Brodogradnja Volume 77 Number 1 (2026) 77108

journal homepage: www.brodogradnja.fsb.hr

Brodogradnja

An International Journal of Naval Architecture and Ocean Engineering for Research and Development

A data-driven framework for attainable ship speed uncertainty under stochastic weather conditions

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

[1] *University of Rijeka, Faculty of Engineering, Vukovarska 58, 51000 Rijeka, Croatia*
[2] *University of Zadar, Maritime Department, Mihovila Pavlinovića 1, 23000 Zadar, Croatia*

ARTICLE INFO

Keywords:
Attainable Ship Speed
Weather Uncertainty
Ship Speed Uncertainty
Data-driven modelling
Voyage Planning

ABSTRACT

This paper presents a data-driven framework for quantifying attainable ship speed uncertainty considering weather forecast uncertainty. The methodology integrates two parallel workflows: weather forecast processing and ship performance simulation. Weather forecast data from NOAA and GFS sources are collected at multiple lead times (0-24h, 24-72h, 72-120h, 120-168h). The data undergo spatial discretisation over a North Atlantic rectangular grid, extracting the main meteorological variables, including significant wave height, peak period, wave direction, wind speed, and wind direction. Ship performance simulations were done using Wärtsilä NaviTrainer NTPRO 5000 and HydroComp NavCad to generate attainable ship speed lookup tables under varying conditions: intended speeds (14.5, 13.5, 12.0 kn), wave heights (0-14 m according to WMO Sea State Codes 0-8), and wave encounter angles (0°-180°). Multiple metrics were used for uncertainty quantification, including RMSE, MAE, Bias, UGR, CRPS, IoA, and FSS for meteorological variables, alongside CMAE for directional parameters. These metrics are subsequently applied to estimated attainable ship speeds, establishing response variable uncertainties. Correlation analysis was conducted between the uncertainty of meteorological variables and the uncertainty in attainable ship speed, providing important insights for estimated time of arrival (ETA) calculations and voyage planning under weather uncertainty.

1. Introduction

Optimal ship routing under varying and severe weather conditions still remains a challenge within the shipping industry. The prediction of attainable ship speed in different sea states needs to be more accurate for voyage planning in the initial phase, as well as along the ship's route. Fuel efficiency and schedule reliability are also affected by these predictions. Ships operating in adverse weather conditions often experience significant speed losses, mainly due to added resistance from waves and wind, degraded propeller performance, and voluntary speed reductions for safety reasons [1,2]. These speed losses result from the interactions between environmental loads and the ship, as demonstrated through bridge simulators such as Wärtsilä NTPRO 5000, which models the real-time ship's dynamics in 6 degrees of freedom (6DOF) [3,4]. Current weather routing systems heavily rely on deterministic speed-power models that assume perfect

* Corresponding author
E-mail address: marijana.marjanovic@riteh.uniri.hr

https://doi.org/10.21278/brod77108

Received 10 August 2025; Received in revised form 16 September 2025; Accepted 3 October 2025
Available online 8 October 2025
ISSN 0007-215X; eISSN 1845-5859

weather forecast accuracy, which is inherently wrong because weather predictions degrade significantly over time. For example, wind speed forecast errors were found to grow non-linearly over five days in the North Atlantic [5]. This uncertainty needs to be considered in voyage planning and route optimization approaches to avoid potentially suboptimal routing decisions. Wu et al. [6] have confirmed that forecast uncertainty grows with the prediction horizon, also providing computational methods for prediction intervals applicable to ship operations. The economic implications are substantial, as stated in [7], where it's stated that speed optimization according to weather conditions could achieve 15-20 % CO_2 reductions for platform supply vessels.

Recent research in attainable ship speed modelling covers different methodologies, though each approach exhibits distinct limitations. While empirical and semi-empirical methods are considered computationally efficient, their applicability is often constrained [8], using 1,477 data points, but this validation focused primarily on specific vessel types and sea conditions, not accounting for generalizability. Semi-empirical models for speed loss in head waves were developed as well [9], although excluding beam and following seas, which limits operational relevance. Methods for estimating added resistance were proposed in [10,11], but validated only against a single vessel type, undermining the broader applicability, while calm-water predictions were improved in [12]. However, validation was still not extended to realistic operational conditions. Some studies focused mainly on added resistance due to diffraction effect [13], considering design and low speed, while leaving high-speed performance unaddressed.

Physics-based models, apart from being computationally complex, consider first-principles hydrodynamics but often have idealised assumptions. The potential-flow approaches, such as strip theory and panel methods [10], neglect viscous effects and nonlinear phenomena, which are fundamental in severe conditions. Even though voyage data was coupled with hindcast weather [14], the reliance on hindcast rather than forecast data limits practical application for voyage planning.

Simulator-based research, while providing controlled environments, often has validation challenges, lacking a comprehensive comparison to full-scale trials [15,16]. EEDI effects were examined using simulators [17], but translating the findings to real-world operations still requires careful consideration of scale effects and environmental complexity. In addition, CFD approaches have improved, even though they are sometimes computationally prohibitive for operational use. A large-scale model tested in natural sea conditions was used to predict wave-induced ship motions and loads, with results validated through measurements and numerical analysis [18]. In [19] the CFD approach was extended to Unsteady Reynolds-Averaged Navier-Stokes (URANS) and Large Eddy Simulation (LES) for parametric rolling, though the computational cost makes such approaches highly impractical for ship route optimization. Insights on trim and swell-induced speed loss were provided in [20,21] but, yet again, only for specific ship types without generalised frameworks. Bow wing concepts were demonstrated as well [22], nonetheless transitioning from CFD validation to practical implementation remains unexplored. Even though simulator environmental modelling was enhanced through CFD [23] it addressed only training applications.

Machine learning (ML) approaches are becoming more popular, but they experience challenges regarding interpretability and generalisation as well. Recent advances in artificial intelligence have shown promising results in maritime applications. A deep reinforcement learning approach for integrated vessel path planning with safe anchorage allocation was developed in [24], demonstrating the potential of AI-based decision-making in complex port operations. While our framework focuses on weather-induced speed loss, it's worth noting that path planning challenges extend to various maritime contexts. The authors in [25] addressed path planning for a manta ray-inspired underwater glider using NSGA-III for multi-objective optimization, considering energy consumption, travel time, and detection range. Similarly, the authors in [26] proposed an adaptive trajectory controller using Soft Actor-Critic combined with PID control for unmanned surface vehicles, achieving improved tracking performance under uncertain conditions. Supervised learning approaches were reviewed in [27], noting how most studies lack cross-validation for different ship types and routes. ML methods for resistance prediction were compared [28,29], but their models require extensive training data, which is often unavailable for newer ships or some specific routes. Multi-objective optimization approaches have gained attention for ship energy efficiency. In [30], the authors developed a comprehensive optimization framework using NSGA-II and TOPSIS that simultaneously considers propeller optimization

and navigation speed, achieving a 2.11 % improvement in propulsion efficiency and 10.39 % reduction in NOx emissions under specific conditions. In [31] and [32], various ML architectures were developed, while domain adaptation was proposed in [33], acknowledging the critical limitation that models trained on one specific ship transfer poorly to others. ML was also integrated into decision support [34,35]; however, neither study addressed weather prediction confidence intervals, which are essential for risk assessment in navigation. Fouling-induced speed loss was investigated in [36], representing one performance degradation mechanism. Comprehensive Arctic benchmarking was provided in [37], where it was revealed that even sophisticated ML approaches struggle with pointwise predictions and achieve accuracy only if averaged over route segments.

When it comes to forecast uncertainty quantification, the research in this area remains largely disconnected from ship speed prediction models. The importance of risk assessment in maritime operations cannot be overlooked. In [38], the authors evaluated operational risks for general cargo ship operators using an AHP-based risk matrix model, identifying critical factors such as delays at ports and perils of the sea as high-risk areas requiring specific management strategies. The connection between ensemble uncertainties and ship fuel consumption was made in [39], but their simplified resistance models may not capture the complex speed-power relationships. The uncertainty visualisation was improved in [40], although it did not address how mariners should integrate this information into routing and speed change decisions. Uncertainty sources are systematically reviewed in [41], but limited guidance is provided on practical implementation. In [42], forecast degradation was quantified with an Adaptive Neuro-Fuzzy Inference System (ANFIS) model, which applies to wind and waves, rather than vessel response. On the other hand, an ANFIS-based model for ship speed prediction was developed in [43]. In [44], first-order probabilistic frameworks were proposed, though linearization assumptions may be invalid for highly nonlinear ship responses.

Although optimization techniques for weather routing were studied [45], computational feasibility for real-time decisions needs to be specifically addressed. Some studies have addressed uncertainty in ship trajectory prediction and collision avoidance. In [46], the authors proposed a collision avoidance decision-making framework for coastal waters that explicitly considers the uncertainty of target ships using AIS data clustering and Gaussian mixture models for trajectory prediction. This approach is particularly relevant for congested coastal areas where multiple vessels interact. Advanced evolutionary algorithms for optimization were proposed in [47] and [48], requiring extensive function evaluations, which are impractical for time-sensitive routing. In [49], digital twin frameworks are presented for Carbon Intensity Indicator (CII) compliance, while more recently, digital twin capabilities were enhanced through reinforcement learning (RL) [50] for adaptive ship performance prediction. When considering specific navigational conditions, Arctic routing was reviewed in [51], acknowledging the lack of data and several model validation challenges in ice-covered waters. Weather forecasts were also integrated with ECDIS interfaces [52], however, without stating the impacts of forecast uncertainties.

Despite extensive research, several research gaps can be identified. Foremost, while individual modelling approaches appear promising, comprehensive frameworks that integrate empirical, physics-based, CFD, and ML methods are still underdeveloped. Most studies validate models under specific conditions without systematic assessment for different ship types, loading conditions, or sea states. Second, although forecast uncertainty is recognised [5,39], very few studies quantify its distribution through attainable ship speed models. Existing research mostly considers speed modelling and uncertainty separately or, in the best case, uses simplified propagation methods. Comparative assessments also usually focus on only one approach rather than evaluating performance under various realistic weather forecast uncertainties. None of the existing frameworks simultaneously develops attainable ship speed models and quantifies the propagation of forecast uncertainty, while also comparing different approaches under operational conditions.

This research aims to address these gaps through several integrated objectives. First, three attainable ship speed models are integrated, combining insights from the navigational simulator experiments (NTPRO 5000), a hydrodynamic and propulsion system simulation tool (NavCad), and real weather forecast data. This data-driven framework integrates empirical formulations and physics-based seakeeping constraints. Unlike recent studies [5,6,35], the quantified forecast uncertainty impact is extended to complete uncertainty propagation chains.

For a practical demonstration of how weather forecast uncertainty affects attainable ship speed in voyage planning, the changes in the uncertainty of time of arrival (ETA) are shown. This enabled the consideration of time-varying forecast skill, spatial error correlation, and model-specific sensitivities. We evaluated not only accuracy, but also the uncertainty of ship speed under forecast uncertainty. We discuss computational efficiency implications for real-time ship routing and operational considerations, along with recommendations for future research.

2. Methodology

The overall research framework for quantifying attainable ship speed uncertainty under stochastic weather conditions is presented in Figure 1.

The proposed framework employs a parallel processing approach that integrates two complementary main data streams: weather forecast uncertainty quantification (left branch, highlighted in blue) and ship performance simulation (right branch, highlighted in pink). The final part of the framework, highlighted in green, presents joint inter-relationship analysis between various uncertainty metrics with applications in voyage planning and ship routing.

The first data stream pertains to the analysis of selected meteorological variables obtained from reliable sources, including weather forecasts and actual sea states (e.g., NOAA, GFS). Forecast data are pre-processed by selected lead times and by sea state according to the Douglas scale. To associate this data with specific spatial locations, we perform a spatial discretisation of the North Atlantic region into a rectangular grid whose nodes define the spatial points of interest. For each node, we extract both forecast and observed (actual) values of key variables, including significant wave height, wave period, wave direction, wind speed, and wind direction. These data are essential for quantifying the uncertainty of the aforementioned variables, i.e. for computing uncertainty metrics (e.g., RMSE, MAE, Bias, UGR, CRPS, IoA, FSS, CMAE) described in Section 2.5. The uncertainty metrics are determined for all variables of interest across the specified lead-time bins and sea-state classes.

The second data stream pertains to the analysis of uncertainty metrics for an attainable ship speed under various sea states. In this context, we first determine the attainable speeds for a selected reference vessel across a range of sea conditions. We employed the simulation environments Wärtsilä NaviTrainer NTPRO 5000 and HydroComp NavCad. Simulations were conducted using the mathematical model of the reference ship for three intended ship speeds, 13 significant wave heights, and 13 encounter wave angles. The NaviTrainer NTPRO 5000 supports simulations with two wave spectra (JONSWAP and Pierson–Moskowitz), whereas NavCad uses a resistance-decomposition approach. The results from all simulation scenarios were organised as lookup tables. Based on these results, it becomes possible to determine the attainable ship speed for both observed (actual) and forecast meteorological inputs from the first data stream of the framework, as a function of intended ship speed, lead time, and sea state. Having obtained attainable-speed values in this manner, we then quantified their uncertainty for all three modelling approaches (JONSWAP, Pierson-Moskowitz, and NavCad). In this regard, we computed the same uncertainty metrics for attainable ship speed as those used for the meteorological non-directional variables.

Having quantified the uncertainties associated with all meteorological variables of interest, together with the attainable ship speed uncertainties obtained from the three alternative approaches, one can subsequently carry out the inter-relationship analyses between predictor (meteorological) and response (ship speed) variables, ultimately providing quantitative insights for voyage planning applications. This integrated approach enables the systematic propagation of weather forecast uncertainties through ship performance models, as described in the following subchapters.

