Marijana Marjanović[1,], Jasna Prpić-Oršić[1], Marko Valčić[1,2]*

Propagation of weather forecast uncertainties through attainable ship speed prediction models

[1] University of Rijeka, Faculty of engineering, Vukovarska 58, Rijeka, 51000, Croatia
[2] University of Zadar, Maritime Department, Ruđera Boškovića 5, Zadar, 23000, Croatia
[*] Corresponding author: marijana.marjanovic@riteh.uniri.hr

Original scientific paper
Received: January 25, 2026
Accepted: April 16, 2026
https://doi.org/10.65776/ep.20.4.4

Abstract

Accurate prediction of weather-induced ship speed loss requires understanding how meteorological forecast uncertainties propagate through vessel performance models. This study presents a comparative assessment of three computational approaches for attainable speed estimation: the Wärtsilä NTPRO 5000 navigation simulator with JONSWAP and Pierson-Moskowitz spectral implementations, and HydroComp NavCad hydrodynamic software. Over 2,000 simulations were conducted for a 28,050 DWT bulk carrier across varying wave heights (0-12 m), encounter angles (0°-180°), and operational speeds (12.0-14.5 knots). Correlation analysis between meteorological predictor variables and ship speed response uncertainties revealed that significant wave height errors exhibit the strongest coupling with speed prediction errors (r = 0.65-0.97), while encounter angle geometry substantially modulates uncertainty propagation patterns. The Pierson-Moskowitz implementation demonstrated the most balanced error characteristics, whereas JONSWAP produced more polarised correlations at extended forecast horizons. Results indicate that uncertainty quantification approaches must account for both model-specific sensitivities and spatially varying forecast skill degradation.

***Keywords**: Weather forecast degradation; Attainable speed modelling; Uncertainty quantification; Correlation analysis.*

1. Introduction

The attainable ship speed under varying weather conditions has been studied using empirical and semi-empirical methods [1,2], physics-based approaches [3,4], and, more recently, data-driven and intelligent models [7,8]. These methods have enabled increasingly accurate speed prediction and demonstrated the potential for substantial fuel savings and emission reductions through weather-aware optimisation [5,6]. Multi-objective frameworks [9], deep reinforcement learning for routing [10], and adaptive control strategies [11] further extend these capabilities. However, most of these approaches remain deterministic and do not account for the propagation of forecast errors through vessel response models.Although uncertainty sources in ship routing have been extensively reviewed [12], existing probabilistic frameworks are typically first-order and rely on linearization [13], which is inadequate for highly nonlinear ship responses. Links between ensemble forecast uncertainty and fuel consumption have been explored [14], yet often with simplified resistance models. Uncertainty has also been recognised in collision avoidance and trajectory prediction [15], underscoring its relevance across maritime operations. While forecast errors are known to grow non-linearly with lead time [5], their transformation through ship performance models remains insufficiently characterised. Most existing frameworks still treat weather forecasting and ship performance as independent components, neglecting the coupled interactions that govern voyage-level uncertainty. This study addresses these limitations by systematically quantifying how meteorological forecast uncertainties propagate through three distinct ship performance modelling approaches. The primary contributions include: (i) comparative uncertainty analysis across Wärtsilä NTPRO 5000 navigation simulator implementations utilizing both JONSWAP and Pierson-Moskowitz spectral formulations, alongside HydroComp NavCad hydrodynamic software; (ii) correlation analysis establishing quantitative relationships between predictor variable (meteorological) and response variable (ship speed) uncertainties across varying sea states and encounter geometries; and (iii) practical demonstration of the framework through a transatlantic case study incorporating estimated time of arrival uncertainty quantification under realistic winter North Atlantic conditions.

2. Ship Performance Simulation Framework

Quantifying attainable ship speed under stochastic weather conditions requires reliable computational tools that capture the interactions between ship hydrodynamics and environmental forces. This study employs three distinct modelling approaches to generate comprehensive speed-loss predictions, enabling a comparative assessment of uncertainty-propagation characteristics.

2.1. Data collection using simulations

Two complementary computational environments were employed for attainable ship speed prediction, representing fundamentally different modelling paradigms.

The Wärtsilä NaviTrainer NTPRO 5000 full-mission bridge simulator served as the primary experimental platform. This DNV Class A-certified system solves vessel motion equations in real time across six degrees of freedom, capturing complex interactions among hull hydrodynamics,

propulsion forces, and environmental disturbances [16,17]. Hydrodynamic forces are decomposed into positional and damping components, determined from experimentally derived coefficients obtained from tank tests. Environmental loading is computed using distinct models for wind and wave effects, with wave-induced forces comprising first-order oscillatory components and second-order mean drift forces that contribute to steady speed loss. Two spectral formulations were implemented: the Pierson-Moskowitz spectrum for fully developed seas and the JONSWAP spectrum for fetch-limited conditions [18].

