SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 10, No. 3, (2024), pp. 325-336
Open Access. Online & Print

10.5281/zenodo.22761657

WEATHER-DRIVEN VARIATIONS IN VESSEL NAVIGATION BEHAVIOR: A DESCRIPTIVE-CORRELATIONAL STUDY USING AIS DATA FROM THE PORT OF PIRAEUS

Luis Eduardo Muñoz Guerrero[1*]

[1]Universidad Tecnológica de Pereira, Facultad de Ingenierías, Email: lemunozg@utp.edu.co

Received: 08/12/2024
Accepted: 25/12/2024

Corresponding Author: Luis Eduardo Muñoz Guerrero (lemunozg@utp.edu.co)

ABSTRACT

Maritime traffic near major ports is shaped by a combination of vessel characteristics and environmental conditions, yet the extent to which weather variables systematically influence navigation behavior remains underexplored at the level of individual, publicly available AIS datasets. This study examines the relationship between meteorological conditions (wind, and a wind-derived sea-state proxy) and vessel navigation behavior (speed, course variability, stopping frequency) in the waters surrounding the Port of Piraeus, Greece, and investigates whether this relationship is moderated by vessel type. Using a stratified subsample of the Piraeus AIS Dataset (133 vessels across five type categories; 294,100 one-minute-resampled position reports spanning August, October and December 2017, chosen to capture summer meltemi, autumn-transition and winter-storm wind regimes) spatially and temporally joined to NOAA wind observations, we apply descriptive statistics, correlation analysis, and multiple regression to test whether wind conditions are associated with significant changes in vessel speed, heading variability, and dwell (stopping) behavior, and whether these effects differ across vessel categories (cargo, tanker, passenger, fishing, other). Wind speed showed a negligible zero-order correlation with vessel speed (Pearson $r = 0.008$) but a small, statistically significant effect once vessel type and its interaction with wind were modeled ($R^2 = 0.072$, $p < .001$): cargo vessels slowed modestly as wind increased, while passenger, tanker and fishing vessels showed a flat-to-slightly-positive wind–speed relationship. Heading variability was weakly and inconsistently associated with wind ($R^2 = 0.013$), with the direction of the effect differing by vessel type. The probability of a sustained stop was significantly, if modestly, related to wind speed (pseudo-$R^2 = 0.104$, $p < .001$) and strongly differentiated by vessel type, with harbor/service ("Other") vessels far more likely to be recorded as stopped than cargo, tanker or fishing vessels. Overall, wind conditions in this sample were consistently significant predictors of navigation behavior at this scale of data, but explained only a small share of total variance, and vessel type was the dominant moderator — supporting a nuanced, type-specific rather than uniform weather-response model of port-approach navigation. This work contributes an applied, reproducible methodology for weather-behavior analysis using open AIS data, with implications for port traffic management and maritime safety planning.

KEYWORDS: AIS Data; Maritime Traffic; Weather Conditions; Vessel Behavior; Descriptive-Correlational Study; Port of Piraeus.

1. INTRODUCTION

Automatic Identification System (AIS) technology has become a cornerstone of maritime surveillance, generating dense, high-frequency records of vessel position, speed, and heading. The growing availability of open AIS datasets has enabled a wide range of applications, from traffic prediction to anomaly detection (Yang et al., 2019). However, much of this literature emphasizes predictive modeling (e.g., trajectory forecasting) over descriptive characterization of why vessels behave as they do under varying environmental conditions.

Weather is a well-established determinant of vessel operations in maritime engineering practice, yet few open, reproducible studies have quantified this relationship using large-scale public AIS data combined with co-located meteorological records. The Piraeus AIS Dataset (Tritsarolis et al., 2022) offers a rare opportunity to do so: it integrates 2.5 years of AIS records from one of the busiest ports in the Mediterranean with spatially and temporally correlated NOAA weather forecasts, already merged and ready for analysis. This is particularly relevant in the Aegean, where strong seasonal Etesian (meltemi) winds are known to reach gale force and disrupt maritime transport (Poupkou et al., 2011).

This study addresses the following research questions:

1. RQ1: Do adverse weather conditions (specifically, high wind speed — the only directly measured meteorological variable in this dataset, see Section 3.3) show a statistically significant association with vessel speed and heading variability?
2. RQ2: Does this association differ depending on vessel type?
3. RQ3: Are stopping/dwell events near the port more frequent under specific weather conditions?

2. LITERATURE REVIEW

The Piraeus AIS Dataset was introduced by Tritsarolis, Kontoulis and Theodoridis (2022) as an open, pre-integrated resource combining AIS kinematic records, vessel static metadata, geographic reference layers, and spatio-temporally matched NOAA weather grids for the port of Piraeus. The same research group has used related versions of this data infrastructure for pattern-mining tasks such as anchorage and co-movement discovery on streaming vessel trajectories (Tritsarolis et al., 2021), building on foundational unsupervised AIS pattern-discovery approaches such as Pallotta et al. (2013), and — like the present study's stop/dwell detection — relies on sustained low-speed segments as the operational signature of anchoring or waiting behavior, though with a focus on real-time pattern discovery rather than weather attribution. The present design also follows the closer antecedent set by Shu et al. (2017) and Zhou et al. (2020), who combined AIS records with external condition data (wind, current, visibility) to quantify their statistical association with vessel speed, heading, and lateral position in port waterways, extending that line of inquiry to the Piraeus context with a stratified, vessel-type-differentiated design.

More broadly, two largely separate literatures motivate this study. The first is an engineering-oriented "weather routing" literature (Zis et al., 2020), which models the effect of wind and wave conditions on fuel consumption, speed loss, and voyage time for specific vessel classes, typically using onboard sensor or noon-report data (e.g., Vitali et al., 2020) rather than open AIS archives. The second is a rapidly growing machine-learning literature on AIS trajectory prediction and anomaly detection (e.g., Capobianco et al., 2021), which treats weather (when included at all) as one of many input features to a predictive model rather than as the object of inferential analysis — a gap noted explicitly in reviews of the maritime anomaly-detection field (Riveiro et al., 2018). Open, large-scale AIS datasets comparable in spirit to the Piraeus dataset — e.g., the U.S. MarineCadastre AIS archive — have primarily supported this predictive/ML strand.

This study is positioned between the two: it uses a public, reproducible AIS dataset (as in the ML strand) but asks a descriptive-correlational, engineering-relevant question (as in the weather-routing strand) — whether, and for which vessel types, wind conditions are statistically associated with observable changes in speed, heading stability, and stopping behavior near a major Mediterranean port. To our knowledge, no prior published analysis has used the Piraeus AIS Dataset specifically for this weather-behavior question, which is the gap this paper addresses.