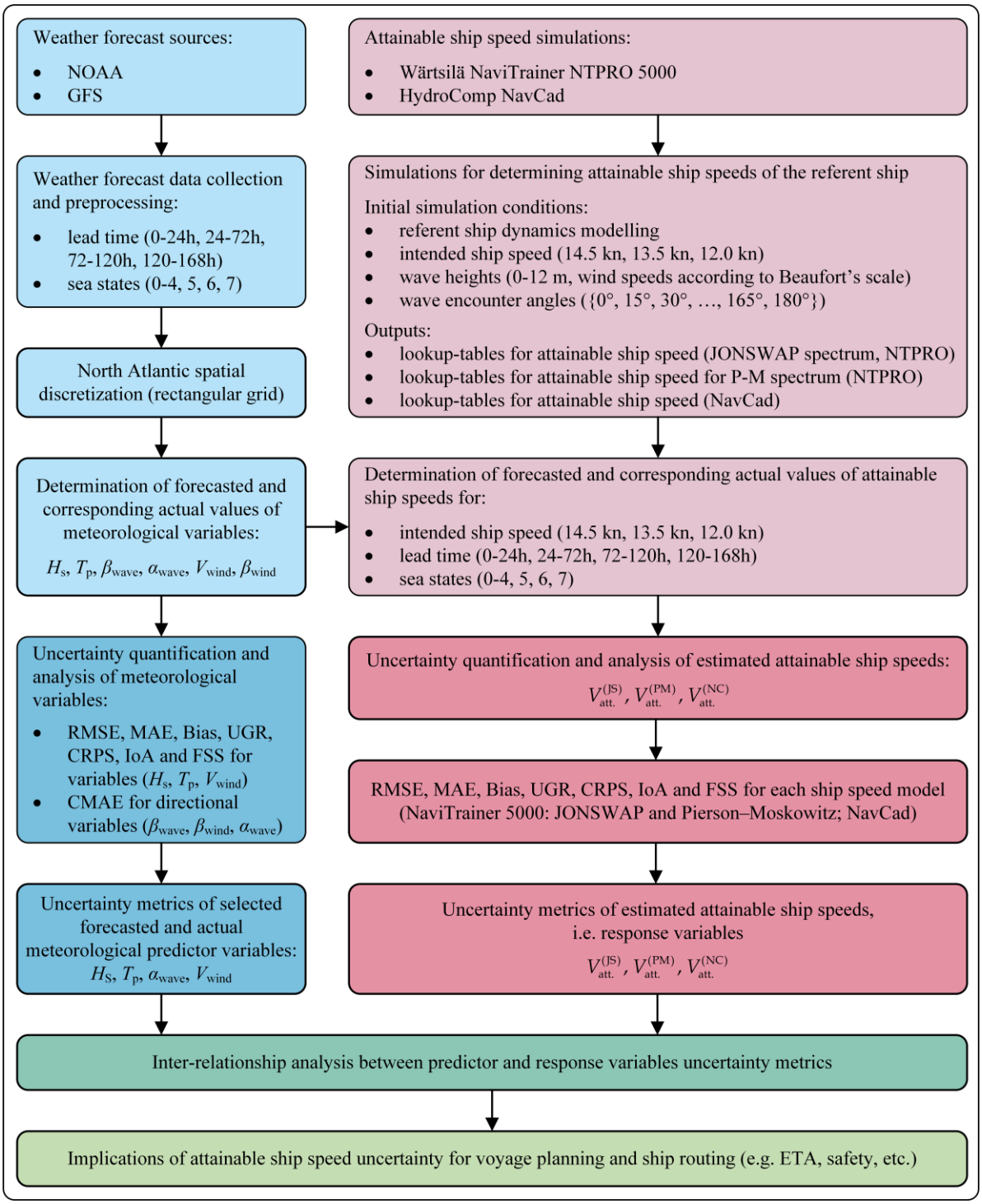


Fig. 1 Integrated framework for quantifying attainable ship speed uncertainty under stochastic weather conditions.

2.1 Experimental Setup

2.1.1 Vessel Characteristics

A 28,050 DWT bulk carrier was chosen as a reference vessel [53] that also served as the basis for the development of a mathematical model for the NTPro 5000 simulator and for NavCad software as well. The vessel in question is a typical medium-sized bulk carrier with principal dimensions of 160.40 m length between perpendiculars (L_{PP}), 27.20 m beam, and 13.60 m depth. With a design draft of 9.819 m, the vessel's gross tonnage is 17,009 tons, and she has a deadweight capacity of 28,189 tons. The propulsion system consists of a two-stroke marine diesel engine rated at 6,150 kW (8,361 PS) at 136 rpm under nominal conditions, although

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 diameter of 5.25 m, a mean pitch of 3.686 m, and a 35-degree skew angle optimized for bulk carrier operations. The service speed is approximately 14 knots, but after analysing the measured data, three speeds that were most prevalent were chosen for simulations. These specifications were carefully integrated into the mathematical model to ensure accurate representation of the vessel's response to environmental loads, which is particularly important for simulating realistic speed loss behaviour in different sea states.

2.1.2 Simulation Environment

Hydrodynamic and propulsion system simulation tools, such as HydroComp's NavCad, employ physics-based models to predict the speed-power characteristics [54,55]. The tool's Analytical Distributed-Volume Method (ADVM) provides an analysis of hull form effects on drag, connecting empirical estimates and full CFD. Meanwhile, real-time simulators such as the Wärtsilä Navi-Trainer Professional 5000 (NTPro 5000) integrate how environmental factors naturally reduce ship speed through physics-based calculations [56,57]. NTPro 5000 is a full-mission bridge simulator that served as the primary platform for ship speed modelling and performance assessment. The core architecture of NTPro 5000 system solves the vessel's motion equations in real time, accounting for 6 degrees of freedom (surge, sway, yaw, heave, roll, pitch) in order to realistically simulate ship dynamics in waves. The system's DNV Class A certification validates its accuracy in ship manoeuvring performance, including acceleration/deceleration, turning, and stopping distances.

The setup workflow for NavCad follows a four-step process starting from the initial inputs of the vessel's parameters, including length, beam, draft, and displacement. Resistance prediction methods can be user-defined [55]. The second phase involves defining the propulsion system parameters. The engine's power curve is linked with the propeller's thrust curve, finding the operating point for each speed where the propeller's required torque equals the engine's available torque and the thrust equals the hull resistance.

Environmental modelling parameters were configured for North Atlantic conditions, including wave spectral characteristics. Within the NTPRO 5000 simulator, one can choose between the Pierson-Moskowitz and JONSWAP formulations to ensure consistency across both computational platforms. Considering that a two-parameter Pierson-Moskowitz (PM) spectrum was created for fully developed wind-generated seas [58], it was our first choice for modelling of environmental conditions. On the other hand, the JONSWAP (JS) spectrum [58] was also used to describe non-fully developed seas.

2.2 Simulation Design

2.2.1 Environmental Conditions

A comprehensive range of environmental conditions was considered to fully capture the ship's performance in different sea states. Wind conditions were simulated according to the Beaufort scale (0-12), with speeds ranging from calm to hurricane force (0-58 knots), considering both relative wind speed and direction effects on ship resistance and stability. Ocean currents were not included in this analysis, as the study focused on the combined effects of wind and wave-induced speed loss, which represent the primary environmental factors affecting ship performance in North Atlantic routes. Wave heights were evaluated from calm conditions up to the sea state that corresponds to 12 m wave height values, number 8 on the Douglas scale. The full spectrum of wave encounter angles from 0° to 180° at 15° increments was covered, represented with the following notation: head seas (0°), quartering seas (45°, 135°), beam seas (90°), and following seas (180°), thereby capturing all possible ship-wave interaction scenarios. Three previously mentioned reference ship speeds were selected for the analysis as follows: 12.0, 13.5, and 14.5 knots.

2.2.2 Ship Speed Loss

The two computational methods used for computing ship speed loss by both NTPro 5000 and NavCad differ in their handling of ship dynamics. NTPro 5000 applies a time-domain simulation approach based on 6-DOF motion equations solved in the body-fixed reference frame [56]. The system integrates the motion equations in real-time, accounting for surge, sway, heave, roll, pitch, and yaw movements. This approach

captures the complex interactions between hull hydrodynamics, propulsion forces, and environmental disturbances. The hydrodynamic forces on the hull are decomposed into positional and damping components, where the positional forces arise from the vessel's drift angle relative to water flow, while damping forces result from the vessel's angular velocities [56]. The simulator determines these forces through experimentally derived coefficients obtained from tank tests, which are stored in tabulated form across the full range of drift angles and yaw rates. When experimental data is unavailable, the system uses trigonometric series expansions and interpolation formulas calibrated against vessel dimensions and hull form characteristics.

Environmental loads in NaviTrainer are computed through distinct models for wind and wave disturbances that account for both steady and dynamic effects. The wave-induced forces consist of first-order oscillatory components that drive the vessel's seakeeping motions and second-order mean drift forces that contribute to steady speed loss and course deviation. These forces are calculated using generalised reduction coefficients that depend on the vessel's draft through the Smith effect, the wavelength to ship length ratio governing force distribution along the hull, and the wave encounter angle. Wind forces are determined from the apparent wind velocity, which combines the true wind with the vessel's motion, acting on the projected lateral and transverse areas above the waterline. The aerodynamic coefficients vary with apparent wind angle and are derived from wind tunnel tests or empirical formulations based on vessel superstructure configuration. The simulator continuously adjusts propeller thrust and rudder angle through the autopilot system to maintain the commanded speed and heading against these environmental loads, with the resulting speed loss emerging naturally from the force balance [56].

NavCad implements a methodical resistance decomposition approach where the total resistance experienced by the vessel is separated into distinct physical components that can be individually calculated and summed [55]. The software's architecture enables the selection of appropriate calculation methods for each resistance component based on vessel type, operational profile, and available data. For bare-hull resistance, NavCad primarily employs the ITTC-1978 correlation line methodology [55], which separates viscous and wave-making resistance components while accounting for Reynolds number effects through form factors. The appendage resistance calculations consider the drag contributions from rudders, bilge keels, shaft brackets, and other hull protrusions, with correction factors applied based on their alignment with the flow field. This component-based methodology enables systematic evaluation of design modifications and their impact on overall resistance.

The environmental resistance components in NavCad are managed through specialised modules that can implement various prediction methods suited to different vessel types and operating conditions. For wind resistance, the software offers multiple calculation approaches ranging from simplified parametric methods suitable for preliminary design to detailed calculations based on vessel-specific wind areas and drag coefficients [55]. Wave-added resistance predictions can utilise regression-based methods derived from systematic model test series, strip theory approaches for slender vessels, or empirical corrections based on sea state and vessel response characteristics. When methods that directly predict speed loss are selected, such as the Aertssen method for weather routing applications, NavCad performs an inverse calculation using the vessel's calm-water resistance curve to convert the speed loss prediction into an equivalent added resistance [55]. This conversion maintains consistency within the software's resistance-based framework while accommodating diverse prediction methodologies developed for specific vessel types and operational scenarios.

2.3 Attainable Ship Speed

The attainable ship speed data was collected through an extensive number of simulations conducted on the NTPro 5000 navigation simulator for:

a. 13 sea states according to various significant wave heights, $H_S \in \{0,1,2,...,12\}$ (m)

b. 13 encounter wave angles, $\alpha_{waves} \in \{0,15,30,...,180\}$ (°)

c. 2 spectra, $S \in$ {'Pierson-Moskowitz', 'JONSWAP'}

d. 2 loading conditions, $L \in \{\text{'Full load','Ballast'}\}$

e. 3 intended referent ship speeds, $V_{\text{ref.}} \in \{12, 13.5, 14.5\}$ (kn)

which gives a total of 2028 simulations.

Parallel simulations were executed in HydroComp's NavCad through its scripting Application Programming Interface (API), which enabled automated batch processing of the identical 1014 simulation scenarios because of no possibilities for explicit wave spectra settings [55]. The NavCad scripting functionality permitted systematic variation of environmental parameters and vessel conditions while maintaining exact correspondence with the NTPro 5000 simulation matrix, thereby facilitating direct comparison between the time-domain and quasi-static computational approaches.

All data processing and analysis were performed using MATLAB R2024b and Python 3.13.5, enabling the development of three distinct lookup table functions for attainable speed computation. This comprehensive wind speed range ensures the framework captures ship performance across all operationally relevant conditions, from port departures in calm weather to severe storm avoidance scenarios typical of North Atlantic winter routes.

While the simulations were initially conducted for wave encounter angles from 0° to 180°, the results were extended to the full 0-360° range by applying symmetrical principles, as ship responses to port and starboard wave encounters are mirror images. For practical navigation and route optimization applications, ship headings were assigned at 15° intervals throughout the complete 0-360° compass range, enabling the calculation of relative encounter angles for any combination of ship course and wave direction. For operational implementation, the attainable ship speed values are obtained through bilinear interpolation between the discrete simulation points in the lookup tables. This ensures smooth transitions for intermediate values of wave height and encounter angle rather than using rounded or nearest-neighbour approximations.