Throughout simulations, an autopilot system maintained commanded heading rather than track-keeping mode, while speed loss emerged naturally from the force balance between thrust and environmental loading. This configuration was chosen deliberately: the lookup tables are intended to characterise the intrinsic wind–wave-induced speed loss for each encounter geometry, isolated from the additional rudder-activity losses that occur when a vessel compensates for drift along a predefined track. In operational routing, such track-keeping losses and drift-induced heading corrections can be superimposed on the lookup-table values as separate corrections at the routing stage.

Parallel simulations utilised HydroComp NavCad, employing quasi-static resistance decomposition methods based on ITTC-1978 correlation procedures [19, 20]. Unlike the time-domain NTPRO approach, NavCad separates total resistance into distinct physical components such as bare-hull viscous resistance, wave-making resistance, appendage drag, and environmental added resistance, which are individually computed and summed. Wave-added resistance predictions employ regression-based methods derived from systematic model test series, with the Aertssen approach [21] enabling direct speed-loss estimation for weather-routing applications.

2.2. Reference vessel and simulation matrix

A 28,050 DWT bulk carrier served as the reference vessel for all simulations, selected for its representative characteristics among medium-sized commercial cargo ships operating transatlantic routes [22]. The principal particulars are summarised in Table 1.

Table 1. Principal particulars of the reference bulk carrier

Parameter	Value	Unit
Length between perpendiculars (LPP)	160.40	m
Beam (B)	27.20	m
Depth (D)	13.60	m
Design draft (T)	9.819	m
Deadweight (DWT)	28,050	tons
Gross tonnage (GT)	17,009	-
Main engine power (MCR)	6,150	kW
Propeller diameter	5.25	m
Service speed	~14.0	knots

The propulsion system consists of a two-stroke marine diesel engine rated at 6,150 kW at 136 rpm under nominal conditions, derated to 5,850 kW at 129 rpm for heavy fuel oil operation. Power transmission is achieved through a direct-drive shaft system connected to a four-bladed fixed-pitch propeller with a mean pitch of 3.686 m and a 35-degree skew angle optimised for bulk carrier operations. These specifications were integrated into both computational platforms to ensure consistent representation of vessel response characteristics.

The simulation matrix was designed to comprehensively cover operationally relevant conditions encountered on North Atlantic routes. Environmental parameters included:

(*i*) Significant wave heights: 13 wave heights spanning $H_S \in \{0, 1, 2, ..., 12\}$ metres, corresponding to WMO Sea State Codes 0-8

(*ii*) Wave encounter angles: 13 orientations $\alpha_{waves} \in \{0°, 15°, 30°, ..., 180°\}$ covering head seas (0°), bow-quartering (15°-75°), beam seas (90°), stern-quartering (105°-165°), and following seas (180°)

(*iii*) Intended reference ship speeds: $V_{ref.} \in \{12.0, 13.5, 14.5\}$ knots

(*iv*) Loading conditions: Full load and ballast configurations

(*v*) Wave spectra: Pierson-Moskowitz and JONSWAP.

Wind speed was not varied as an independent parameter but was coupled to significant wave height through the Beaufort–Douglas sea-state correspondence, consistent with the fully developed sea assumption of the Pierson–Moskowitz spectrum and the wind–wave coupling intrinsic to the JONSWAP formulation. Wind direction was kept aligned with the wave direction, reflecting typical North Atlantic conditions in which swell and wind sea co-propagate; this simplification is the same as that adopted in the underlying simulation framework [23]. This parameterisation yielded 2,028 unique simulation scenarios for NTPRO 5000 (accounting for both spectral formulations and loading conditions) and 1,014 scenarios for NavCad (single spectral treatment) [23]. While simulations were conducted for encounter angles 0°-180°, results were extended to the full 360° range by applying symmetry principles, as port and starboard wave encounters produce mirror-image responses.

2.3. Attainable Speed Formulation

Neglecting ocean current effects, the attainable ship speed $V_{att.}$ can be expressed as a function of intended reference speed and prevailing sea state conditions

$$V_{\text{att.}} = f\left(V_{\text{ref.}}, H_S, T_p, \alpha_{\text{waves}}\right) \quad (1)$$

where $V_{ref.}$ denotes the intended reference speed (knots), H_S represents significant wave height (m), T_p indicates peak wave period (s), and α_{waves} specifies the encounter

wave angle (degrees). The functional relationship is determined empirically, with results organised into lookup tables that enable rapid interpolation for arbitrary input combinations.