3. DATA AND METHODS

3.1.1. Data Source

The Piraeus AIS Dataset (Tritsarolis, Kontoulis and Theodoridis, 2022) contains:

1. AIS dynamic data: kinematic records <timestamp, vessel_id, lon, lat, heading, course, speed>
2. AIS static data: vessel type and country

3. Weather data: NOAA-derived meteorological forecasts, spatially and temporally joined to the AIS records
4. Geographic reference data: ESRI shapefiles of the area of interest (port, coastline)
5. Coverage: May 9, 2017 – December 26, 2019; >244 million AIS records

3.1.2. Sampling Strategy

Given the dataset's scale (>244 million kinematic records across 2017–2019), a stratified subsample was drawn rather than processing the full corpus, to keep the analysis feasible for a single researcher without dedicated computing infrastructure. Concretely:

1. Vessel selection. The AIS static table (6,230 vessels) was classified into five categories from the declared AIS shiptype code — Cargo (70–79), Tanker (80–89), Passenger (60–69), Fishing (30), and Other (all remaining codes, including service/harbor craft, unclassified and default codes). 100 vessels were sampled at random (fixed seed) from each category, yielding a candidate pool of 500 vessels (100 per category). Of these, only 133 unique vessels actually produced dynamic (kinematic) records in at least one of the three selected months and were retained in the final analytic sample — the remainder of the candidate pool was simply inactive (no transmitted positions) during August, October or December 2017. This candidate-to-active attrition is itself informative: it indicates that most registered vessels in the static table are not continuously present near Piraeus, consistent with a port that serves a mix of regular and occasional/seasonal traffic.

2. Time window. Three non-consecutive months of 2017 (the dataset's first year) were selected to maximize real variation in wind regime: August (summer meltemi/Etesian winds, a persistent and well-characterized seasonal regime over the Aegean; Tyrlis and Lelieveld, 2013; Poupkou et al., 2011), October (autumn transition), and December (winter storm season).

3. Trajectory extraction and resampling. All dynamic pings belonging to the candidate vessels were streamed (chunked reads) out of the three-monthly AIS dynamic files (8.0M, 4.3M and 6.0M raw rows respectively) and resampled to at most one position report per vessel per minute, to bound file size while preserving enough temporal density for rolling-window feature derivation.

4. Weather join. Each resampled position was matched to the nearest NOAA weather grid station (of ~45 stations covering the area, spatial nearest-neighbor) and the temporally nearest reading from that station's ~3-hour-cadence time series (merge_asof, nearest direction), an approach consistent with prior work coupling large-scale AIS voyage data with external gridded weather/hindcast sources (Vitali et al., 2020).

5. Result. 294,100 vessel-minute observations from 133 vessels across the three months (137,086 in August, 60,513 in October, 99,699 in December; 3,198 additional rows were dropped because they fell at the start of a vessel's trajectory, before a full 30-minute heading-variability window could be computed).

This sampling frame — a stratified vessel pool rather than a full census, and three selected months rather than the full 2.5-year span — is documented explicitly as a scope limitation in Section 6.

3.1.3. Variables

Table 0: Study Variables, Their Type, Source and Analytical Role.

Variable	Type	Source	Role
Vessel speed (knots)	Continuous	AIS dynamic	Dependent
Heading variability (std. dev. over trajectory segment)	Continuous	AIS dynamic (derived)	Dependent
Dwell/stop frequency near port	Count	AIS dynamic (derived)	Dependent
Wind speed	Continuous	NOAA weather	Independent
Sea state / wave-related proxy (if available)	Continuous/Ordinal	NOAA weather	Independent
Vessel type	Categorical	AIS static	Moderator
Time of day / season	Categorical	Derived from timestamp	Control

Derived variable definitions. Because course-over-ground is a circular (angular) quantity — 0° and 359° are adjacent, not extreme values — heading variability is computed as the circular standard deviation of course (via the resultant vector length of unit vectors on the unit circle, cf.

Mardia and Jupp, 2000) over a rolling 30-minute per-vessel window, rather than a linear standard deviation, which would overstate variability for vessels crossing the 0°/360° boundary. A dwell/stop event is defined as a sustained period of speed below 1.0 knot for at least 10 consecutive minutes, to distinguish genuine stopping (anchoring, berthing, drifting) from brief low-speed readings caused by AIS position noise, following AIS-based operational definitions of vessel waiting/stopping used elsewhere in the literature (Pallotta et al., 2013; Franzkeit et al., 2020). Both thresholds (window length, speed cutoff, minimum duration) are implementation defaults in scripts/04_build_real_dataset.py (mirrored in scripts/03_derive_kinematic_features.py) and are noted as candidates for sensitivity checks in future work. The NOAA weather layer distributed with the dataset provides wind speed and gust (among other atmospheric variables) but no direct sea-state or wave-height measurement; sea_state is therefore computed as a standard 0–12 Beaufort scale category derived deterministically from wind speed (in knots), following the conventional Beaufort thresholds — a proxy whose operational relevance is supported by evidence that the relative contribution of wind- versus wave-induced speed loss shifts systematically across Beaufort levels (Kim et al., 2017). Because it is a monotonic transform of wind speed rather than an independently measured variable, sea_state is highly collinear with wind_speed_kt (Pearson r ≈ 0.95 in the analytic sample) and is reported in the correlation matrix for completeness but excluded from the regression models, which use wind_speed_kt directly.

Equation 1. For course-over-ground readings θi (i=1,…,n; in radians) within a vessel's rolling 30-minute window, the mean resultant length is R=1nicosθi2+isinθi2, and heading variability is reported as the circular standard deviation S=−2lnR, converted from radians to degrees.

3.1.4. Analytical Approach

1. Descriptive statistics of speed, heading variability, and dwell frequency, stratified by vessel type and weather condition category
2. Correlation analysis (Pearson/Spearman as appropriate) between weather variables and navigation behavior variables
3. Multiple regression (or ANCOVA) modeling navigation behavior as a function of weather variables, vessel type, and their interaction, to test moderation (RQ2), following the precedent of estimating weather effects on vessel speed separately by vessel category rather than in aggregate (Zhou et al., 2020)
4. Assumption checks: normality, homoscedasticity, multicollinearity (VIF) as applicable

3.1.5. Tools

Python (pandas, geopandas for shapefile handling, statsmodels/scipy for statistical testing, matplotlib/seaborn for visualization).