The encounter wave angle $\alpha_{\text{waves}} \in [0, 2\pi\rangle$, as depicted in Figure 2, can be expressed in terms of the ship heading $\psi \in [0, 2\pi\rangle$ and meteorological wave direction $\beta_{\text{waves}} \in [0, 2\pi\rangle$ as:

$$\alpha_{\text{waves}} = \begin{cases} \beta_{\text{waves}} - \psi, & \text{for } \psi \le \beta_{\text{waves}} \\ 2\pi + \beta_{\text{waves}} - \psi, & \text{for } \psi > \beta_{\text{waves}}. \end{cases} \tag{1}$$

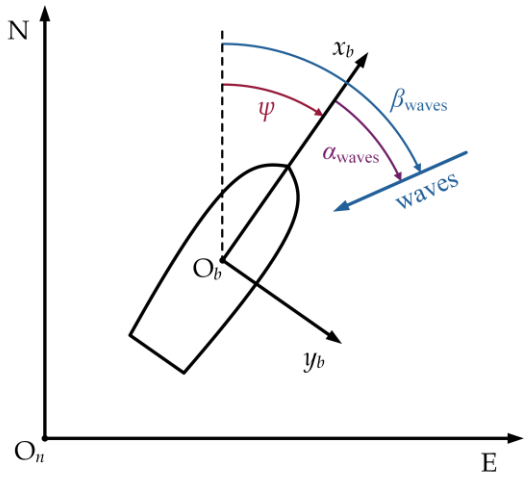


Fig. 2 The definition of the encounter wave angle relative to the ship's heading

If one neglects the wind and ocean current loads, the attainable ship speed $V_{\text{att.}}$ can be expressed as a function of intended referent ship speed $V_{\text{ref.}}$ and current sea conditions represented by the significant wave height H_{S}, wave period T_{P} and encounter wave angle α_{waves}, which yields:

$$V_{\text{att.}} = f(V_{\text{ref.}}, H_{\text{S}}, T_{\text{P}}, \alpha_{\text{waves}}). \tag{2}$$

The sea conditions can be actual, for actual values of $H_{\text{S}}^{(\text{act.})}$, $T_{\text{P}}^{(\text{act.})}$ and $\alpha_{\text{waves}}^{(\text{act.})}$, and forecasted, for forecast values of $H_{\text{S}}^{(\text{for.})}$, $T_{\text{P}}^{(\text{for.})}$ and $\alpha_{\text{waves}}^{(\text{for.})}$. The attainable ship speed $V_{\text{att.,actual}}$ in actual conditions is:

$$V_{\text{att.,actual}} = f(V_{\text{ref.}}, H_{\text{S}}^{(\text{act.})}, T_{\text{P}}^{(\text{act.})}, \alpha_{\text{waves}}^{(\text{act.})}) \tag{3}$$

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

$$V_{\text{att.,predicted}} = f(V_{\text{ref.}}, H_{\text{S}}^{(\text{for.})}, T_{\text{P}}^{(\text{for.})}, \alpha_{\text{waves}}^{(\text{for.})}). \tag{4}$$

As previously pointed out, the attainable ship speed was calculated by means of the navigational simulator Wärtsilä NTPRO 5000, both for the Pierson–Moskowitz spectrum ($V_{\text{att.}}^{(\text{PM})}$) and the JONSWAP spectrum ($V_{\text{att.}}^{(\text{JS})}$), and as well as with the HydroComp NavCad software ($V_{\text{att.}}^{(\text{NC})}$).

Several assumptions that define the operational limitations of the developed models need to be acknowledged. While the theoretical framework presented in Equations (2)-(4), as well as those from Section 2.6, includes peak wave period (T_{p}) as an independent variable affecting attainable ship speed, the practical implementation in this study requires clarification. In the NTPRO 5000 simulations, when using the Pierson-Moskowitz spectrum, the peak wave period is not an independently controllable parameter but rather a derived quantity determined by the spectrum formulation.

Similarly, in the JONSWAP spectrum implementation, the peak period remains coupled to the significant wave height through the wind-wave relationship. This coupling reflects the physical reality that in wind-generated seas, wave height and period are not independent but evolve together according to the fetch, duration, and wind speed [56]. Therefore, Tables 1-4 present attainable speeds as functions of the encounter angle, with T_{p} implicitly included through the spectral relationships rather than as an independent variable. This simplification is justified for the North Atlantic routes studied here, where swell and wind seas typically align, and the Pierson-Moskowitz assumption of fully developed seas is reasonable.

For the NavCad simulations, the software's internal wave resistance algorithms similarly couple wave period and height based on standard wave statistics for the specified sea states, consistent with the approach used in classification society guidelines [55].

Throughout all simulations, the ship's heading was initially set to 000° while environmental load angles were varied across all encounter angles. An autopilot system was utilised for course-tracking under varying environmental conditions [56]. For each environmental condition, simulations were started at initial speeds, with the autopilot system adjusting the ship speed to maintain course.

Another limitation within the lookup tables manifests as NaN (Not a Number) values, representing sea states where the autopilot system could no longer maintain the demanded course. When environmental loads exceeded the ship's directional control capability, characterised by excessive yaw rates and vertical motions including severe roll, pitch, and heave amplitudes, or when the autopilot was effectively "thrown off" course, the corresponding speed values were designated as NaN. This approach effectively sets the operational boundary beyond which navigation becomes dangerous or impossible. In the ship route optimization and decision support context, these NaN values serve as indicators for areas that require avoidance or course and speed alteration decisions. This essentially translates into navigational risk zones where the vessel cannot safely maintain its intended heading due to the severity of environmental conditions and associated ship motions.

However, that does not mean that the vessel could not sail even under such sea states with a substantial voluntary reduction in speed, but this scenario is penalised by route optimisation itself. The overview of simulation results, for several selected cases, is given in Tables 1, 2, 3 and 4.

Table 1 Simulated attainable ship speeds obtained with NTPRO 5000 for the Pierson–Moskowitz spectrum, full load conditions, and intended ship speed of 14.5 kn

	Encounter wave angles α_{waves} (°)												
H_s (m)	0	15	30	45	60	75	90	105	120	135	150	165	180
0	14.50	14.50	14.50	14.50	14.50	14.50	14.50	14.50	14.50	14.50	14.50	14.50	14.50
1	14.11	14.09	14.06	14.04	14.10	14.20	14.27	14.31	14.33	14.35	14.37	14.37	14.38
2	13.52	13.53	13.47	13.37	13.32	13.51	13.70	13.85	13.89	13.98	14.06	14.09	14.11
3	12.55	12.50	12.30	12.05	11.97	12.33	12.80	13.04	13.09	13.40	13.59	13.67	13.68
4	10.93	10.96	10.82	10.90	11.11	11.22	11.52	11.69	11.95	12.56	13.09	13.31	13.41
5	10.15	10.30	10.48	10.88	10.82	11.14	10.83	9.21	11.63	12.51	13.00	13.28	13.40
6	9.47	9.60	9.81	9.97	10.27	10.97	10.35				12.81	13.17	13.37
7	9.19	9.33	9.55	9.74	10.25	10.98					12.76	13.15	13.36
8	8.64	8.90	9.17	9.35	8.80	8.44					12.11	12.94	13.20
9	7.96	8.13	8.21	7.71								12.50	12.91
10	7.52	7.51	7.69	7.41									
11	7.23	7.09	6.99										
12	6.90	6.60	6.36										

Table 2 Simulated attainable ship speeds obtained with NavCad, full load conditions, and intended ship speed of 14.5 kn

	Encounter wave angles α_{waves} (°)												
H_s (m)	0	15	30	45	60	75	90	105	120	135	150	165	180
0	14.50	14.50	14.50	14.50	14.50	14.50	14.50	14.50	14.50	14.50	14.50	14.50	14.50
1	13.84	13.84	13.86	13.88	13.91	13.95	13.99	14.04	14.08	14.11	14.14	14.15	14.16
2	12.87	12.87	12.89	12.93	12.97	13.03	13.09	13.15	13.22	13.27	13.32	13.35	13.36
3	11.93	11.94	11.97	12.02	12.09	12.17	12.23	12.31	12.38	12.45	12.50	12.53	12.55
4	11.26	11.27	11.28	11.31	11.35	11.41	11.48	11.56	11.65	11.73	11.79	11.81	11.82
5	10.67	10.68	10.70	10.74	10.79	10.80	10.82	10.87	10.93	11.00	11.06	11.10	11.11
6	10.07	10.08	10.09	10.11	10.15	10.20	10.26				10.40	10.43	10.45
7	9.43	9.44	9.46	9.50	9.56	9.64					10.10	10.16	10.17
8	9.03	9.04	9.06	9.09	9.13	9.19					9.60	9.65	9.67
9	8.38	8.41	8.52	8.68								9.20	9.22
10	7.42	7.46	7.58	7.77									
11	6.46	6.51	6.64										
12	5.50	5.55	5.70										

Table 3 Simulated attainable ship speeds obtained with NTPRO 5000 for the Pierson–Moskowitz spectrum, full load conditions, and intended ship speed of 12.0 kn

	Encounter wave angles α_{waves} (°)												
H_s (m)	0	15	30	45	60	75	90	105	120	135	150	165	180
0	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00
1	11.61	11.59	11.53	11.53	11.59	11.71	11.78	11.83	11.88	11.88	11.90	11.92	11.92
2	10.97	10.98	10.91	10.79	10.73	10.84	11.11	11.28	11.32	11.47	11.59	11.64	11.65
3	9.82	9.76	9.61	9.55	9.38	9.50	9.85	10.11	10.16	10.81	11.09	11.22	11.26
4	8.48	8.54	8.60	8.61	8.81	8.66	8.91			9.63	10.48	10.97	11.17
5	8.17	8.25	8.35	8.54	8.41	8.59	7.32			9.45	10.38	10.89	11.07
6	7.68	7.75	7.77	7.86	8.00	8.34					10.17	10.84	11.04
7	7.40	7.45	7.49	7.52	7.66	7.63					9.99	10.83	11.01
8	6.79	6.97	7.15	7.32	6.56	6.12						10.60	10.87
9	6.19	6.28	6.30	6.04									10.62
10	5.82	5.81	5.98	5.83									
11	5.56	5.43	5.20										
12	5.30	5.05											

Table 4 Simulated attainable ship speeds obtained with NavCad, full load conditions, and intended ship speed of 12.0 kn

	Encounter wave angles α_{waves} (°)												
H_s (m)	0	15	30	45	60	75	90	105	120	135	150	165	180
0	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00	12.00
1	11.61	11.61	11.62	11.64	11.66	11.68	11.70	11.72	11.75	11.76	11.78	11.79	11.79
2	11.22	11.23	11.24	11.27	11.31	11.35	11.40	11.45	11.49	11.53	11.56	11.57	11.58
3	10.83	10.84	10.87	10.91	10.96	11.03	11.10	11.17	11.24	11.29	11.33	11.36	11.37
4	9.90	9.92	9.98	10.08	10.20	10.34	10.50		10.80	10.92	11.02	11.08	11.10
5	9.06	9.09	9.17	9.31	9.48	9.68	9.90			10.49	10.63	10.71	10.74
6	8.22	8.26	8.36	8.54	8.76	9.02				10.06	10.24	10.34	10.38
7	7.80	7.82	7.89	8.01	8.15	8.32					9.11	9.18	9.20
8	6.84	6.87	6.96	7.09								8.53	8.56
9	5.88	5.91	6.02	6.18									7.92
10	4.92	4.96	5.08	5.27									
11	3.96	4.01	4.14										
12	3.00												

2.4 Weather Forecast Data Processing

Weather forecast uncertainties arise from multiple intersecting sources throughout the entire prediction chain. Initial condition uncertainties stem from sparse oceanic observations, with satellite altimeters measuring wave heights with ±(0.25-0.5) m errors and covering only narrow tracks separated by 200-500 km, while in-situ buoys are distributed hundreds of kilometres apart, creating substantial data voids over open oceans. Numerical weather prediction models like NOAA's Global Forecasting System (GFS), operating at 0.25° resolution (~27 km), cannot resolve sub-grid phenomena and rely on imperfect parameterisations for processes like wave generation, dissipation, and air-sea interactions, with these physical approximations contributing 30-40 % of total forecast error [59]. Data assimilation compounds these uncertainties through observation operator errors, simplified background error covariance assumptions, and quality control procedures that reject 15-30 % of observations. As forecasts evolve, chaotic error growth doubles synoptic-scale errors every 2-3 days, imposing a theoretical predictability limit of approximately two weeks. Post-processing adds further uncertainty through spatial and temporal interpolation. Converting from 6-hourly model output to a continuous time series can introduce 10-20 % error for rapidly changing parameters. For wave parameters, the WaveWatch III (WW3) global wave analysis assimilates significant wave height (H_s) measurements from satellite altimeters, including Jason-3 and Sentinel-3A/B, complemented by in-situ observations from NOAA's NDBC buoy network, providing both global coverage and ground-truth validation at approximately 0.5° resolution with 6-hourly updates [60]. Wave direction (β_{waves}) information is derived from directional wave buoys and Sentinel-1 SAR wave mode data when available. Wind field verification relies on the GFS 0.25° analysis, which incorporates ASCAT scatterometer observations from MetOp-B and MetOp-C satellites for ocean surface wind vectors, AMSR2 microwave radiometer-derived wind speeds, and conventional ship and buoy reports [61,62].

Satellite observations provide data as well, though each measurement system has inherent limitations. Radar altimeters measure significant wave height with approximately 10 % accuracy (±0.5 m for moderate seas) by analysing radar pulse broadening, but only sample along narrow ground tracks with multi-day revisit periods [62,63]. Scatterometers provide wide-swath wind coverage, enabling near-daily global coverage, though rain contamination and coastal proximity compromise data quality. Synthetic Aperture Radar (SAR) captures directional wave spectra but cannot resolve waves shorter than 150-200 m due to velocity bunching effects. These diverse observations ultimately undergo data assimilation procedures including quality control screening, bias correction, spatial thinning to ~100 km spacing, and optimal weighting with model backgrounds. This produces analysis fields that are not pure observations but rather model-observation blends. Essentially, this observational uncertainty of approximately 0.2-0.3 m in wave height analyses sets a fundamental limit on achievable forecast accuracy. When modern wave models achieve similar RMSE values at short lead times, they approach the theoretical limit of predictability given current observational constraints.