The sea conditions can be actual, for actual values of $H_S^{(act.)}$, $T_P^{(act.)}$ and $\alpha_{waves}^{(act.)}$, and forecasted, for forecast values of $H_S^{(for.)}$, $T_P^{(for.)}$ and $\alpha_{waves}^{(act.)}$. The attainable ship speed $V_{att.,actual}$ in actual conditions is

$$V_{att.,actual} = f\left(V_{ref.}, H_S^{(act.)}, T_P^{(act.)}, \alpha_{waves}^{(act.)}\right) \quad (2)$$

i.e. the predicted attainable ship speed $V_{att.,predicted}$ in forecasted conditions is

$$V_{att.,predicted} = f\left(V_{ref.}, H_S^{(for.)}, T_P^{(for.)}, \alpha_{waves}^{(for.)}\right) \quad (3)$$

For operational implementation, attainable speed values are obtained through bilinear interpolation between discrete simulation points, ensuring smooth transitions for intermediate wave heights and encounter angles. The encounter wave angle is computed from ship heading ψ and meteorological wave direction β_{waves} according to

$$\alpha_{waves} = \begin{cases} \beta_{waves} - \psi, & for\ \psi \le \beta_{waves} \\ 2\pi + \beta_{waves} - \psi, & for\ \psi > \beta_{waves} \end{cases}. \quad (4)$$

While the theoretical formulation for attainable ship speed includes peak wave period T_p as an independent variable, practical implementation couples period to significant wave height through the spectral relationships. In the Pierson-Moskowitz formulation, peak period is not independently controllable but rather a derived quantity determined by the spectrum parameterisation. Similarly, JONSWAP implementation maintains period-height coupling through wind-wave relationships. This simplification reflects the physical reality that wave height and period evolve together according to fetch, duration, and wind speed characteristics [18], and is justified for the North Atlantic routes studied (e.g., Rotterdam–New York), where swell and wind seas typically align.

Simulation outputs occasionally yielded undefined (NaN) values representing sea states where the autopilot system could no longer maintain the demanded course, i.e. conditions characterised by excessive yaw rates, severe roll and pitch amplitudes, or loss of directional stability. These values effectively define operational boundaries beyond which navigation becomes hazardous, serving as indicators for route optimisation algorithms to identify avoidance zones or mandate course and speed alterations.

3. Weather Forecast Uncertainty

The meteorological dataset was organised into four forecast lead-time classes, each aligned with a distinct level of operational decision-making. The short-range window (0–24 h) represents immediate onboard planning and was sampled at 0, 6, 12, 18, and 24 hours. Tactical voyage adjustments are supported by the medium-range window (24–72 h), comprising forecasts at 30, 36, 42, 48, 54, 60, 66, and 72 hours. The extended-range horizon (72–120 h) addresses strategic routing decisions and includes predictions at 78, 84, 90, 96, 102, 108, 114, and 120 hours. Finally, the long-range window (120–168 h) provides advance planning capability through outputs at 126, 132, 138, 144, 150, 156, 162, and 168 hours. This structured temporal partitioning enables a systematic evaluation of forecast skill decay with increasing lead time, which is essential for analysing how uncertainty propagates into attainable ship speed estimates.

For each lead-time class, both forecasted and observed values were collected for six meteorological variables directly influencing ship performance: significant wave height H_s (m), representing the mean height of the highest one-third of waves; peak wave period T_p (s), characterising the dominant spectral component; wind speed V_{wind} (m/s) at the 10 m reference level; wave direction β_{waves} (deg), defined as the direction from which waves propagate relative to geographic north; wind direction β_{wind} (deg), defined analogously for the atmospheric flow; and the encounter angle α_{waves} (deg), describing the relative orientation between vessel heading and incoming wave direction.

All data processing and uncertainty analysis were implemented within a Python-based workflow tailored for large-scale meteorological datasets, employing specialised libraries (cfgrib and eccodes for GRIB decoding, xarray and pandas for spatio-temporal alignment, and NumPy, SciPy and scikit-learn for statistical evaluation) within Python 3.13.

4. Uncertainty metrics

The Root Mean Square Error (RMSE) [24], can be defined as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(F_i - O_i)^2}. \quad (5)$$

The quadratic form of the RMSE places a disproportionately high weight on large deviations, rendering the metric especially responsive to sporadic but severe errors. This property is advantageous for revealing extreme discrepancies in predicted ship speed that may have a decisive influence on routing outcomes and schedule reliability.