4. RESULTS

Results are based on the 294,100-observation analytic sample described in Section 3.2 (133 vessels; August, October, December 2017). Full statistical output is archived in outputs/ (table1_descriptive_stats.csv, table2_correlation_pearson.csv, table2_correlation_spearman.csv, table3_model_summary.csv, table_vif.csv, model1_speed_ols_summary.txt, model2_heading_ols_summary.txt, model3_stop_logit_summary.txt, figure1_stop_rate_by_type_wind.png, figure2_correlation_heatmap.png, figure3_speed_vs_wind.png, figure4_heading_vs_wind.png, figure5_predicted_prob_stop.png, figure6_spatial_activity_density.png).

4.1.1. Descriptive Statistics (Table 1, Figure 1)

Mean speed by vessel type and wind category (Calm < 5 kt, Moderate 5–15 kt, Strong 15–25 kt, Severe > 25 kt) is summarized below (mean [std], n); heading variability by vessel type and wind is examined via regression rather than a separate descriptive table (Section 4.3, Table 3b, Figure 4):

Table 1: Mean Speed in Knots [Std. Dev.], N, By Vessel Type and Wind Category (Calm < 5 Kt, Moderate 5–15 Kt, Strong 15–25 Kt, Severe > 25 Kt).

Vessel type	Calm	Moderate	Strong	Severe
Cargo	2.18 [3.86], n=3,547	2.42 [4.12], n=12,608	1.25 [3.38], n=2,379	–
Fishing	3.71 [3.61], n=4,068	2.90 [3.62], n=9,387	4.87 [3.21], n=1,651	–
Other	0.50 [1.89], n=18,880	0.41 [1.99], n=65,167	0.30 [1.45], n=14,579	–
Passenger	2.21 [4.78], n=30,921	2.44 [5.02], n=88,824	3.18 [5.76], n=13,406	11.63 [4.07], n=61
Tanker	2.59 [3.66], n=5,604	3.40 [3.85], n=20,767	2.88 [3.92], n=2,251	–

Two patterns stand out before any inferential test. First, the "Other" category — the catch-all for non-Cargo/Tanker/Passenger/Fishing ship type codes, which in practice includes tugs, pilot boats, and other harbor/service craft — has both the lowest mean speed (0.3–0.5 kt) and the highest stop rate (91–95%) of any category, and dominates the sample by row count (98,626 of 294,100 rows, 34%), consistent with such vessels spending most of their transmitting time moored or on station rather than transiting. Second, the "Severe" wind bin (>25 kt) contains only 61 observations, all from Passenger vessels underway at high mean speed (11.63 kt) with zero recorded stops — too small an n to generalize from, but plausibly reflecting that ferries either keep moving or suspend AIS-visible operation in the most severe conditions rather than idling at sea. Figure 1 displays the stop rate by vessel type across wind categories, making the "Other" category's persistently high stop rate, and Fishing's comparatively low and wind-sensitive stop rate, directly visible.

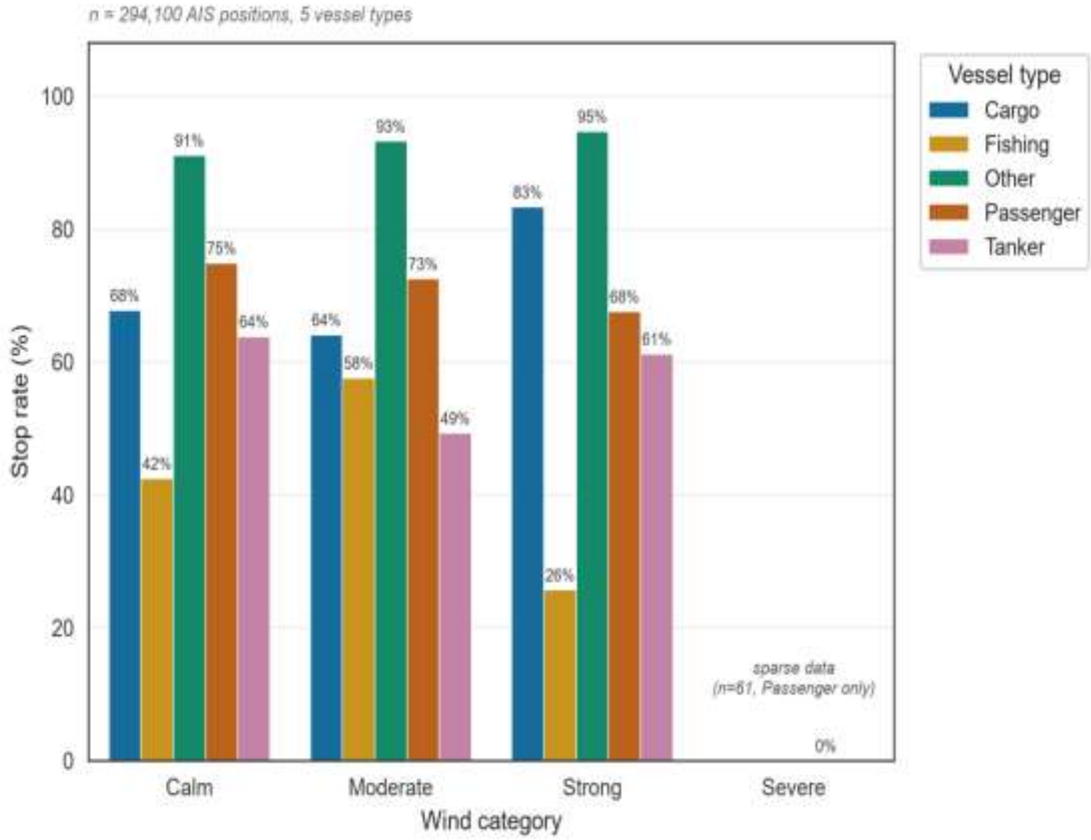

Figure 1: Stop Rate (%) By Vessel Type and Wind Category.

4.1.2. Correlation Analysis (Table 2, Figure 2)

At the zero-order (bivariate) level, wind speed was essentially uncorrelated with vessel speed (Pearson r = 0.008; Spearman ρ = 0.007) and weakly negatively correlated with heading variability (Pearson r = −0.078; Spearman ρ = −0.086). As expected, given its construction, sea_state correlated almost perfectly with wind_speed_kt (r = 0.952) and is not treated as an independent predictor (Section 3.3). is_stopped correlated strongly and negatively with speed_knots (Pearson r = −0.793; Spearman ρ = −0.840), which is expected by construction (stopped vessels have near-zero speed by definition) and serves mainly as a validity check on the derived variables rather than a substantive finding.