The spatiotemporal matching methodology ensures appropriate forecast-observation comparisons through careful interpolation and alignment procedures. Spatial matching uses bilinear interpolation for scalar variables (H_s, T_p, V_{wind}) while utilising vector component interpolation for directional quantities to maintain circular consistency. Near coastal boundaries, the methodology switches to nearest-neighbour interpolation to prevent land contamination. Temporal alignment restricts forecast-observation pairs to those within a ±3-hour window, balancing data availability with temporal consistency. Multi-level quality control procedures are implemented from initial sensor-level checks through variational quality control (VarQC) within the assimilation systems, ensuring robust verification statistics while maximising data utilisation [59,62].

The weather forecast dataset that was used in this study was obtained for a period extending from early January through the end of April 2025. This temporal coverage was purposely selected in order to capture the diverse meteorological conditions which are characteristic of the North Atlantic winter and early spring seasons, when weather varies significantly. The data were sourced from the NOAA GFS, which can be publicly accessed on their servers. The acquisition process involved systematic automated retrieval of forecast outputs across multiple initialisation cycles, with new forecasts issued every 6 hours (00, 06, 12, 18 UTC) and extending to 168-hour lead times, ensuring comprehensive temporal coverage and enabling uncertainty quantification across the full forecast horizon.

The spatial domain was focused on the North Atlantic Ocean area, with forecast data extracted at 2,619 discrete geographical points of the rectangular grid bounded by 20°N to 65°N latitude and 70°W to 10°W longitude. Grid points located over land masses were automatically excluded from the analysis, ensuring that only oceanic data points relevant to ship routing were retained. These points were strategically positioned at 50 nautical mile intervals, as shown in Figure 3, to capture mesoscale weather patterns, while maintaining practical data processing requirements, which created a spatial resolution that balances computational efficiency with coverage that's adequate for ship routing applications.

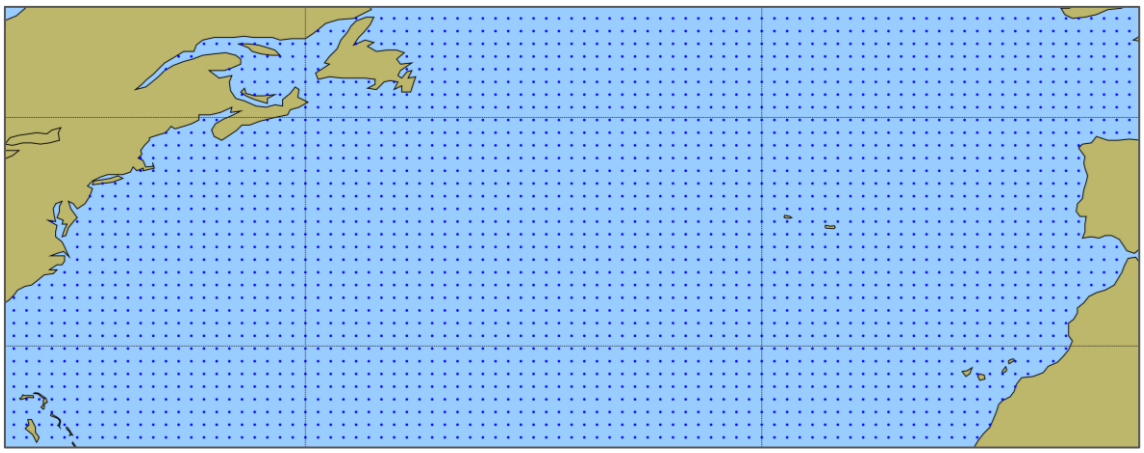


Fig. 3 The rectangular grid of points 50 NM apart in the North Atlantic

The analysis of weather variables was divided into four distinct forecast lead time windows, each serving specific operational planning requirements. Short-range forecasts (0-24 hours) were considered for capturing immediate operational decisions with new forecast outputs at 0, 6, 12, 18, and 24-hour intervals. Medium-range forecasts (24-72 hours) were analysed for tactical voyage planning with data points at 30, 36, 42, 48, 54, 60, 66, and 72 hours. Extended-range forecasts (72-120 hours) can enable strategic route optimization with predictions at 78, 84, 90, 96, 102, 108, 114, and 120 hours. Finally, long-range forecasts (120-168 hours) served to provide advance planning capabilities with outputs at 126, 132, 138, 144, 150, 156, 162, and 168 hours. This temporal segmentation enables an accurate assessment of forecast skill degradation across different planning horizons, which is crucial for understanding the propagation of uncertainty in ship speed predictions.

The dataset included both forecasted and actual observed values for six primary meteorological variables important for ship performance estimation: significant wave height (H_s) measured in meters, representing the average height of the highest one-third of waves; peak wave period (T_p) in seconds, indicating the dominant

wave period in the spectrum; wind speed (V_{wind}) recorded in meters per second at the 10-meter reference height; wave direction (β_{waves}) specified in degrees, indicating the direction from which waves propagate related to the North axis; wind direction (β_{wind}) measured in degrees, denoting the direction from which wind originates related to the North axis; and encounter angle (α_{waves}) calculated in degrees, representing the relative angle between vessel heading and wave direction.

The computational framework for data processing utilised Python libraries optimized for handling large meteorological datasets. The processing pipeline included: (*i*) data decoding using libraries such as cfgrib and eccodes for GRIB2 format interpretation; (*ii*) spatial-temporal organisation for efficient manipulation and storage of multi-dimensional arrays; (*iii*) statistical analysis, distribution fitting and uncertainty quantification (*iv*) calculation of specialised metrics including CRPS; and (*v*) structured data organisation and batch processing workflows. The data architecture was constructed to handle the computational demands of analysing multiple forecast cycles while maintaining data integrity and enabling easier workflows. Special attention was given to the unique dataset structure, where each meteorological state was replicated for 25 different ship headings (0-360° at 15° intervals) to comprehensively assess encounter angle effects on ship performance.

2.5 Uncertainty Metrics

For the uncertainty analysis of non-directional meteorological variables across multiple forecast horizons (0-24h, 24-72h, 72-120h, and 120-168h) and associated attainable ship speeds, the following metrics have been used, where n denotes the number of observations, and F_i and O_i are i-th forecasted and observed values, respectively.

The Root Mean Square Error (RMSE), as one of the most fundamental accuracy measures in forecast verification [64], can be defined as:

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

A perfect forecast yields RMSE = 0 [64]. The squaring operation, however, makes RMSE particularly sensitive to outlier errors, which aids in identifying systematic biases in ship speed predictions that could impact route planning.

The Mean Absolute Error (MAE) provides a linear measure of average forecast error magnitude [64], and it can be noted as:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|F_i - O_i\right|. \tag{6}$$

Unlike RMSE, MAE weights all errors equally, making it less sensitive to outliers [64]. The relationship between RMSE and MAE provides insights into error distribution characteristics. When RMSE values exceed MAE, it indicates the presence of large outlier errors, as RMSE "penalises large errors more", whereas MAE weights all errors linearly [64]. In practice, both metrics are often advised for comprehensive model comparisons [64].

Bias quantifies systematic forecast tendencies, revealing whether a model consistently over- or under-predicts, and it can be noted as [65]:

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

Positive bias indicates systematic over-prediction, while negative values suggest under-prediction [54]. Unlike RMSE and MAE, bias can approach zero values even with large errors if positive and negative deviations cancel out. For that reason, bias should always be interpreted alongside magnitude-based metrics to distinguish between compensating errors and genuine accuracy [65].

The Index of Agreement, also known as Willmott's Index [66], provides a standardised measure (range 0 to 1) of how well a model's predictions match observations, which is relative to the variability in the observations [66]. It is defined as:

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

where $\bar{O}$ denotes the mean value of all observations.

A value of 1 indicates perfect agreement, while 0 suggests the model performs no better than using the observed mean as a constant predictor [66]. This normalisation makes IoA particularly useful for comparing model performance across variables with different scales and variabilities.

Fractions Skill Score (FSS) is a spatial verification metric designed for high-resolution forecasts of categorical events (e.g., rain exceeding some threshold) [67]. It usually ranges from 0, i.e. no skill, to 1, which represents a perfect forecast. Instead of comparing forecast and observed values point-by-point, FSS compares the fractional coverage of an event within neighbourhoods around each point [67]. The forecast and observation fields are first converted into binary maps (whether an event or no-event above a threshold is present), then a moving window (neighbourhood) is used to calculate the fraction of grid points with the event in both fields. FSS is hence defined as [67]:

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

where MSE_f represents the MSE between the forecast fraction $F(n)$ of forecast grid points exceeding threshold with neighbourhood $n = 1,\ldots,N$, and observed fractions $\mathrm{O}(n)$ of observed grid points exceeding threshold with neighbourhood $n = 1,\ldots,N$, while $\mathrm{MSE}_{f,ref}$ is the reference MSE representing a forecast with no skill, i.e. the worst case scenario, and N denotes the total number of neighbourhoods. One application of FSS is for determining the spatial scale at which a forecast has useful skill; for example, a precipitation forecast might achieve FSS > 0.5 only when evaluated over a 50-km neighbourhood, suggesting reliability at that scale even if exact placement is off [67]. In the case of comparing multiple models, the model with a higher FSS for a given scale provides better spatial accuracy [67].

Continuous Ranked Probability Score (CRPS) extends forecast evaluation to probabilistic predictions, assessing the accuracy of a forecast distribution by comparing it to the observed outcome [68]:

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

where $\sigma = \alpha|f|$ is the assumed standard deviation, α is the uncertainty factor (10 % by default), f is the forecast value, i.e. the distribution mean μ, $z = (O - \mu)/\sigma$ is the standardised difference, O is the actual observed value, $\Phi(z)$ is the standard normal cumulative distribution function (CDF), and:

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

is the standard normal probability density function (PDF).

CRPS generalises the Mean Squared Error to probability distributions [68]. A CRPS value of 0 is considered ideal and is achieved if the forecast assigns all probability to the correct outcome [68,69]. Notably, for deterministic forecasts, CRPS reduces exactly to MAE, making it a proper scoring rule that accounts for both accuracy and appropriate uncertainty quantification. CRPS is commonly used by atmospheric and climate

centres, such as ECMWF, to evaluate ensemble weather forecasts for variables like temperature, precipitation, and wave height.

Uncertainty Growth Rate (UGR) quantifies how the forecast uncertainty evolves with increasing lead time, which can provide insights into predictability limits [70]. The value characterises the exponential growth of forecast error or ensemble spread [5]. One can differ linear UGR:

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

where $\mathrm{RMSE}(h) = \sqrt{\frac{1}{N_h}\sum_{i=1}^{N_h}[F_i^{(h)} - O_i]^2}$ is the RMSE at lead time h, N_h is the number of forecast observation pairs at lead time h, $F_i^{(h)}$ is the i-th forecast value at lead time h, O_i is the i-th observed value and the exponential UGR:

$$\mathrm{UGR}_{\mathrm{exp.}}(h) = \frac{\mathrm{d}\ln(\mathrm{RMSE}(h))}{\mathrm{d}h} \tag{13}$$

where $\mathrm{RMSE}(h) = \sigma_0 e^{\lambda h}$ is the exponential model for the RMSE at the lead time h, σ_0 is the initial uncertainty of RMSE at $h = 0$, and λ is the exponential growth rate parameter. Essentially, UGR quantifies how forecast uncertainty increases with lead time through linear (absolute units/hour) or exponential (relative %/hour) growth rates [70].

For the uncertainty of directional variables like meteorological wave direction β_{waves} and encounter wave angles α_{waves}, the Circular Mean Absolute Error (CMAE) was used. CMAE is defined as:

$$\mathrm{CMAE} = \frac{180}{\pi} \cdot \frac{1}{n}\sum_{i=1}^{n}\left|\mathrm{atan2}(\sin\Delta_i, \cos\Delta_i)\right| \tag{14}$$

where $\Delta_i = F_i - O_i$, expressed in radians. The metric is used because of the 360° discontinuity; a forecast of 1° and an observation of 359° differ by only 2°, not 358° [65]. The transformation ensures that all angular differences fall within the range [-180°, 180°], which makes capturing the minimal angular distance between forecast and observation possible.

2.6 Voyage Planning and ETA estimation

Let the route be defined by a sequence of $n+1$ waypoints: $W = \{W_0, W_1, ..., W_n\}$ where each waypoint is defined as:

$$W_i = \{\varphi_i, \lambda_i, d_i\} \tag{15}$$

where $\varphi_i \in [-90°, 90°]$ is latitude, $\lambda_i \in [-180°, 180°]$ is longitude, and d_i is the distance to the next waypoint in nautical miles, $i = 1, ..., n$.

The cumulative distance to the waypoint i is defined as:

$$D_i = \sum_{j=0}^{i-1} d_j, \tag{16}$$

with $D_0 = 0$, while the total route distance is:

$$D_{\mathrm{total}} = D_n = \sum_{j=0}^{n-1} d_j. \tag{17}$$

For a given distance $d \in [0, D_{\text{total}}]$ along the route, the $k-\text{th}$ segment is identified such that $D_k \le d < D_{k+1}$. The position $P(d) = (\varphi(d), \lambda(d))$ at distance d can be calculated using linear interpolation:

$$f = \frac{d - D_k}{D_{k+1} - D_k} \in [0,1] \tag{18}$$

where $\varphi(d) = \varphi_k + f(\varphi_{k+1} - \varphi_k)$ and $\lambda(d) = \lambda_k + f(\lambda_{k+1} - \lambda_k)$.