In contrast, the Mean Absolute Error (MAE) offers a linear measure of the average error magnitude [24] and is defined as

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|F_i - O_i|. \quad (6)$$

Bias quantifies systematic forecast tendencies, which reveals if a model consistently over- or under-predicts, and is noted as [25]

$$\text{Bias} = \frac{1}{n}\sum_{i=1}^{n}(F_i - O_i)\,. \tag{7}$$

Unlike RMSE and MAE, the bias may approach zero even in the presence of substantial errors, provided that positive and negative deviations compensate each other. Consequently, bias must be evaluated in conjunction with magnitude-based metrics in order to discriminate between true predictive accuracy and mere cancellation of opposing errors [25].

The Index of Agreement (Willmott's Index) [26] is a normalised metric ranging from 0 to 1 that quantifies the degree of correspondence between model predictions and observations relative to their variability. It is defined as

$$\text{IoA} = 1 - \frac{\sum_{i=1}^{n}(F_i - O_i)^2}{\sum_{i=1}^{n}(|F_i - \bar{O}| + |O_i - \bar{O}|)^2}. \tag{8}$$

The Fractions Skill Score (FSS) is a spatial verification metric for high-resolution forecasts of threshold-based events [27]. Ranging from 0 (no skill) to 1 (perfect agreement), it evaluates forecast performance by comparing the fractional occurrence of an event within local neighbourhoods rather than at individual grid points. Forecast and observed fields are first binarised, after which a moving window computes the local event fractions in both fields. The FSS is then defined as [27]

$$\text{FSS} = 1 - \frac{\text{MSE}_f}{\text{MSE}_{f,ref}} = 1 - \frac{\sum_{n=1}^{N}[O(n)-F(n)]^2}{\sum_{n=1}^{N}[O^2(n)+F^2(n)]} \tag{9}$$

where MSE_f denotes the mean squared error between the forecast fractions $F(n)$ and the observed fractions $O(n)$ for neighbourhoods $n = 1,\ldots,N$, and $\text{MSE}_{f,ref}$ is the reference MSE corresponding to a no-skill forecast, i.e. the worst-case scenario. Here, N denotes the total number of neighbourhoods.

The Continuous Ranked Probability Score (CRPS) generalises forecast verification to probabilistic predictions by evaluating the agreement between a forecast distribution and the corresponding observed value [28]

$$\text{CRPS} = \sigma\left[z(2\Phi(z) - 1) + 2\phi(z) - \frac{1}{\sqrt{\pi}}\right] \tag{10}$$

where $\sigma=\alpha\,|\,f\,|$ is the assumed standard deviation, is the uncertainty factor (10 % by default), f is the forecast value, $z = \frac{O-\mu}{\sigma}$ is the standardised difference, O is the actual observed value, $\Phi(z)$ is the standard normal cumulative distribution function (CDF), and

$$\phi(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}} \tag{11}$$

is the standard normal probability density function (PDF). CRPS generalises the Mean Squared Error to probabilistic forecasts by evaluating entire predictive distributions [28]. A value of zero is ideal and occurs when all probability is assigned to the observed outcome [28,29]. For deterministic forecasts, CRPS reduces exactly to MAE, making it a proper scoring rule that reflects both accuracy and uncertainty.

The Uncertainty Growth Rate (UGR) measures how forecast uncertainty increases with lead time, providing insight into the limits of predictability [30]. A linear form of UGR may be distinguished

$$\text{UGR}_{\text{lin.}}(h) = \frac{\text{d}\big(\text{RMSE}(h)\big)}{\text{d}h} \tag{12}$$

where RMSE(h) is theRMSE at lead time h, Nh, is the number of forecast observation pairs at lead time h, $F_i^{(h)}$ is the i-th forecast value at lead time h, is the i-th observed value.

For directional quantities such as meteorological wave direction and encounter wave angle, uncertainty was quantified using the Circular Mean Absolute Error, defined as

$$\text{CMAE} = \frac{180}{\pi}\cdot\frac{1}{n}\sum_{i=1}^{n}|\text{atan2}(\sin\Delta_i, \cos\Delta_i)| \tag{13}$$

It accounts for the circular nature of angles, ensuring that only the smallest directional difference is evaluated.

5. Uncertainty Analysis

For every location and time, a forecast–observation pair is formed by combining the forecasted value with the corresponding observed value of the same variable. These pairs are grouped in their respective sea-state bins.