Table 2: Zero-Order Correlation Between Key Variables (Pearson R / Spearman P).

Variable pair	Pearson r	Spearman ρ
speed_knots – wind_speed_kt	0.008	0.007
heading_std – wind_speed_kt	-0.078	-0.086
sea_state – wind_speed_kt	0.952	0.966
is_stopped – speed_knots	-0.793	-0.840

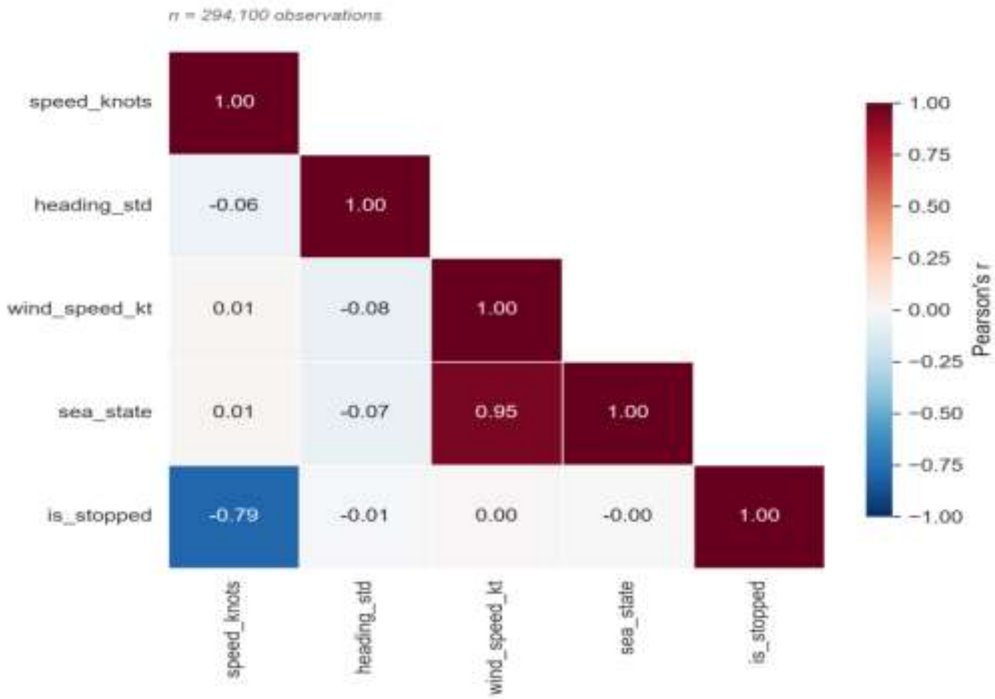

Figure 2: Pearson Correlation Matrix of Key Variables.

4.1.3. Regression Models (Equations 2–3, Tables 3a–3d, Figures 3–5)

Equation 2. Models 1 and 2 (OLS) share the specification $Yi=\beta0+\beta1\cdot windi+\beta2\cdot Fishingi+\beta3\cdot Otheri+\beta4\cdot Passengeri+\beta5\cdot Tankeri+\gamma2\cdot windi\times Fishingi+\gamma3\cdot windi\times Otheri+\gamma4\cdot windi\times Passengeri+\gamma5\cdot windi\times Tankeri+\varepsilon i$, where Yi is speed_knots (Model 1, Table 3a) or heading_std (Model 2, Table 3b), and Fishingi, Otheri, Passengeri, Tankeri are vessel-type dummy variables (reference category: Cargo).

Model 1 — speed~wind×vessel_type (OLS; specification in Equation 2, coefficients in Table 3a). $R^2 = 0.072$ ($F = 2{,}527$, $p < .001$). For the reference category (Cargo), speed decreased significantly with wind speed ($\beta = -0.077$ kt per knot of wind, $p < .001$). This slope differed significantly by vessel type: the wind×type interaction terms were positive and significant for Fishing (+0.118), Other (+0.067), Passenger (+0.144), and Tanker (+0.104), meaning the negative wind-speed relationship seen for Cargo was attenuated or reversed for every other category — net slopes were approximately flat for Other (−0.010), and mildly positive for Fishing (+0.041), Tanker (+0.027) and Passenger (+0.067). In other words, only Cargo vessels showed a clear "slow down in wind" pattern in this sample; the other categories did not. Figure 3 plots the speed–wind relationship separately for each vessel type, with a fitted trend line per panel; the near-flat, divergent per-type slopes visible there are consistent with the interaction coefficients above, while the strong separation in typical speed level across panels (Passenger vessels transiting near 15–20 kt versus "Other"/Cargo craft clustered near zero) is what the aggregate correlation in Table 2 obscures.

Table 3a: Model 1 (Speed~Wind×Vessel_Type, OLS) Coefficient Estimates. $R^2 = 0.072$, $F(9, 294090) = 2{,}527$, $P < .001$. Reference Category: Cargo.

Term	Coefficient	Std. error	p-value
Intercept	2.960	0.063	< .001
Vessel type: Fishing	0.021	0.091	.816
Vessel type: Other	-2.458	0.069	< .001
Vessel type: Passenger	-1.071	0.067	< .001
Vessel type: Tanker	0.014	0.083	.871
Wind speed (kt)	-0.077	0.006	< .001
Wind × Fishing	0.118	0.009	< .001
Wind × Other	0.067	0.006	< .001
Wind × Passenger	0.144	0.006	< .001
Wind × Tanker	0.104	0.008	< .001

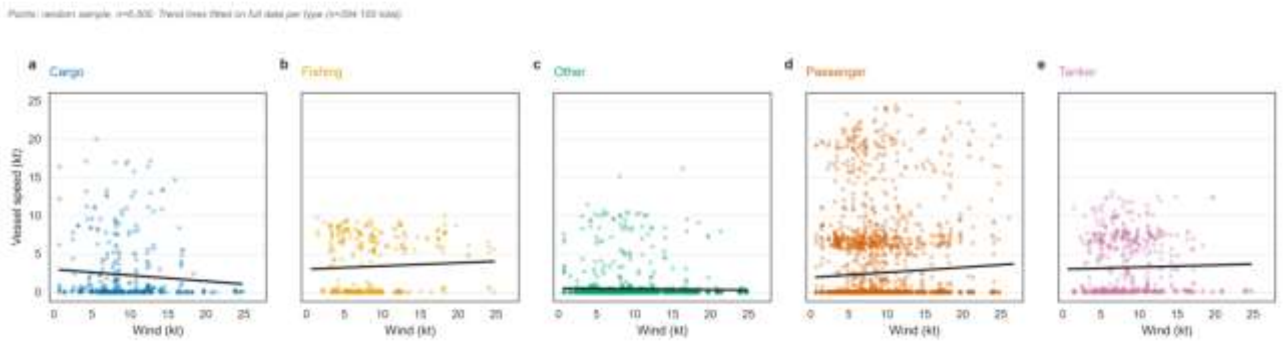

Figure 3: Vessel Speed Vs. Wind Speed, Faceted by Vessel Type (A-E), With Per-Type Fitted Regression Trend (Model 1).