For the purpose of weather forecasts, the North Atlantic region has been divided into a rectangular discrete grid $G = \{(\varphi_i^g, \lambda_i^g, t_j) : i \in I, j \in J\}$, where all neighbouring points are 50 nm apart, as shown in Figure 3. For each grid point, we have weather parameters:

$$W_{ij} = (H_{\text{S},ij}, T_{\text{p},ij}, \beta_{\text{wave},ij}, \alpha_{\text{wave},ij}, V_{\text{wind},ij}, \beta_{\text{wind},ij}). \tag{19}$$

Thus, the weather at any point $W_{ij} = W(\varphi, \lambda, t)$ is based on the nearest point:

$$(i^*, j^*) = \arg\min_{i,j} d_{ij} \tag{20}$$

where $d_{ij} = \sqrt{(\varphi_i^g - \varphi)^2 + (\lambda_i^g - \lambda)^2}$.

Temporal selection was based on indices for which criteria $|t_j - t| \le \Delta t_{\max}$, $\Delta t_{\max} = 3\,\text{h}$, is valid. The forecast data was handled with a lead time τ, i.e. as $t_{\text{issue}} = t - \tau$, where t_{issue} is the forecast issue date and time and τ is the forecast hour.

Uncertainty metrics are classified according to:

a. Wave height class: $H_{\text{S}} \in \{[0,2.5], [2.5,4], [4,6], [6,9]\}$ (m)

b. Wave encounter angle class: $\alpha_{\text{wave}} \in \{\text{Head, Bow-Quartering, Beam, Stern-Quartering, Following}\}$

c. Lead time class: $\tau \in \{[0,24], [24,72], [72,120], [120,168]\}$ (h).

Error metric lookup function:

$$\varepsilon(H_{\text{S}}, T_p, V_{\text{wind}}, \alpha_{\text{wave}}, \tau) = \{\text{RMSE, MAE, Bias, UGR, CRPS, IoA, FSS}\} \tag{21}$$

retrieve required uncertainty metrics for a given sea state condition expressed in terms of H_{S}, T_p, V_{wind}, α_{wave} and τ.

For this practical implementation example, the attainable ship speed is calculated based on (2), i.e. as:

$$V_{\text{att.}}^{(\text{PM})} = f_{\text{PM}}(V_{\text{ref.}}, H_{\text{S}}, T_p, \alpha_{\text{wave}}). \tag{22}$$

Uncertainty of $V_{\text{att.}}^{(\text{PM})}$ is estimated based on associated uncertainty metrics $(H_{\text{S}}, \alpha_{\text{wave}}, \tau)$, while the actual ship speed follows a normal distribution around the predicted value, i.e.:

$$V_{\text{att.,actual}}^{(\text{PM})} \sim \mathcal{N}\left(\mu_{V_{\text{att.}}^{(\text{PM})}}, \sigma^2_{V_{\text{att.}}^{(\text{PM})}}\right) \tag{23}$$

where $\mu_{V_{\text{att.}}^{(\text{PM})}} = V_{\text{att.,predicted}}^{(\text{PM})} + \text{Bias}(H_{\text{S}}, \alpha_{\text{wave}}, \tau)$ and $\sigma_{V_{\text{att.}}^{(\text{PM})}} = \text{RMSE}(H_{\text{S}}, \alpha_{\text{wave}}, \tau)$.

The term (23) can be rewritten in a more explicit way as:

$$V_{\text{att.,actual}}^{(\text{PM})} = V_{\text{att.,predicted}}^{(\text{PM})} + \text{Bias}(H_{\text{S}}, \alpha_{\text{wave}}, \tau) + \varepsilon \tag{24}$$

where $\varepsilon \sim \mathcal{N}(0, \mathrm{RMSE}^2(H_S, \alpha_{\mathrm{wave}}, \tau))$.

Ship speed confidence intervals (CI) are defined as:

$$V_{\mathrm{att.}}^{(\mathrm{PM})} \in [V_{\mathrm{att.,predicted}}^{(\mathrm{PM})} + \mathrm{Bias} - k \cdot \mathrm{RMSE},\ V_{\mathrm{att.,predicted}}^{(\mathrm{PM})} + \mathrm{Bias} + k \cdot \mathrm{RMSE}] \tag{25}$$

where k is a coverage factor, e.g., $k = 1.96$ for 95 % confidence.

Given the remaining distance $d_{\mathrm{rem.}}$ (nm) and uncertain attainable ship speed $V_{\mathrm{att.}}^{(\mathrm{PM})}$ (kn), estimated time of arrival ETA (h) can be written as:

$$\mathrm{ETA} = t_{\mathrm{current}} + \frac{d_{\mathrm{rem.}}}{V_{\mathrm{att.}}^{(\mathrm{PM})}}, \tag{26}$$

where t_{current} is the current time.

If one, based on (26), defines a function:

$$g(V_{\mathrm{att.}}^{(\mathrm{PM})}) = t_{\mathrm{current}} + \frac{d_{\mathrm{rem.}}}{V_{\mathrm{att.}}^{(\mathrm{PM})}}, \tag{27}$$

then:

$$g'(V_{\mathrm{att.}}^{(\mathrm{PM})}) = -\frac{d_{\mathrm{rem.}}}{(V_{\mathrm{att.}}^{(\mathrm{PM})})^2}, \quad g''(V_{\mathrm{att.}}^{(\mathrm{PM})}) = \frac{2d_{\mathrm{rem.}}}{(V_{\mathrm{att.}}^{(\mathrm{PM})})^3}. \tag{28}$$

A Taylor expansion of the function (27) about the point $V_{\mathrm{att.}}^{(\mathrm{PM})} = \mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}}$ gives:

$$g(V_{\mathrm{att.}}^{(\mathrm{PM})}) = g(\mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}}) + g'(\mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}})(V_{\mathrm{att.}}^{(\mathrm{PM})} - \mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}}) + \frac{1}{2} g''(\mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}})(V_{\mathrm{att.}}^{(\mathrm{PM})} - \mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}})^2 + \ldots. \tag{29}$$

By dropping the higher-order terms, the approximation of (29) in terms of the first-order Taylor expansion yields:

$$g(V_{\mathrm{att.}}^{(\mathrm{PM})}) \approx g(\mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}}) + g'(\mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}})(V_{\mathrm{att.}}^{(\mathrm{PM})} - \mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}}). \tag{30}$$

If one takes $\mathbb{E}[\cdot]$ of both sides of (30), it can be written:

$$\mathbb{E}[g(V_{\mathrm{att.}}^{(\mathrm{PM})})] \approx \underbrace{g(\mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}})}_{\text{constant}} + g'(\mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}}) \underbrace{\mathbb{E}[V_{\mathrm{att.}}^{(\mathrm{PM})} - \mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}}]}_{\mathbb{E}[V_{\mathrm{att.}}^{(\mathrm{PM})}] - \mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}} = 0} \tag{31}$$

which finally yields:

$$\mathbb{E}[g(V_{\mathrm{att.}}^{(\mathrm{PM})})] = \mathbb{E}[\mathrm{ETA}] \approx g(\mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}}) = t_{\mathrm{current}} + \frac{d_{\mathrm{rem.}}}{\mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}}}, \text{ i.e.} \tag{32}$$

$$\mu_{\mathrm{ETA}} := t_{\mathrm{current}} + \frac{d_{\mathrm{rem.}}}{\mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}}}. \tag{33}$$

From the variance point of view, if one takes $\mathrm{Var}[\cdot]$ of both sides of (30), it yields:

$$\mathrm{Var}[g(V_{\mathrm{att.}}^{(\mathrm{PM})})] \approx \mathrm{Var}[g(\mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}})] + \mathrm{Var}[g'(\mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}})(V_{\mathrm{att.}}^{(\mathrm{PM})} - \mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}})]. \tag{34}$$

Following (23), (26), (27), (28), (34) and simple variance rules such as $\mathrm{Var}(k) = 0$, $\mathrm{Var}(k \cdot X) = k^2 \cdot \mathrm{Var}(X)$ and $\mathrm{Var}(X - k) = \mathrm{Var}(X)$, where k is a constant and X is a random variable, one can write as follows:

$$\mathrm{Var}(\mathrm{ETA}) \approx (g'(\mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}}))^2 \mathrm{Var}(V_{\mathrm{att.}}^{(\mathrm{PM})}) \tag{35}$$

$$\sigma_{\mathrm{ETA}}^2 := \mathrm{Var}(\mathrm{ETA}) \approx \left(\frac{d_{\mathrm{rem.}}}{\mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}}^2} \right)^2 \sigma_{V_{\mathrm{att.}}^{(\mathrm{PM})}}^2 = \left(\frac{d_{\mathrm{rem.}}}{\mu_{V_{\mathrm{att.}}^{(\mathrm{PM})}}^2} \right)^2 \mathrm{RMSE}^2(H_{\mathrm{S}}, \alpha_{\mathrm{wave}}, \tau). \tag{36}$$

Finally, the ETA distribution can be written based on (33) and (36) as:

$$\mathrm{ETA} \sim \mathcal{N}(\mu_{\mathrm{ETA}}, \sigma_{\mathrm{ETA}}^2). \tag{37}$$

For a voyage with multiple segments, each with different conditions $(H_{\mathrm{S}}, \alpha_{\mathrm{wave}}, \tau)$, one can also define the segment time uncertainty. Namely, for any segment i with distance d_i and conditions $(H_{\mathrm{S},i}, \alpha_{\mathrm{wave},i})$, time t_i can be expressed as:

$$t_i = \frac{d_i}{V_{\mathrm{att.},i}^{(\mathrm{PM})}} \tag{38}$$

and therefore $\sigma_{t_i}^2$ can be written as:

$$\sigma_{t_i}^2 = \left(\frac{d_i}{\mu_{V_{\mathrm{att.},i}^{(\mathrm{PM})}}^2} \right)^2 \cdot \mathrm{RMSE}^2(H_{\mathrm{S},i}, \alpha_{\mathrm{wave},i}, \tau_i). \tag{39}$$

Finally, the total ETA uncertainty, under the independence assumption, yields as:

$$\sigma_{\mathrm{ETA,total}}^2 = \sum_i \sigma_{t_i}^2. \tag{40}$$

As the voyage progresses and forecast lead time changes, one can write:

$$\mathrm{RMSE}(\tau) = \mathrm{RMSE}_{\mathrm{base}} \cdot g(\tau) \tag{41}$$

where $g(\tau)$ is an increasing function capturing forecast degradation and $\mathrm{RMSE}_{\mathrm{base}}$ is the baseline RMSE at the initial or shortest forecast lead time.

Finally, ETA point estimate with uncertainty, which is one specific point in time, can be expressed as:

$$\mathrm{ETA} = \mu_{\mathrm{ETA}} \pm k \cdot \sigma_{\mathrm{ETA}}, \tag{42}$$

where μ_{ETA} is the expected (mean) arrival time, k is the coverage factor and σ_{ETA} is the standard deviation of ETA in hours. For instance, if $k = 1.96,$ which corresponds to 95 % predictive intervals, there's a 95 % probability the actual arrival will fall within bounds $\mathrm{ETA} \in [\mathrm{ETA}_{\mathrm{lower}}, \mathrm{ETA}_{\mathrm{upper}}]$, where $\mathrm{ETA}_{\mathrm{lower}}$ presents the earliest likely arrival if conditions are favourable, and $\mathrm{ETA}_{\mathrm{upper}}$ presents the latest likely arrival if conditions are unfavourable. In this context, the coverage probability statement can be written as:

$$\mathbb{P}(\mathrm{ETA}_{\mathrm{actual}} \in [\mathrm{ETA}_{\mathrm{lower}}, \mathrm{ETA}_{\mathrm{upper}}]) = 0.95. \tag{43}$$

From all the above, one can finally conclude that the ship will arrive at the destination port by the target date/time t_{target} with probability $\mathbb{P}$ that can be expressed as:

$$\mathbb{P}\{\mathrm{ETA} \le t_{\mathrm{target}}\} = \Phi\left(\frac{t_{\mathrm{target}} - \mu_{\mathrm{ETA}}}{\sigma_{\mathrm{ETA}}} \right) \tag{44}$$

where $\Phi(\cdot)$ is the cumulative distribution function (CDF) of the standard normal distribution.

3. Results and Discussion

3.1 The Uncertainty Analysis

The uncertainty analysis uses spatial pooling instead of temporal tracking of individual grid points in the North Atlantic area. Each grid point is classified into sea state bins $b \in \{1,2,3,4\}$, which corresponds to the observed significant wave height ranges from the set $H_S \in \{[0, 2.5), [2.5, 4), [4, 6), [6, 9)\}$ (m), respectively. A forecast-analysis pair $(H_S^{(for.)}, H_S^{(act.)})$ consists of a forecasted variable value $H_S^{(for.)}$ and its corresponding observed value $H_S^{(act.)}$ at the same location and time.

The methodology works as follows. Sea states are divided into bins indexed by b, where each bin b has boundaries $H_{S,\min,b}^{(act.)} = \min(H_{S,b}^{(act.)})$ and $H_{S,\max,b}^{(act.)} = \max(H_{S,b}^{(act.)})$. At each time step, a grid point is assigned to a bin b if its observed value $H_{S,b}^{(act.)}$ satisfies $H_{S,\min,b}^{(act.)} \le H_{S,b}^{(act.)} < H_{S,\max,b}^{(act.)}$. The forecasted value $H_S^{(for.)}$ at that same point is then paired with this observed value $H_S^{(act.)}$ to compute uncertainty metrics for a bin b. Over the observed period, which in our case was four months, each bin accumulates a number of forecast-observation variable pairs. This approach assumes statistical stationarity of forecast errors within each sea state bin. The errors are treated as independent of geographic location, temporal evolution patterns, and synoptic weather conditions.

