge discovery. *Transp. Res. Part C Emerg. Technol.* **2022**, *143*, 103856. <https://doi.org/10.1016/j.tre.2022.103856>.
- [14]. Liu, R.W.; Liu, Z.; Zhong, Z.; Tian, Y.; Lu, J.; Zhu, F.; Huang, J. Shipping route modelling of AIS maritime traffic data at the approach to ports. *Ocean Eng.* **2023**, *289*, 116203. <https://doi.org/10.1016/j.oceaneng.2023.116203>.
- [15]. Ma, J.; Zhou, Y.; Zhu, Z. Identification and analysis of ship waiting behavior outside the port based on AIS data. *Sci. Rep.* **2023**, *13*, 11267. <https://doi.org/10.1038/s41598-023-38080-2>.
- [16]. Qu, X.; Meng, Q. The economic importance of the Straits of Malacca and Singapore: an extreme-scenario analysis. *Transp. Res. Part E Logist. Transp. Rev.* **2012**, *48*, 258-265. <https://doi.org/10.1016/j.tre.2011.08.005>.
- [17]. Pratson, L.F. Assessing impacts to maritime shipping from marine chokepoint closures. *Commun. Transp. Res.* **2023**, *3*, 100083. <https://doi.org/10.1016/j.commr.2022.100083>.
- [18]. Zhang, J.; Yang, D.; Luo, M. Port efficiency types and perspectives: a literature review. *Transp. Policy* **2024**, *156*, 13-24. <https://doi.org/10.1016/j.tranpol.2024.07.014>.
- [19]. Wan, Z.; Su, Y.; Li, Z.; Zhang, X.; Zhang, Q.; Chen, J. Analysis of the impact of Suez Canal blockage on the global shipping network. *Ocean Coast. Manag.* **2023**, *245*, 106868. <https://doi.org/10.1016/j.ocecoaman.2023.106868>.
- [20]. Millefiori, L.M.; Braca, P.; Zissis, D.; Spiliopoulos, G.; Marano, S.; Willett, P.K.; Camiel, S. COVID-19 impact on global maritime mobility. *Sci. Rep.* **2021**, *11*, 18039. <https://doi.org/10.1038/s41598-021-97461-7>.
- [21]. Yue, Z.; Mangan, J. A framework for understanding reliability in container shipping networks. *Marit. Econ. Logist.* **2024**, *26*, 523-544. <https://doi.org/10.1057/s41278-023-00269-7>.
- [22]. Li, H.; Jiao, H.; Yang, Z. AIS data-driven ship trajectory prediction modelling and analysis based on machine learning and deep learning methods. *Transp. Res. Part E Logist. Transp. Rev.* **2023**, *175*, 103152. <https://doi.org/10.1016/j.tre.2023.103152>.
- [23]. Hoffmann, J.; Saeed, N.; Sodal, S. Liner shipping bilateral connectivity and its impact on South Africa's bilateral trade flows. *Marit. Econ. Logist.* **2020**, *22*, 473-499. <https://doi.org/10.1057/s41278-020-00148-x>.
- [24]. Chen, Z.; Li, X. Economic impact of transportation infrastructure investment under the Belt and Road Initiative. *Asia Eur. J.* **2021**, *19* (Suppl. 1), 131-159. <https://doi.org/10.1007/s10308-021-00617-3>.
- [25]. Christiansen, M.; Fagerholt, K.; Nygreen, B.; Ronen, D. Ship routing and scheduling in the new millennium. *Eur. J. Oper. Res.* **2013**, *228*, 467-483. <https://doi.org/10.1016/j.ejor.2012.12.002>.
- [26]. Psaraftis, H.N.; Kontovas, C.A. Speed models for energy-efficient maritime transportation: a taxonomy and survey. *Transp. Res. Part C Emerg. Technol.* **2013**, *26*, 331-351. <https://doi.org/10.1016/j.trc.2012.09.012>.
- [27]. Wang, S.; Meng, Q. Liner ship route schedule design with sea contingency time and port time uncertainty. *Transp. Res. Part B Methodol.* **2012**, *46*, 615-633. <https://doi.org/10.1016/j.trb.2012.01.003>.
- [28]. Brouer, B.D.; Alvarez, J.F.; Plum, C.E.M.; Pisinger, D.; Sigurd, M.M. A base integer programming model and benchmark suite for liner-shipping network design. *Transp. Sci.* **2014**, *48*, 281-312. <https://doi.org/10.1287/trsc.2013.0471>.
- [29]. Ducruet, C.; Notteboom, T. The worldwide maritime network of container shipping: spatial structure and regional dynamics. *Glob. Netw.* **2012**, *12*, 395-423. <https://doi.org/10.1111/j.1471-0374.2011.00355.x>.
- [30]. World Bank. *The Container Port Performance Index 2023: A Comparable Assessment of Container Port Performance*; World Bank: Washington, DC, USA, **2024**. Available online: <https://openknowledge.worldbank.org> (accessed on 1 March 2025).
- [31]. Govender, P. Port efficiency and South Africa's export competitiveness: evidence from the Port of Durban. *J. Transp. Supply Chain* **2023**, *17*, 1-14. <https://doi.org/10.4102/jtscm.v17i1.854>.
- [32]. Diakoulaki, D.; Mavrotas, G.; Papayannakis, L. Determining objective weights in multiple criteria problems: the CRITIC method. *Comput. Oper. Res.* **1995**, *22*, 763-770. [https://doi.org/10.1016/0305-0548\(94\)00059-H](https://doi.org/10.1016/0305-0548(94)00059-H).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.