Model 2 — heading_std~wind×vessel_type (OLS; same specification as Equation 2, coefficients in Table 3b). $R^2 = 0.013$ ($F = 446.4$, $p < .001$) — statistically significant given the large sample, but explaining very little variance. The Cargo reference slope was negative ($\beta = -0.55°$ per knot, $p < .001$); interactions were significant in both directions (Fishing +1.13, net positive ≈ +0.57; Other −0.61, net negative ≈ −1.17; Passenger +0.22, net negative ≈ −0.34; Tanker +0.42, net negative ≈ −0.13). No consistent direction of wind effect on course stability emerged across vessel types. Figure 4 plots heading variability against wind speed separately for each vessel type: the near-flat, shallow and inconsistently signed slopes visible across panels are a direct visual counterpart to the weak, direction-inconsistent interaction terms reported above.

Table 3b: Model 2 (Heading_Std~Wind×Vessel_Type, OLS) Coefficient Estimates. $R^2 = 0.013$, $F(9, 294090) = 446.4$, $P < .001$. Reference Category: Cargo.

Term	Coefficient	Std. error	p-value
Intercept	42.725	0.593	< .001
Vessel type: Fishing	-14.887	0.850	< .001
Vessel type: Other	7.931	0.643	< .001
Vessel type: Passenger	0.696	0.628	.268
Vessel type: Tanker	2.545	0.775	.001
Wind speed (kt)	-0.554	0.055	< .001
Wind × Fishing	1.126	0.083	< .001
Wind × Other	-0.613	0.060	< .001
Wind × Passenger	0.218	0.059	< .001
Wind × Tanker	0.425	0.076	< .001

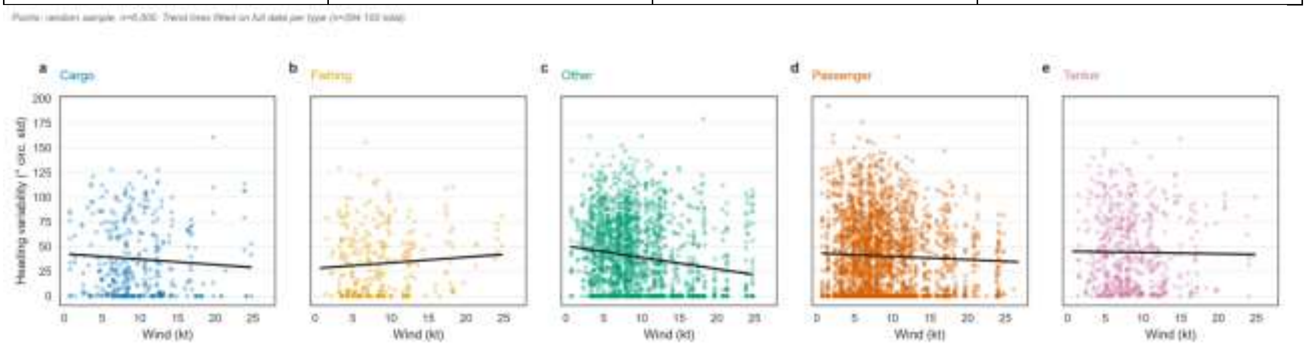

Figure 4: Heading Variability Vs. Wind Speed, Faceted by Vessel Type (A-E), With Per-Type Fitted Regression Trend (Model 2).

Equation 3. Model 3 (logistic regression, Table 3c):

logitPis_stoppedi=1=lnPis_stoppedi=11−Pis_stoppedi=1=β0+β1·windi+β2·Fishingi+β3·Otheri+β4·Passengeri+β5·Tankeri, with the same dummy coding and reference category (Cargo) as Equation 2.

Model 3 — Pstopped~wind+vessel_type (logistic; specification in Equation 3, coefficients in Table 3c). Pseudo-$R^2 = 0.104$ (LLR $p < .001$). Wind speed was a significant, small negative predictor of stopping ($\beta = -0.0116$ per knot, $p < .001$; odds ratio ≈ 0.988 per knot), i.e., vessels were very slightly less likely to be in a sustained stop as wind increased, controlling for type — consistent with wind discouraging idle drifting rather than encouraging it.

Vessel type was a much stronger predictor: relative to Cargo, the odds of being stopped were sharply higher for Other (β = +1.87, OR ≈ 6.5) and moderately higher for Passenger (β = +0.24, OR ≈ 1.27), and lower for Fishing (β = −0.73, OR ≈ 0.48) and Tanker (β = −0.61, OR ≈ 0.54). Figure 5 translates these coefficients into predicted probabilities across the observed wind range: the five vessel-type curves are near-parallel and clearly separated, while each individual curve's negative slope illustrates the small, counter-intuitive wind effect reported above.

Table 3c: Model 3 (Pstopped~Wind+Vessel_Type, Logistic) Coefficient Estimates. Pseudo-R^2 = 0.104, LLR P < .001. Reference Category: Cargo.

Term	Coefficient	Std. error	p-value	Odds ratio
Intercept	0.831	0.018	< .001	2.30
Vessel type: Fishing	-0.731	0.023	< .001	0.48
Vessel type: Other	1.872	0.020	< .001	6.50
Vessel type: Passenger	0.241	0.017	< .001	1.27
Vessel type: Tanker	-0.610	0.020	< .001	0.54
Wind speed (kt)	-0.012	0.001	< .001	0.988