Over the analysed four-month period (January–April 2025), each bin accumulates a large number of such pairs, from which uncertainty metrics are computed. This approach assumes that forecast errors within a given sea state are statistically stationary and independent of geographic location, temporal evolution, and synoptic regime. Although temporal correlations are not preserved when grid points transition between bins, this methodology is well aligned with ship routing practice. A vessel encounters sea states spatially along its route rather than at fixed locations, and routing decisions therefore require error statistics that are conditioned on sea state rather than on position. Spatial pooling provides robust statistics for each sea state while remaining computationally efficient.

The non-directional variables exhibit systematic forecast degradation with lead time for significant wave height, wave period, and wind speed. For wave height, RMSE increases from approximately 0.05–0.12 m at 24 h to 0.64–1.04 m at 168 h in moderate seas, showing near-linear growth, while severe sea states display accelerated

error growth beyond 72 h. MAE remains about 15–20 % lower than RMSE across all conditions, indicating the persistent presence of operationally relevant outliers. In calmer seas, relative uncertainty becomes pronounced despite small absolute errors, whereas in higher sea states, predictability limits emerge at longer horizons.

Wave period forecasts are markedly more stable, with RMSE increasing from roughly 0.11–0.17 s at 24 h to 1.16–1.40 s at 168 h in moderate conditions, and exhibiting weaker dependence on sea state. Probabilistic verification confirms good calibration throughout. These patterns demonstrate that forecast uncertainty evolves in a variable- and sea-state-dependent manner, providing a necessary basis for analysing its propagation into attainable ship-speed predictions.

Wind speed exhibits the highest relative uncertainty among all analysed variables. In moderate seas, RMSE increases from approximately 0.45–0.85 m/s at 24 h to 3.04–4.37 m/s at 168 h, while in higher sea states it can exceed 8.0 m/s at maximum lead times. The associated exponential UGR of about 1.3–1.45 % per hour is comparable to the 1.5–1.8 % observed for wave height, indicating similar predictability limits. Across all variables, forecast uncertainty shows a non-linear dependence on sea state. Although absolute errors generally grow with severity, relative uncertainty follows a U-shaped pattern, with the largest values in very calm and extreme conditions. This demonstrates that forecast reliability depends not only on lead time but also on environmental intensity.

Directional variables require circular metrics, with CMAE accounting for angular discontinuity. An example of results is shown in Figure 1. Wind direction variability differs markedly across sea states: in moderate seas, CMAE increases from about 3.5–4.8° at 24 h to 38–46° at 168 h, while in calm conditions it can exceed 50° even at short lead times. In severe weather, stronger atmospheric forcing yields 20–30 % lower CMAE than in moderate seas. The steepest growth occurs in the 24–72 h window, where CMAE increases by 5–20°, coinciding with key tactical planning horizons. Wave direction forecasts are markedly more stable. In moderate seas, CMAE grows from roughly 1.3–3.4° at 24 h to 16.6–31.4° at 168 h, with calm conditions reaching about 40° only at long lead times. In sea states 6–9, CMAE remains below 30° even at extended horizons. Encounter wave angle exhibits the most complex behaviour, with CMAE ranging from 1.3–3.4° at 24 h to 16.6–31.4° at 168 h in moderate seas. Uncertainty is up to 40 % higher in following seas than in head seas, as small directional errors can shift conditions from benign to hazardous. In sea states above 7, CMAE may increase by up to 25° within a single 24-hour update, indicating predictability barriers not captured by current ensemble systems and highlighting a critical asymmetry for route optimization.

Cross-correlation analysis shows that encounter-angle uncertainty cannot be approximated as a linear combination of wind and wave direction errors. Their coupling varies strongly with sea state, with correlation coefficients increasing from about 0.3 in calm conditions (sea states 0–2) to 0.85 in storms (sea states 7–9), reflecting the dominance of unified atmospheric systems in severe weather. When combined with the inherent 15° heading variability of autopilot control, this interaction produces compound uncertainties that exceed root-sum-square estimates by 15–25 %.