This grouping method does not preserve temporal correlations when grid points transition between different sea state bins. Each grid point location contributes independently to different bins as the weather conditions change. However, this spatial pooling approach is well-suited for ship routing applications. The ship encounters sea states spatially along her routes, not at fixed points. Thus, route optimization requires error statistics for each sea state, independent of specific locations. The large sample sizes collected for each bin in this study ensure robust statistics while maintaining computational efficiency.

3.1.1 Non-Directional Meteorological Variable Uncertainty

The analysis of the uncertainty of non-directional meteorological variables reveals distinct patterns of forecast degradation across the three primary variables: significant wave height (H_s), wave period (T_p), and wind speed (V_{wind}). While the complete analysis encompasses different sea states (H_s = 0-12 m), Figure 4 presents representative uncertainty metrics for sea state 5 (H_s = 2.5-4 m) as an illustrative example of the observed patterns.

For significant wave height across all analysed sea states, RMSE values demonstrate consistent growth patterns from short-range to extended forecasts. In sea state 5 (Figure 4a), RMSE increases from 0.05-0.12 m at 24 h lead time to 0.64-1.04 m at 168 h lead time, exhibiting nearly linear growth. These ranges correspond to differences caused by various RMSE values for each encounter angle class. This pattern aligns with findings in [71], who reported similar linear degradation in North Atlantic wave hindcasts, though our exponential growth rate of 1.5-1.8 % per hour is notably lower than the 2.3 % reported in [6] for their ANFIS-based predictions.

The MAE consistently tracks 15-20 % below RMSE across all sea states, indicating persistent outlier errors that affect operational planning. In calmer conditions (sea states 2-3), the relative uncertainty increases despite lower absolute errors, while severe conditions (sea states 7-9) show accelerated error growth beyond 72-hour lead times, consistent with the predictability limits identified in [39].

Wave period predictions exhibit more stable uncertainty characteristics across the full range of conditions. The analysis reveals RMSE growth from 0.11-0.17 s (24 h) to 1.16-1.4 s (168 h) over the forecast horizon for moderate seas (Figure 4b), with proportionally smaller increases in both calm and severe conditions. The CRPS values indicate well-calibrated probabilistic forecasts throughout.

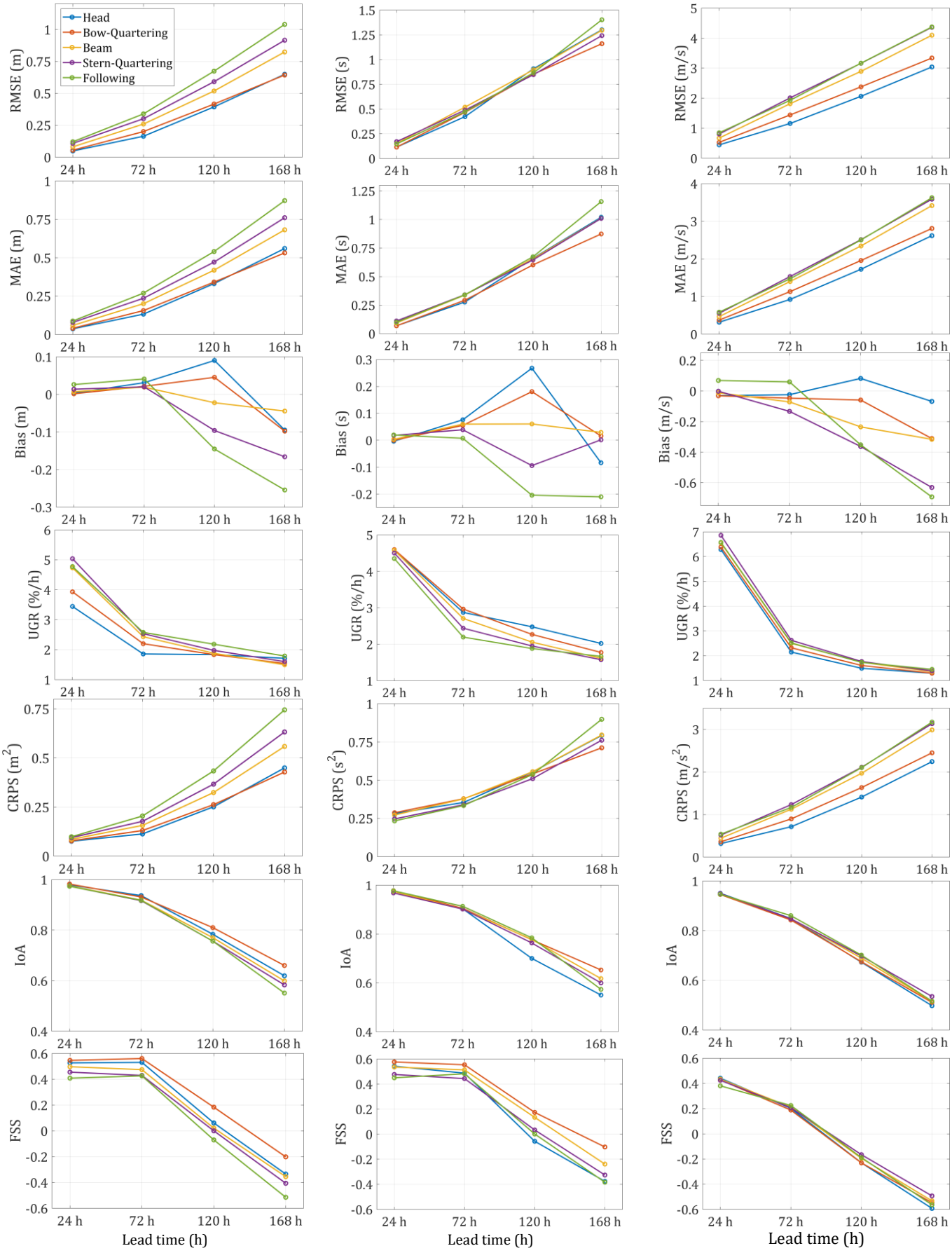


Fig. 4 Uncertainty metrics of meteorological non-directional predictor variables for the sea state 5 (H_S = 2.5-4 m): significant wave height H_S (*a*), wave period T_p (*b*) and wind speed V_{wind} (*c*).

Wind speed forecasts consistently demonstrate the highest relative uncertainty among all meteorological variables analysed. The pattern observed in sea state 5 (Figure 4c), with RMSE increasing from 0.45-0.85 m/s (24 h) to 3.04-4.37 m/s (168 h), is amplified in higher sea states where RMSE can exceed 8.0 m/s at maximum lead times. Corresponding exponential UGR of 1.3-1.45 % per hour is relatively close to 1.5-1.8 % of the observed one for wave height, suggesting similar predictability characteristics.

The correlation between sea state severity and forecast uncertainty reveals non-linear relationships across all three variables. While absolute errors generally increase with sea state, the relative uncertainty (RMSE normalised by mean values) shows a U-shaped pattern, with the highest relative errors in very calm (sea states 0-1) and extreme conditions (sea states 10-12). This pattern has important implications for ship speed predictions, as it suggests that forecast reliability varies not only with lead time but also with the prevailing environmental severity.

3.1.2 Directional Meteorological Variable Uncertainty

The uncertainty characteristics of directional meteorological variables required specialised metrics to account for their circular nature, with the Circular Mean Absolute Error (CMAE) properly handling the 360° discontinuity inherent in directional data. The comprehensive analysis across different sea states reveals complex patterns in directional forecast degradation, with Figure 5 again presenting representative results for sea state 5 (H_s = 2.5-4 m).

Meteorological wind direction uncertainty exhibits pronounced variability across different sea conditions. While Figure 5a shows CMAE values increasing from 3.54-4.81° at 24 h to 38.3-46.1° at 168 h for moderate seas, the analysis reveals that directional uncertainty is strongly modulated by sea state severity. In calm conditions (sea states 0-2), CMAE can exceed 50° even at short lead times due to weak pressure gradients and variable wind patterns. Conversely, during severe weather (sea states 8-10), the stronger atmospheric forcing produces more coherent wind fields, resulting in CMAE values 20-30 % lower than in moderate conditions. The steepest uncertainty growth consistently occurs in the 24-72 h window across all sea states, where CMAE increases by 5-20°, substantially exceeding the 10° increase reported in [44] for Mediterranean conditions. This accelerated degradation in the medium range has critical implications for voyage planning, as it coincides with key tactical decision horizons.

Wave direction forecasts demonstrate markedly superior stability compared to wind direction across the entire spectrum of sea conditions analysed. The CMAE growth from 1.34-3.44° at 24 h to 16.6-31.4° at 168 h observed in moderate seas (Figure 5b) represents the median behaviour, with calm conditions showing only marginally higher uncertainty (CMAE reaching 40° at 168h) despite the challenges of predicting swell propagation in light winds. Notably, in sea states 6-9, wave direction CMAE remains below 30° even at extended lead times, reflecting the dominance of well-defined swell systems. The analysis further reveals that the wave direction forecast skill shows minimal sensitivity to the choice of wave spectrum (JONSWAP vs. Pierson-Moskowitz), contrasting with the spectrum-dependent speed loss variations reported in [1].

The encounter wave angle uncertainty, synthesising both meteorological forecast errors and navigational considerations, presents the most complex patterns across different operational conditions. The CMAE ranges from 1.34-3.44° at 24 h to 16.6-31.4° at 168 h, as shown for sea state 5 in Figure 5c. It should be noted that the encounter wave angle categories (head seas, bow-quartering, beam, stern-quartering, and following) shown in both Figures 4 and 5 were derived by calculating encounter angles for 25 different ship headings (0-360° at 15° intervals) at each grid point, then grouping the results according to the relative angle between the meteorological wave direction and each hypothetical ship heading. This systematic approach allowed us to assess uncertainty patterns across all possible encounter scenarios without specifying a particular route. In following seas (encounter angles 150-180°), uncertainty is amplified by up to 40 % compared to head seas, as small directional changes can shift the encounter angle between favourable following seas and dangerous quartering conditions. This asymmetry, not previously documented in the literature, has profound implications for routing algorithms that typically assume symmetric uncertainty distributions. The non-linear growth pattern intensifies in sea states above 7, where CMAE can increase by 25° within a single 24-hour forecast update cycle, suggesting predictability barriers not captured by current ensemble forecasting systems.

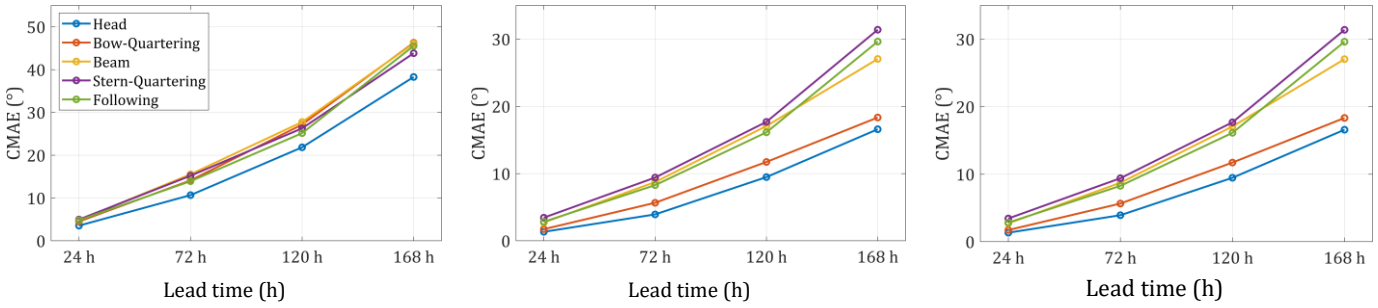


Fig. 5 Uncertainty metrics of meteorological directional predictor variables for the sea state 5 (H_s = 2.5-4 m): meteorological wind direction β_{wind} (*a*), meteorological wave direction β_{wave} (*b*) and encounter wave angle α_{wave} (*c*).

Cross-correlation analysis between directional variables reveals that encounter angle uncertainty cannot be treated as a simple linear combination of its components. The coupling between wind and wave direction uncertainties varies significantly with sea state: correlation coefficients range from 0.3 in calm conditions to 0.85 in storm conditions, indicating that unified atmospheric systems drive both wind and wave fields during severe weather. This coupling effect, combined with the 15° average heading uncertainty inherent in autopilot course-keeping, produces compound uncertainties that consistently exceed root-sum-square estimates by 15-25 %. These findings challenge the independence assumptions underlying current probabilistic routing systems and suggest that Monte Carlo approaches may be necessary for accurate uncertainty propagation. The temporal evolution of directional uncertainty also exhibits distinct diurnal patterns not apparent in the magnitude variables, with CMAE typically 10-15 % higher during nighttime forecast initialisations, possibly reflecting reduced observational data availability.

3.1.3 Attainable Ship Speed Uncertainty

The transformation of meteorological forecast uncertainties into ship speed prediction errors determines the effectiveness of operational weather routing. This analysis examines how three modelling frameworks, NTPRO 5000 (with Pierson-Moskowitz and JONSWAP spectra) and NavCad, propagate weather forecast uncertainties to attainable speed estimates across different operational conditions and sea states. The following Figure 6 shows the uncertainty propagation from meteorological variables to attainable ship speed predictions for an intended speed of 14.5 knots under sea state 5 conditions (H_s = 2.5-4 m). The NTPRO 5000 JONSWAP spectrum results (Figure 6a) show RMSE values increasing from 0.06-0.11 knots at 24 h to 0.45-0.82 knots at 168 h lead time. The uncertainty growth exhibits a quasi-linear pattern, with a notable acceleration after 72 hours. MAE values remain consistently lower than RMSE by approximately 20 %, indicating the presence of outlier predictions that significantly impact error statistics. The bias fluctuates between -0.03 and +0.5 knots, suggesting minimal systematic error in the JONSWAP-based predictions. The IoA maintains values above 0.4 throughout most of the forecast period, demonstrating robust model performance even with increasing uncertainty.