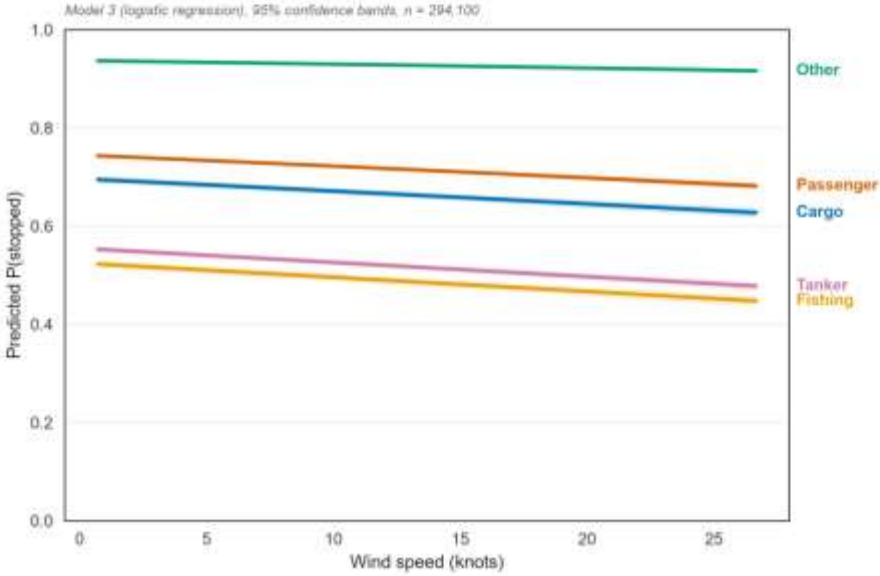

Figure 5: Predicted Probability of a Sustained Stop by Wind Speed and Vessel Type (Model 3).

Model diagnostics (Table 3d). Variance inflation factors for the main-effects predictors shared across all three models were all below the conventional threshold of 5 (wind speed VIF = 1.01; vessel-type dummies ranged from 1.72 to 4.49), indicating no problematic multicollinearity beyond the already-noted, and separately excluded, wind_speed_kt/sea_state relationship (Section 3.3). Residual diagnostics for the two OLS models (Models 1–2) indicated significant non-normality (Jarque–Bera p < .001 for both), consistent with the right-skewed, non-negative nature of speed and heading-variability data rather than a modeling artifact, and mild positive residual autocorrelation (Durbin–Watson ≈ 1.79–1.80 for both models), plausibly reflecting the resampled, within-vessel time-series structure of the data (Section 3.2). Given the very large sample size (n = 294,100), these violations are expected to inflate the precision (understate the standard errors) of the coefficient estimates somewhat rather than bias their sign or approximate magnitude, and do not alter the substantive pattern of results; robustness to cluster-robust (by-vessel) standard errors is noted as a candidate refinement in Section 6.

Table 3d: Variance Inflation Factors (VIF) For the Main-Effects Predictors (Wind Speed and Vessel-Type Dummies) Shared Across Models 1-3.

Predictor	VIF
Wind speed (kt)	1.01
Vessel type: Fishing	1.72
Vessel type: Other	4.20

Vessel type: Passenger	4.49
Vessel type: Tanker	2.30

4.1.4. Spatial Distribution of Vessel Activity (Figure 6)

To contextualize the preceding statistical results geographically, Figure 6 maps the spatial density of all AIS positions in the analytic sample (left) against the subset flagged as stopped (right), using the port boundary and Saronic Gulf coastline shapefiles distributed with the dataset (Tritsarolis et al., 2022) for reference. All positions concentrate along a narrow corridor funneling into the Piraeus port entrance, consistent with a single dominant approach channel. Stopped positions, by contrast, are markedly more dispersed across the surrounding gulf rather than concentrated inside the port polygon itself, forming diffuse clusters offshore and along the coastline that are consistent with informal anchoring or waiting areas rather than berthing. This spatial pattern is consistent with the anchorage-driven interpretation of stop behavior proposed in Section 5, and motivates the anchorage-specific follow-up analysis suggested there and by Tritsarolis et al. (2021).

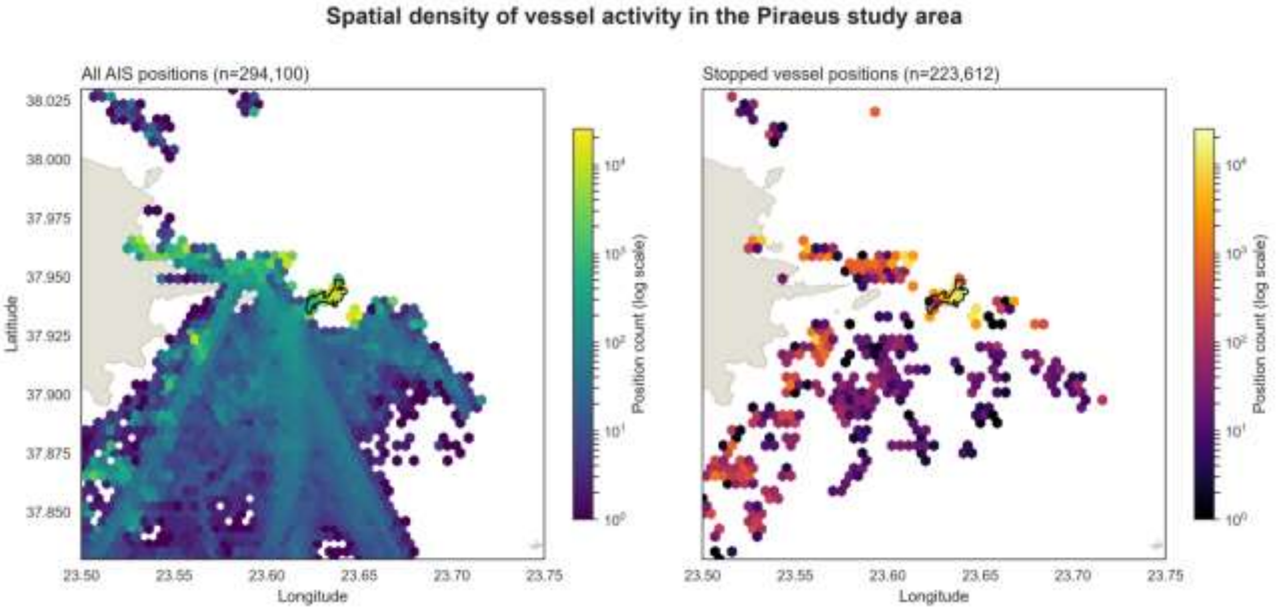

Figure 6: Spatial Density of All AIS Positions (Left) Versus Stopped Vessel Positions (Right), Piraeus Study Area.

4.1.5. Summary By Research Question

5. RQ1 (does wind associate with speed/heading changes?): Yes, but the effects are small in magnitude and explain little variance ($R^2 = 0.07$ and 0.01 respectively) despite being statistically significant at this sample size — a pattern of small-but-significant explanatory power consistent with other AIS-based studies of weather–speed relationships (Dreyer, 2021).