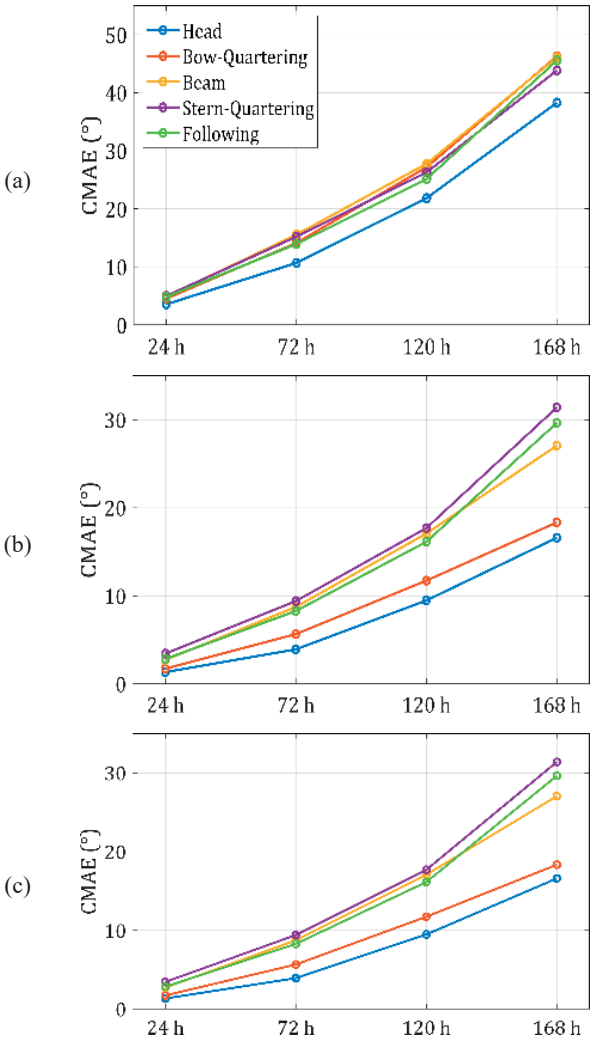

Fig. 1. Uncertainty measures for directional meteorological predictors in sea state WMO five, corresponding to significant wave heights between two point five and four metres: wind direction (*a*), wave direction (*b*), and encounter wave angle (*c*).

Uncertainty propagation to attainable ship speed varies across modelling approaches and operational conditions. For an intended speed of 14.5 kn in moderate seas, the JONSWAP-based NTPRO model exhibits RMSE growth from approximately 0.06–0.11 kn at 24 h to about 0.45–0.82 kn at 168 h, following an almost linear trend with acceleration beyond 72 h. MAE remains around 20 % lower, indicating the influence of occasional large deviations, while bias remains close to zero and agreement indices indicate stable performance.

The Pierson–Moskowitz implementation yields slightly higher uncertainty, with RMSE approaching 0.93 kn at maximum lead time, reflecting increased sensitivity to wave-field variability. Its spatial skill degrades more rapidly, and probabilistic calibration weakens beyond the medium range. NavCad produces the lowest overall uncertainty, with RMSE reaching roughly 0.76 kn at 168 h. Its stepped growth pattern reflects the quasi-static nature of the resistance model, while a small positive bias of about 0.25 kn indicates a tendency to overestimate speed loss. The relatively constant uncertainty growth rate of approximately 2% per hour suggests predictable behaviour, which is advantageous for risk-aware routing.

Correlation analysis was performed across all simulation cases to quantify how meteorological forecast errors propagate into attainable ship speed uncertainty. Pearson coefficients were computed between corresponding uncertainty metrics of predictor and response variables for 3 modelling approaches (NTPRO–JONSWAP, NTPRO–PM, NavCad), 2 reference speeds (12.0 and 14.5 kn), sea states 0–7, and 5 encounter-angle groups.

Three robust patterns emerge. First, a clear hierarchy governs uncertainty transfer: wave height shows the strongest and most consistent correlations with speed errors (typically 0.65–0.97 for RMSE and MAE), followed by wave period (0.45–0.85), while wind-related variables exhibit the widest variability (0.26–0.99). This reflects the dominant role of wave-induced resistance in speed loss. Second, the strength of the correlation depends strongly on the encounter geometry. Head seas (0–30°) yield the highest positive correlations, beam seas (75–105°) the weakest and most scattered, and following seas (150–180°) exhibit bimodal behaviour with either strong positive or negative coupling, indicating threshold effects in stern-wave interactions. Third, coupling intensifies with lead time: short-range forecasts (0–24 h) show moderate correlations (0.65–0.85), whereas long-range horizons (120–168 h) produce either very strong (>0.95) or very weak (<0.30) relationships, revealing increasing determinism or complete decoupling as forecast skill degrades.

For the 14.5 kn case under the Pierson–Moskowitz model, wave-height RMSE and MAE correlations remain consistently high (0.223–0.973), peaking in head seas (0.870–0.970) and increasing with lead time. Bias correlations are strongly negative (−0.713 to −1.000), indicating that wave-height overprediction directly induces speed underprediction. Wind-speed correlations are weaker by 8–10 % (0.259–0.996) and show predominantly positive bias, implying that wind overprediction tends to coincide with speed overprediction and suggesting limitations in aerodynamic modelling.