The Pierson-Moskowitz spectrum implementation in NTPRO 5000 (Figure 6b), however, produces slightly higher uncertainty levels, with RMSE reaching 0.93 knots at maximum lead time. This 12 % increase compared to JONSWAP results reflects the different spectral characteristics, particularly in fetch-limited conditions typical of the North Atlantic. The FSS values show more rapid degradation, falling below 0.3 after 96 hours, suggesting that the Pierson-Moskowitz spectrum may be more sensitive to spatial variations in wave field predictions. CRPS values indicate good probabilistic calibration in short to medium-range forecasts but deteriorate notably beyond 72 hours. NavCad predictions (Figure 6c) exhibit the lowest uncertainty levels among the three models, with RMSE reaching 0.76 knots at 168-hour lead time. The distinct stepped pattern in uncertainty growth corresponds to the quasi-static resistance calculation approach, which responds more dramatically to discrete changes in environmental conditions. The positive bias averaging 0.25 knots suggests that NavCad's hydrodynamic model tends to overestimate speed loss in dynamic conditions. Despite higher absolute errors, the UGR remains relatively constant at approximately 2 % per hour, indicating predictable uncertainty growth that could be valuable for risk-based planning.

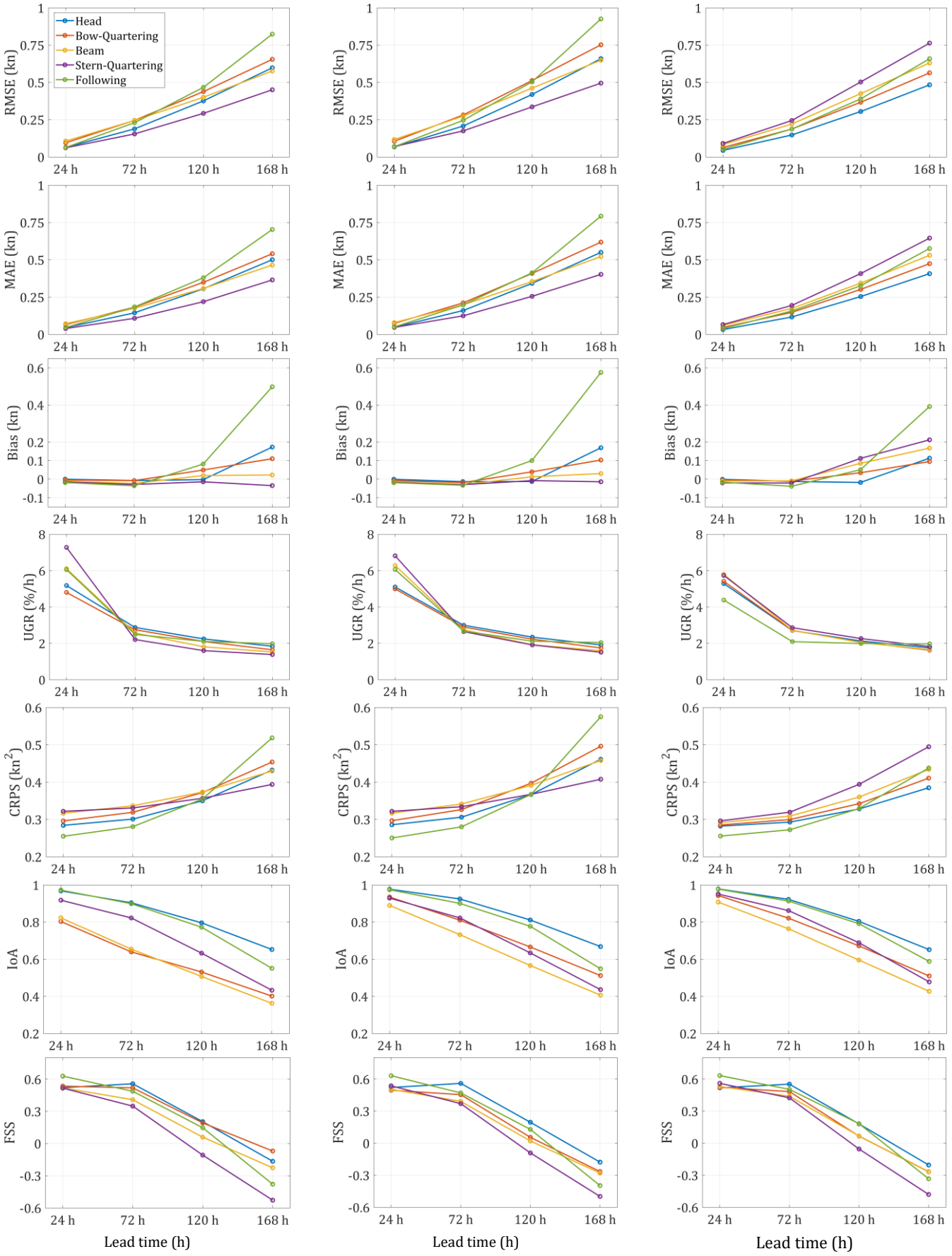


Fig. 6 Attainable ship speed uncertainty metrics for NTPRO 5000 JONSWAP spectrum (*a*), NTPRO 5000 Pierson–Moskowitz (*b*) and NavCad (*c*). Intended ship speed 14.5 kn and sea state 5 (H_s = 2.5-4 m).

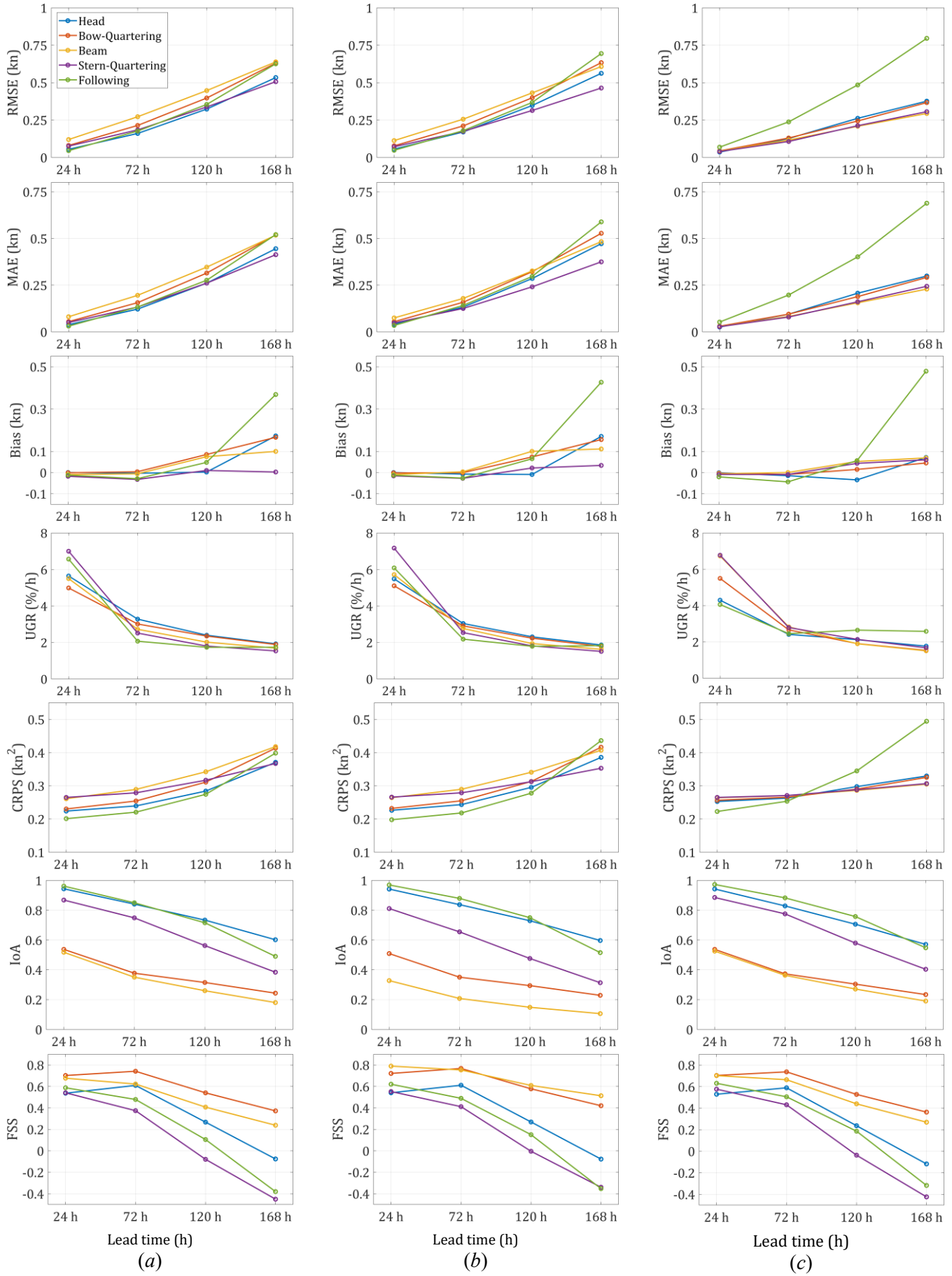


Fig. 7 Attainable ship speed uncertainty metrics for NTPRO 5000 JONSWAP spectrum (*a*), NTPRO 5000 Pierson–Moskowitz (*b*) and NavCad (*c*). Intended ship speed 12 kn and sea state 5 (H_s = 2.5-4 m).

Attainable ship speed uncertainty for a lower intended speed of 12.0 knots under the same sea state 5 conditions is presented in Figure 7, revealing important speed-dependent characteristics in uncertainty propagation. The NTPRO 5000 JONSWAP results (Figure 7a) demonstrate markedly different uncertainty behaviour compared to the 14.5-knot case. RMSE values grew from 0.05-0.12 knots at 24 h to 0.51-0.64 knots at 168 h, representing a 10 % average decrease in absolute uncertainty compared to higher speed operations. The positive bias averaging +0.21 knots at 168 h indicates systematic over-prediction of attainable speeds at lower intended speeds, suggesting that the vessel struggles more to maintain course in challenging conditions when operating below optimal speed.

The Pierson-Moskowitz implementation (Figure 7b) shows even more pronounced uncertainty growth at lower speeds, with RMSE reaching an average of 0.59 knots at maximum lead time. The IoA drops below 0.85 after 96 hours, indicating significant forecast skill degradation. The FSS pattern reveals interesting threshold behaviour, with sharp drops at 72-hour intervals, suggesting that the model's sensitivity to environmental conditions increases at lower operational speeds.

NavCad predictions (Figure 7c) exhibit the most dramatic response to speed reduction for the following seas, with RMSE values reaching 0.8 knots at 168-hour lead time, i.e. a 20 % increase over the 14.5-knot scenario. The variable bias pattern, oscillating between +0.05 and +0.48 knots, suggests that the quasi-static resistance model struggles to capture the non-linear speed-power relationships at lower speeds. The stepped uncertainty growth pattern becomes more pronounced, with distinct jumps corresponding to sea state transitions. UGR analysis reveals an average exponential growth rate of 2 % per hour at 168 h, similar to UGR at a nominal speed of 14.5 knots.

3.2 Correlation Analysis Between Meteorological and Ship Speed Uncertainties

Correlation between attainable ship speeds metrics and selected meteorological variables uncertainty metrics was conducted by means of the Pearson correlation coefficient:

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \tag{45}$$

where X_i is the predictor variable value (e.g., $\mathrm{RMSE}(H_\mathrm{S})$), Y_i is the response variable value (e.g., $\mathrm{RMSE}(V_{\mathrm{att.}}^{(\mathrm{PM})})$), $\bar{X}$ and $\bar{Y}$ are the mean values of the predictor and response variables, respectively, and n is the number of data points.

Considering the excessive amount of result data, only the correlation between corresponding uncertainty metrics was pointed out (e.g. $r(\mathrm{RMSE}(H_\mathrm{S}), \mathrm{RMSE}(V_{\mathrm{att.}}^{(\mathrm{PM})}))$). Therefore, only the diagonal elements of correlation matrices were analysed for the purpose of this work, according to:

$$r_{\mathrm{diag.}}^{(m,l)} = r\{X^{(m)}, Y^{(m)}\} \tag{46}$$

for each metric m and lead time l, where $m \in \{\mathrm{RMSE, MAE, Bias, UGR, CRPS, IoA, FSS}\}$ and $l \in \{0\text{-}24\ \mathrm{h}, 24\text{-}72\ \mathrm{h}, 72\text{-}120\ \mathrm{h}, 120\text{-}168\ \mathrm{h}\}$.

The correlation analysis between meteorological and ship speed uncertainty metrics was conducted systematically across all simulation conditions to quantify how weather forecast errors propagate through different ship performance models. The Pearson correlation coefficient (45) was computed for corresponding uncertainty metrics between predictor (meteorological) and response (ship speed) variables, focusing on diagonal elements of the correlation matrices as expressed in (46). This comprehensive analysis covered three modelling approaches (NTPRO 5000 JONSWAP, NTPRO 5000 Pierson-Moskowitz, and NavCad), two intended speeds (12.0 and 14.5 knots), all sea states (0-7), and five encounter angle groups.