6. RQ2 (does vessel type moderate the wind relationship?): Yes, strongly — the wind×type interaction was significant in both OLS models, and the sign/magnitude of the wind effect differed by category, with Cargo the only type showing a clear slow-down pattern.

7. RQ3 (does wind predict stopping frequency?): Yes, in a small-but-significant, and counter-to-naive-expectation, negative direction; vessel type (especially the harbor/service "Other" category) is a far stronger predictor of stopping than wind.

5. DISCUSSION

The central finding of this study is not that weather is irrelevant to vessel behavior near Piraeus, but that its measurable effect at the resolution of this dataset is small, type-dependent, and in one respect (stopping behavior) counter to a naive "bad weather causes vessels to slow down and wait" narrative — broadly consistent with prior AIS-based evidence that wind measurably affects vessel speed while heading/course is comparatively more driven by other factors such as currents and traffic encounters (Shu et al., 2017). Two results are worth unpacking in light of maritime operations practice.

First, the wind–speed relationship (Model 1) was negative only for Cargo vessels; Passenger, Tanker and Fishing vessels showed flat or mildly positive slopes.

A plausible operational explanation is that Piraeus-area traffic near the port approach is dominated by short-haul, schedule-driven categories (ferries, coastal tankers, local fishing) whose masters may prioritize holding schedule or reaching shelter over speed reduction, whereas cargo vessels on longer approach legs have more slack to trade speed for fuel economy or comfort in worse conditions. This is consistent with, though not a direct test of, the weather-routing literature's emphasis on voyage-economics-driven speed adjustment, which this descriptive dataset cannot separate from scheduling constraints or from the classic distinction between involuntary speed loss (added resistance in a seaway) and voluntary speed reduction by the master (Prpić-Oršić and Faltinsen, 2012). In absolute terms, the magnitude of the Cargo slowdown found here is modest relative to physically estimated speed losses in a seaway (on the order of ~1 knot in head seas at Beaufort 6; Kim et al., 2017), consistent with the small effect size observed ($R^2 = 0.072$).

Second, the negative (rather than positive) association between wind speed and the probability of a sustained stop (Model 3) is the most counter-intuitive result. One interpretation is a survivorship/selection effect in AIS-visible behavior: vessels genuinely idling or drifting at anchor in calm conditions are common and safely observable, whereas in stronger wind vessels at anchor may need to actively maneuver (dragging anchor, holding station) rather than sit still long enough to register as "stopped" under our 10-minute sustained-low-speed definition — or masters may simply avoid anchoring in worse conditions and instead keep way on. This is a hypothesis the current descriptive design cannot adjudicate and would require finer-grained anchorage-specific analysis (cf. the anchorage-detection approach of Tritsarolis et al., 2021) to test directly. The strong differentiation of stopping behavior by vessel type found here (Model 3) parallels AIS-based evidence of systematic differences in vessel waiting times by type at anchorage (Franzkeit et al., 2020), reinforcing vessel type as a first-order determinant of dwell behavior rather than a secondary control.

Third, the outsized role of the "Other" vessel-type category — both in row count and in stop probability — is a data-composition artifact worth flagging for practitioners reusing this dataset: it functions as a residual bucket for non-standard AIS shiptype codes and, in this sample, behaves like a population of largely stationary harbor/service craft rather than transiting traffic. Analyses that do not separate it from genuine "unclassified transiting vessel" cases risk conflating two very different behavioral populations.

For port traffic management, the practical implication is that a single, uniform weather-response rule (e.g., "increase safety margins uniformly for all vessels above wind threshold X") is not well supported by this data — the relationship between wind and both speed and stopping behavior is heterogeneous enough across vessel categories (cf. Zhou et al., 2020) that type-specific guidance would better reflect observed behavior than a one-size-fits-all rule, at least for the wind ranges and vessel mix observed in this three-month, single-port sample. This echoes prior evidence that speed variability itself is spatially linked to safety-relevant zones in busy port waters (Zhang et al., 2019), reinforcing the practical value of AIS-based, type-differentiated monitoring.

6. LIMITATIONS

1. Single-port, three-month context. Results are drawn from Piraeus, Greece, and from August, October and December 2017 only (chosen to span summer meltemi, autumn-transition and winter-storm wind regimes); generalizability to other ports, vessel mixes, or years is untested.

2. Stratified vessel sampling, not a full census. 133 of ~500 randomly-sampled candidate vessels (100 per type category) actually appeared in the three selected months' dynamic data; the analysis does not cover the dataset's full ~6,230-vessel static registry or its full 2017–2019 span (see Section 3.2).

3. No independent sea-state measurement. The dataset provides wind speed/gust but no wave height or sea-state observation; sea_state is a deterministic Beaufort-scale transform of wind speed ($r \approx 0.95$ with wind_speed_kt) and was therefore excluded from the regression models to avoid reporting a collinear variable as if it were independent evidence. Relatedly, AIS-derived speed alone cannot distinguish involuntary speed loss (added resistance in a seaway) from voluntary speed reduction by the vessel's master (Prpić-Oršić and Faltinsen, 2012), a distinction this descriptive design does not attempt to resolve.

4. Weather variables are model/observation-grid derived (NOAA), not vessel-local in-situ readings, and were matched to each AIS position via nearest-station/nearest-time joins (~45 stations, ~3-hour cadence), which introduces some spatial and temporal imprecision — a vessel's true local wind may differ from its matched station reading, especially near the temporal midpoint between readings, a limitation also noted when coupling large-scale AIS data with gridded weather/hindcast sources (Vitali et al., 2020).

5. 1-minute trajectory resampling and fixed derivation thresholds. Heading variability (30-minute rolling circular std) and stop detection (>1.0 kt sustained <10 min) use single implementation defaults rather than a range tested via sensitivity analysis; results could shift somewhat under different threshold choices.

6. Non-independent observations and non-normal residuals in the OLS models. Residual diagnostics (Section 4.3, Table 3d) show significant non-normality (Jarque–Bera $p < .001$) and mild positive autocorrelation (Durbin–Watson ≈ 1.79–1.80) for Models 1–2, consistent with repeated, within-vessel time-series observations rather than independent draws; conventional OLS standard errors likely understate true uncertainty somewhat, and cluster-robust (by-vessel) standard errors are a natural refinement for future work, though the large sample size makes this unlikely to change the substantive conclusions.

7. "Other" vessel-type category is heterogeneous. As discussed in Section 5, it aggregates all non-Cargo/Tanker/Passenger/Fishing AIS shiptype codes and, empirically, behaves like a largely stationary harbor/service-craft population; it should not be read as "unclassified transiting vessels."