The JONSWAP implementation produces more polarised behaviour, with frequent near-perfect correlations (±1.000) at extended lead times, indicating overly deterministic uncertainty transfer under fully developed sea assumptions. CRPS correlations are largely negative in the head and following seas (−0.949 to −0.134), implying that improved probabilistic wind forecasts may reduce speed prediction reliability under strongly nonlinear conditions. UGR correlations vary widely (−0.858 to 0.890), with the strongest positive coupling in stern-quartering seas and extreme negative values in bow-quartering seas at short horizons, highlighting the strong geometry dependence of predictability.

6. Discussion and Conclusions

This study demonstrates that weather forecast uncertainty propagates into attainable ship speed in a structured yet highly nonlinear manner. By organising forecast observation pairs within sea-state bins, the proposed framework aligns uncertainty quantification with the way vessels experience weather in practice, spatially along their routes rather than at fixed locations. This enables statistically robust, operationally meaningful error characterisation that is directly applicable to routing algorithms.

The results show that forecast degradation is strongly variable-dependent. Significant wave height exhibits near-linear error growth in moderate conditions and accelerated degradation in severe seas, whereas wave period remains comparatively stable and well-calibrated. Wind speed displays the highest relative uncertainty, with nonlinear dependence on sea state and pronounced errors in both very calm and extreme conditions. Directional variables reveal even more complex behaviour: wind direction uncertainty is highly sensitive to sea state and forecast horizon, whereas wave direction remains comparatively stable due to the coherence of swell systems. Encounter wave angle, synthesising meteorological and navigational uncertainty, exhibits strong asymmetry between head and following seas, with up to 40 % higher uncertainty in following conditions. This asymmetry has direct implications for routing, as small directional errors can shift a vessel from benign following seas into hazardous quartering regimes.

Propagation to attainable ship speed preserves and amplifies these structures. All models exhibit increasing speed uncertainty with lead time, but with distinct characteristics. The JONSWAP-based implementation shows quasi-linear growth with limited bias, the Pierson–Moskowitz implementation is more sensitive to spatial variability and degrades faster beyond the medium range, while NavCad produces lower absolute uncertainty but introduces a systematic tendency to overestimate speed loss. These differences highlight that uncertainty is not solely a property of the forecast, but also of the vessel response model through which it is filtered.

Several limitations should be acknowledged. The analysis is restricted to a single vessel type and operational speed range. Although the simulation matrix covers wave heights up to 12 m, the correlation analysis focuses on sea states 0-7, reflecting typical routing conditions rather than survival scenarios. Extreme environments, where vessels would normally reroute, are therefore underrepresented. Ocean currents were neglected, and the coupling between wind and wave errors is based on the characteristics of the GFS-based dataset and the North Atlantic region.

Future work should extend the framework to multiple ship types and loading conditions, incorporate ocean current fields and fuel consumption models, and evaluate ocean

area-specific behaviour across different forecast systems. Finally, embedding the proposed uncertainty models into real-time optimisation algorithms will allow systematic assessment of risk-aware routing strategies in operational environments.

This work reframes weather routing from a deterministic optimisation problem into an uncertainty-aware decision process. Coupling meteorological predictability with vessel response provides a physically grounded basis for risk-informed navigation, where efficiency and reliability are treated as inseparable objectives.

Acknowledgement. This work has been supported by the Croatian Science Foundation under the project HRZZ-IP-2022-10-2821.

7. References

[1] Kim, M., Hizir, O., Turan, O., Day, S., Incecik, A., 2017. Estimation of added resistance and ship speed loss in a seaway. Ocean Engineering, 141, 465-476.

[2] Lang, X., Mao, W., 2020. A semi-empirical model for ship speed loss prediction at head sea and its validation by full-scale measurements. Ocean Engineering, 209, 107494.

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

[4] Jiao, J., Sun, S., Ren, H., 2016. Predictions of wave induced ship motions and loads by large-scale model measurement at sea and numerical analysis. Brodogradnja, 67(2), 81-100.

[5] Marjanović, M., Prpić-Oršić, J., Turk, A., Valčić, M., 2025. Anomalous Behavior in Weather Forecast Uncertainty: Implications for Ship Weather Routing. Journal of Marine Science and Engineering, 13(6), 1185.

[6] Ding, J.-F., Tseng, W.-J., Sung, Y.-J., 2024. An evaluation of operational risks for general cargo ship operators. Brodogradnja, 75(1), 75101.