8. AIS data reflects declared/transmitted vessel type and position, which may not always be accurate — static fields such as declared vessel type have historically shown substantially higher error rates than dynamic fields such as speed and position, the latter reported accurate to within roughly 1% (Harati-Mokhtari et al., 2007) — and vessel identifiers in this dataset are anonymized/hashed, preventing external validation against vessel registries.

9. Descriptive-correlational design does not establish causality. Observed associations between wind and behavior may partly reflect confounds (season, time of day, traffic scheduling) not modeled here.

7. CONCLUSION

Using a stratified, three-month subsample of the Piraeus AIS Dataset (294,100 vessel-minute observations from 133 vessels), this study found that wind speed is a statistically significant but modest and vessel-type-dependent predictor of navigation behavior near the Port of Piraeus. Wind explained little variance in speed ($R^2 = 0.072$) or heading stability ($R^2 = 0.013$) on its own, but its effect differed meaningfully across vessel categories (RQ2), with only Cargo vessels showing a consistent slow-down pattern under stronger wind. Contrary to a simple "bad weather → more waiting" expectation, stronger wind was associated with a slightly lower, not higher, probability of sustained stopping (RQ3), while vessel type — particularly the harbor/service "Other" category — was a far stronger predictor of stop behavior than weather itself. These findings support treating vessel type as a first-order moderator, not a control variable, in future weather-behavior analyses of AIS data, and suggest that practical, weather-responsive port traffic guidance should be differentiated by vessel category rather than applied uniformly. Future work should extend the time window across full seasons and years, test sensitivity to the heading/stop-detection thresholds used here, and investigate the anomalous stop–wind relationship directly (e.g., via anchorage-specific analysis) to distinguish genuine behavioral response from AIS-visibility artifacts.

Author Contributions: The author was solely responsible for conceptualization, data curation, formal analysis, methodology, and writing of this article.

REFERENCES

Capobianco, S., Millefiori, L. M., Forti, N., Braca, P. and Willett, P. (2021) Deep learning methods for vessel trajectory prediction based on recurrent neural networks. IEEE Transactions on Aerospace and Electronic Systems, Vol. 57, No. 6, 4329–4346.

Dreyer, L. O. (2021) Safe speed for Maritime Autonomous Surface Ships – The use of Automatic Identification System data. Proceedings of the 31st European Safety and Reliability Conference (ESREL 2021), Angers, France.

Franzkeit, J., Pache, H. and Jahn, C. (2020) Investigation of vessel waiting times using AIS data. In Freitag, M., Haasis, H.-D., Kotzab, H. and Pannek, J. (Eds.) Dynamics in Logistics: Proceedings of the 7th International Conference LDIC 2020, Lecture Notes in Logistics, 70–78. Cham, Springer.

Harati-Mokhtari, A., Wall, A., Brooks, P. and Wang, J. (2007) Automatic Identification System (AIS): Data reliability and human error implications. Journal of Navigation, Vol. 60, No. 3, 373–389.

Kim, M., Hizir, O., Turan, O., Day, S. and Incecik, A. (2017) Estimation of added resistance and ship speed loss in a seaway. Ocean Engineering, Vol. 141, 465–476.

Mardia, K. V. and Jupp, P. E. (2000) Directional Statistics. Chichester, Wiley.

Pallotta, G., Vespe, M. and Bryan, K. (2013) Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction. Entropy, Vol. 15, No. 6, 2218–2245.

Poupkou, A., Zanis, P., Nastos, P., Papanastasiou, D., Melas, D., Tourpali, K. and Zerefos, C. (2011) Present climate trend analysis of the Etesian winds in the Aegean Sea. Theoretical and Applied Climatology, Vol. 106, No. 3–4, 459–472.

Prpić-Oršić, J. and Faltinsen, O. M. (2012) Estimation of ship speed loss and associated CO2 emissions in a seaway. Ocean Engineering, Vol. 44, 1–10.

Riveiro, M., Pallotta, G. and Vespe, M. (2018) Maritime anomaly detection: A review. WIREs Data Mining and Knowledge Discovery, Vol. 8, No. 5, e1266.

Shu, Y., Daamen, W., Ligteringen, H. and Hoogendoorn, S. P. (2017) Influence of external conditions and vessel encounters on vessel behavior in ports and waterways using Automatic Identification System data. Ocean Engineering, Vol. 131, 1–14.

Tritsarolis, A., Chondrodima, E., Pelekis, N. and Theodoridis, Y. (2022) Vessel Collision Risk Assessment using AIS Data: A Machine Learning Approach. Proceedings of the 23rd IEEE International Conference on Mobile Data Management (MDM), Paphos, Cyprus.

Tritsarolis, A., Kontoulis, Y. and Theodoridis, Y. (2022) The Piraeus AIS dataset for large-scale maritime data analytics. Data in Brief, Vol. 40, 107782.

Tritsarolis, A., Kontoulis, Y., Pelekis, N. and Theodoridis, Y. (2021) MaSEC: Discovering anchorages and co-movement patterns on streaming vessel trajectories. Proceedings of the 17th International Symposium on Spatial and Temporal Databases (SSTD).

Tyrlis, E. and Lelieveld, J. (2013) Climatology and dynamics of the summer Etesian winds over the eastern Mediterranean. Journal of the Atmospheric Sciences, Vol. 70, No. 11, 3374–3396.

Vitali, N., Prpić-Oršić, J. and Guedes Soares, C. (2020) Coupling voyage and weather data to estimate speed loss of container ships in realistic conditions. Ocean Engineering, Vol. 210, 106758.

Yang, D., Wu, L., Wang, S., Jia, H. and Li, K. X. (2019) How big data enriches maritime research – a critical review of Automatic Identification System (AIS) data applications. Transport Reviews, Vol. 39, No. 6, 755–773.

Zhang, L., Meng, Q. and Fwa, T. F. (2019) Big AIS data based spatial-temporal analyses of ship traffic in Singapore port waters. Transportation Research Part E: Logistics and Transportation Review, Vol. 129, 287–304.

Zhou, Y., Daamen, W., Vellinga, T. and Hoogendoorn, S. P. (2020) Impacts of wind and current on ship behavior in ports and waterways: A quantitative analysis based on AIS data. Ocean Engineering, Vol. 213, 107774.

Zis, T. P. V., Psaraftis, H. N. and Ding, L. (2020) Ship weather routing: A taxonomy and survey. Ocean Engineering, Vol. 213, 107697.