[7] Valčić, M., Antonić, R., Tomas, V., 2011. ANFIS Based Model for Ship Speed Prediction. Brodogradnja, 62(4), 373-382.

[8] Moreira, L., Vettor, R., Guedes Soares, C., 2021. Neural net-work approach for predicting ship speed and fuel consumption. Journal of Marine Science and Engineering, 9(2), 119.

[9] Lu, D., Wang, A., Gan, H., Su, Y., Ao, X., 2025. A multi-objective collaborative optimization method of ship energy efficiency based on NSGA-II and TOPSIS. Brodogradnja, 76(3), 76301.

[10] Shin, G.-H., Yang, H., 2025. Deep reinforcement learning for integrated vessel path planning with safe anchorage allocation. Brodogradnja, 76(3), 76305.

[11] Guan, W., Xi, Z., Cui, Z., Zhang, X., 2025. Adaptive trajectory controller design for unmanned surface vehicles based on SAC-PID. Brodogradnja, 76(2), 76206.

[12] Ksciuk, J., Kuhlemann, S., Tierney, K., Koberstein, A., 2023. Uncertainty in maritime ship routing and scheduling: A literature review. European Journal of Operational Research, 308(2), 499-524.

[13] Vettor, R., Bergamini, F., Guedes Soares, C., 2021. A comprehensive approach to account for weather uncertainties in ship route optimization. Journal of Marine Science and Engineering, 9(12), 1434.

[14] Vettor, R., Guedes Soares, C., 2022. Reflecting the uncertainties of ensemble weather forecasts on the predictions of ship fuel consumption. Ocean Engineering, 250, 111009.

[15] Gao, J., Zhang, Y., 2024. Ship collision avoidance decision-making research in coastal waters considering uncertainty of target ships. Brodogradnja, 75(2), 75203.

[16] Wärtsilä, 2011. Wärtsilä Navigation simulator NTPRO 5000. Ship Speed Modeling in Wärtsilä NTPRO 5000.

[17] Zhang, S., Cheng, H., Deng, Z., Mei, L., Ding, L., Guo, C., Wang, X., Zhao, G., 2023. Navigational Safety Assessment of Ten-Thousand-Ton Vessels in Ship Tunnels by Ship Simulations. Water, 15(20), 3584.

[18] Fossen, T.I., 2011. Handbook of Marine Craft Hydrodynamics and Motion Control. John Wiley & Sons, Ltd., Hoboken, USA.

[19] HydroComp, 2023. NavCad - The premier hydrodynamic and propulsion system simulation software. https://www.hydrocompinc.com/solutions/navcad/

[20] HydroComp, 2024. HydroComp NavCad 2024 User's Guide. HydroComp, Inc.

[21] Bassam, A.M., Phillips, A.B., Turnock, S.R., Wilson, P.A., 2015. Ship voyage energy efficiency assessment using ship simulators. VI International Conference on Computational Methods in Marine Engineering (MARINE 2015), Rome, Italy, 15-17 June 2015.

[22] Yan, D., Chen, C., Gan, W., Sasa, K., He, G., Yu, H., 2025. Carbon intensity indicator (CII) compliance: Applications of ship speed optimization on each level using measurement data. Marine Pollution Bulletin, 212, 117593.

[23] Marjanović, M., Valčić, M., Prpić-Oršić, J., Barić, M., 2026. A data-driven framework for attainable ship speed uncertainty under stochastic weather conditions. Brodogradnja, 77(1), 77108.

[24] Buizza, R., Leutbecher, M., 2015. The forecast skill horizon. *Quarterly Journal of the Royal Meteorological Society*, 141, 3366-3382.

[25] Wilks, D.S., 2019. *Statistical Methods in the Atmospheric Sciences*, 4th ed. Academic Press, Cambridge, MA, USA.

[26] Willmott, C.J., Robeson, S.M., Matsuura, K., 2011. A refined index of model performance. *International Journal of Climatology*, 32(13), 2088-2094.

[27] Antonio, B., Aitchison, L., 2025. How to derive skill from the Fractions Skill Score. https://arxiv.org/abs/2311.11985 (accessed 5th July 2025)

[28] Hersbach, H., 2000. Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting*, 15, 559-570.

[29] Gneiting, T., Balabdaoui, F., Raftery, A.E., 2007. Probabilistic Forecasts, Calibration and Sharpness. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 69, 243-268.

[30] Rodwell, M.J., Wernli, H., 2023. Uncertainty growth and forecast reliability during extratropical cyclogenesis. *Weather and Climate Dynamics*, 4(3), 591-615.