The overall correlation analysis reveals several fundamental patterns in uncertainty propagation that persist across different modelling frameworks and operational conditions. First, the correlation strength

between meteorological and ship speed uncertainties demonstrates clear hierarchical relationships: wave height uncertainties show the strongest and most consistent correlations with speed prediction errors (typically 0.65-0.97 for RMSE/MAE), followed by wave period (0.45-0.85), while wind-related variables exhibit the most variable correlations (0.26-0.99). This hierarchy reflects the dominant role of wave-induced resistance in ship speed loss mechanisms. Second, the correlation patterns exhibit strong dependencies on encounter angle geometry. Head seas (0-30°) consistently produce the highest positive correlations for magnitude-based metrics (RMSE, MAE), indicating direct error propagation. Beam seas (75-105°) show the weakest and most variable correlations, reflecting the complex lateral dynamics not fully captured in the models. Following seas (150-180°) demonstrate unique bimodal behaviour, with correlations either strongly positive or strongly negative depending on the specific metric and lead time, suggesting threshold effects in surf-riding and broaching conditions. While this analysis includes sea states 0-7, the influence of extreme sea conditions (sea states 8-9) on uncertainty propagation was not extensively investigated. This study forms part of a broader ship weather routing framework, where operational conditions, rather than survival conditions, are of primary interest. Vessels typically avoid or reroute around such extreme conditions rather than attempting to maintain their course.

Third, the temporal evolution of correlations reveals increasing coupling strength with forecast lead time. Short-range forecasts (0-24 h) show moderate correlations (0.65-0.85), while extended-range forecasts (120-168 h) exhibit either very strong (>0.95) or very weak (<0.3) correlations, indicating that uncertainty relationships become more deterministic or completely decouple as forecast skill degrades. The NTPRO 5000 Pierson–Moskowitz implementation demonstrates the most balanced correlation patterns across all conditions.

As illustrated in Table 5 for the 14.5-knot case with wave height uncertainties, RMSE and MAE correlations maintain consistently strong positive relationships (0.223-0.973), with head seas showing the highest values (0.870-0.970). The correlation strength increases systematically with lead time, reaching maximum values at 120-168 hours. Bias correlations reveal physically consistent inverse relationships (-0.713 to -1.000) for head and following seas, where wave height over-prediction leads to speed under-prediction, particularly pronounced in head and following seas. Wind speed correlations for the same Pierson–Moskowitz configuration exhibit more complex patterns. While RMSE and MAE maintain positive correlations (0.259-0.996), they are approximately 8-10 % weaker than corresponding wave height correlations. Notably, bias correlations show predominantly positive values, contrary to wave height patterns, indicating that wind over-prediction coincides with speed over-prediction, a counterintuitive result suggesting limitations in aerodynamic modelling.

The JONSWAP spectrum implementation, although not shown in tables here, produced more polarised correlation patterns. Perfect or near-perfect correlations (±1.000) appeared frequently, particularly at extended lead times, suggesting oversimplified uncertainty relationships. This deterministic behaviour likely stems from the JONSWAP spectrum's assumption of fully developed seas, which may not adequately represent the variable fetch conditions typical of North Atlantic operations. It should be emphasised that these comparative findings between the Pierson-Moskowitz and JONSWAP spectrum implementations are specific to the North Atlantic region, as different ocean basins with distinct fetch characteristics, swell patterns, and storm systems may yield significantly different relative performance between the two spectral formulations. CRPS value correlations are predominantly negative here in the head and following seas (-0.949 to -0.134), suggesting that improved probabilistic wind forecasts actually reduce speed prediction reliability in these conditions. This counterintuitive result may reflect the non-linear aerodynamic effects that are not fully captured in the simulation model.

UGR correlations reveal highly variable relationships ranging from -0.858 to 0.890, with notable transitions between negative and positive values across different conditions and lead times. The strongest positive correlations appear in stern-quartering seas, suggesting that forecast degradation rates are most predictive of speed uncertainty growth when waves approach from the quarter. Conversely, bow-quartering seas show the most extreme negative correlation at short lead times, transitioning to moderate positive values at medium range.

Table 5 Correlation between uncertainty metrics of attainable ship speed $V_{\text{att.}}^{(\text{PM})}$ and uncertainty metrics of significant wave height H_{S} in case of intended ship speed 14.5 kn and sea states 0-7.

		Pearson correlation coefficient $r(x, y)$						
Encounter angles	Lead time (h)	RMSE($V_{\text{att.}}^{(\text{PM})}$), RMSE($H_{\text{S}}$)	MAE($V_{\text{att.}}^{(\text{PM})}$), MAE($H_{\text{S}}$)	Bias($V_{\text{att.}}^{(\text{PM})}$), Bias($H_{\text{S}}$)	UGR($V_{\text{att.}}^{(\text{PM})}$), UGR($H_{\text{S}}$)	CRPS($V_{\text{att.}}^{(\text{PM})}$), CRPS($H_{\text{S}}$)	IoA($V_{\text{att.}}^{(\text{PM})}$), IoA($H_{\text{S}}$)	FSS($V_{\text{att.}}^{(\text{PM})}$), FSS($H_{\text{S}}$)
Head seas	0-24	0.870	0.885	-0.993	-0.663	-0.990	0.888	-0.748
	24-72	0.931	0.937	-0.991	-0.441	-0.992	-0.574	0.882
	72-120	0.970	0.961	-0.971	0.436	0.917	0.523	0.894
	120-168	0.920	0.910	-0.957	0.814	0.841	-0.958	-0.974
Bow-Quartering	0-24	0.872	0.864	-0.996	-0.858	-0.977	-0.025	-0.911
	24-72	0.959	0.955	-0.978	0.380	-0.932	0.973	-0.468
	72-120	0.956	0.937	-0.925	0.657	0.520	-0.502	-0.016
	120-168	0.970	0.965	-0.969	0.183	0.979	0.138	-0.615
Beam	0-24	0.652	0.648	-0.035	0.157	-0.978	-0.809	-0.291
	24-72	0.743	0.742	-0.982	0.544	-0.932	-0.939	0.934
	72-120	0.682	0.589	-0.532	0.718	-0.505	-0.985	0.988
	120-168	0.842	0.795	-0.970	0.559	0.784	-0.997	0.998
Stern-Quartering	0-24	0.223	0.244	-0.752	0.890	-0.982	-0.711	-0.530
	24-72	0.973	0.951	-0.965	0.687	0.859	-0.506	0.742
	72-120	0.973	0.941	-0.978	0.142	1.000	-0.680	0.861
	120-168	0.971	0.954	-0.930	0.742	0.994	-0.927	0.858
Following	0-24	0.928	0.963	-0.713	0.236	-0.985	-0.848	0.946
	24-72	0.944	0.937	-0.980	-0.552	-0.997	-0.949	0.866
	72-120	0.950	0.955	-0.998	-0.068	-0.619	-0.990	0.961
	120-168	0.562	0.538	-1.000	0.176	-0.132	-0.635	-0.906

Head seas demonstrate a systematic evolution from negative correlations at shorter lead times to positive values at extended forecasts, indicating a reversal in the relationship between forecast degradation and speed uncertainty as the forecast horizon extends. Following seas exhibit weak and inconsistent correlations, suggesting that forecast degradation rates have limited predictive value for speed uncertainty in these favourable conditions.

Looking at Table 6, which presents the correlation patterns for the NavCad implementation at 14.5 knots, several distinctive characteristics emerge that differentiate it from the Pierson–Moskowitz model. The NavCad framework exhibits remarkably high correlation coefficients for RMSE and MAE in most conditions, frequently approaching perfect correlation (0.987-1.000), particularly in head seas and bow-quartering angles. This near-deterministic behaviour reflects NavCad's quasi-static resistance calculation methodology, where environmental inputs translate more directly into speed predictions without the dynamic motion effects captured by the time-domain NTPRO simulations. The most interesting feature in Table 6 is the dramatic variation in correlation strength across different encounter angles. Following seas demonstrate particularly anomalous behaviour, with RMSE correlations as low as 0.081 at 0-24 h lead time, jumping to near-perfect correlation (1.000) at 72-120 h. This extreme variability suggests that NavCad's resistance-based approach fails to fully capture the complex dynamics of circumstances that occur in following seas, where both resistance and lateral forces contribute to speed loss in non-linear ways. The beam seas correlations, while stronger than in the Pierson–Moskowitz model, still show the weakest overall values (0.424-0.787 for RMSE), confirming that the quasi-static approach has fundamental limitations in representing lateral hydrodynamic interactions that become dominant when waves approach from abeam. Lower operational ship speeds fundamentally alter correlation structures, as demonstrated in the 12.0-knot analyses (Tables 7 and 8 show selected examples). The Pierson–Moskowitz implementation at 12.0 knots reveals several distinct changes: overall correlation magnitudes decrease by 15-20 % compared to 14.5 knots, negative correlations become more prevalent across all metrics, and encounter angle sensitivity increases dramatically. Beam seas, which showed moderate correlations at higher speeds, exhibit weak or even negative correlations (-0.717 for wave height RMSE) at lower speeds. Comparing correlation patterns across the three modelling frameworks reveals

fundamental differences in uncertainty propagation mechanisms. JONSWAP consistently produces moderate, physically interpretable correlations that vary smoothly with conditions.

Table 6 Correlation between uncertainty metrics of attainable ship speed $V_{\text{att.}}^{(\text{NC})}$ and uncertainty metrics of significant wave height H_{S} in case of intended ship speed 14.5 kn and sea states 0-7.

		Pearson correlation coefficient $r(x, y)$						
Encounter angles	Lead time (h)	RMSE($V_{\text{att.}}^{(\text{NC})}$), RMSE($H_{\text{S}}$)	MAE($V_{\text{att.}}^{(\text{NC})}$), MAE($H_{\text{S}}$)	Bias($V_{\text{att.}}^{(\text{NC})}$), Bias($H_{\text{S}}$)	UGR($V_{\text{att.}}^{(\text{NC})}$), UGR($H_{\text{S}}$)	CRPS($V_{\text{att.}}^{(\text{NC})}$), CRPS($H_{\text{S}}$)	IoA($V_{\text{att.}}^{(\text{NC})}$), IoA($H_{\text{S}}$)	FSS($V_{\text{att.}}^{(\text{NC})}$), FSS($H_{\text{S}}$)
Head seas	0-24	0.997	0.997	-0.999	-0.059	-0.993	0.933	-0.784
	24-72	0.994	0.993	-1.000	0.480	-0.985	-0.687	0.882
	72-120	0.996	0.996	-0.999	0.797	0.916	0.578	0.941
	120-168	0.991	0.987	-0.995	0.990	0.971	-0.876	0.997
Bow-Quartering	0-24	0.997	0.997	-0.969	0.511	-0.995	0.044	-0.956
	24-72	0.999	0.999	-0.850	0.953	-0.977	0.969	-0.662
	72-120	0.988	0.988	-0.994	0.959	0.494	-0.579	-0.143
	120-168	0.987	0.990	-0.994	0.983	0.977	0.173	-0.747
Beam	0-24	0.757	0.787	0.828	-0.016	-0.996	-0.709	-0.343
	24-72	0.656	0.711	-0.961	0.470	-0.966	-0.857	0.910
	72-120	0.424	0.422	-0.634	0.762	-0.441	-0.947	0.974
	120-168	0.450	0.466	-0.855	0.913	-0.045	-0.968	0.980
Stern-Quartering	0-24	0.793	0.935	-0.980	0.904	-0.992	-0.663	-0.492
	24-72	0.695	0.930	-0.820	0.598	-0.901	-0.436	0.740
	72-120	0.998	0.999	-0.954	0.558	0.903	-0.639	0.848
	120-168	0.998	0.999	-0.940	0.957	0.966	-0.911	0.863
Following	0-24	0.081	0.153	0.132	0.709	-0.988	-0.844	0.744
	24-72	0.928	0.916	-0.890	0.303	-0.985	-0.943	0.906
	72-120	1.000	1.000	-0.997	0.853	0.981	-0.988	0.998
	120-168	1.000	1.000	-0.994	0.720	0.980	-0.676	-0.902

Table 7 Correlation between uncertainty metrics of attainable ship speed $V_{\text{att.}}^{(\text{PM})}$ and uncertainty metrics of significant wave height H_{S} in case of intended ship speed 12.0 kn and sea states 0-7.

		Pearson correlation coefficient $r(x, y)$						
Encounter angles	Lead time (h)	RMSE($V_{\text{att.}}^{(\text{PM})}$), RMSE($H_{\text{S}}$)	MAE($V_{\text{att.}}^{(\text{PM})}$), MAE($H_{\text{S}}$)	Bias($V_{\text{att.}}^{(\text{PM})}$), Bias($H_{\text{S}}$)	UGR($V_{\text{att.}}^{(\text{PM})}$), UGR($H_{\text{S}}$)	CRPS($V_{\text{att.}}^{(\text{PM})}$), CRPS($H_{\text{S}}$)	IoA($V_{\text{att.}}^{(\text{PM})}$), IoA($H_{\text{S}}$)	FSS($V_{\text{att.}}^{(\text{PM})}$), FSS($H_{\text{S}}$)
Head seas	0-24	0.794	0.780	-0.998	-0.979	-0.982	0.730	-0.921
	24-72	0.872	0.883	-0.992	-0.577	-0.998	-0.683	-0.773
	72-120	0.947	0.937	-0.992	0.031	0.773	0.474	-0.583
	120-168	0.781	0.753	-0.972	0.452	0.315	-0.917	0.793
Bow-Quartering	0-24	0.672	0.615	-0.998	-0.605	-0.968	-0.038	-0.910
	24-72	0.902	0.872	-0.961	0.150	-0.848	0.964	-0.794
	72-120	0.873	0.802	-0.906	0.029	-0.809	-0.493	-0.147
	120-168	0.929	0.933	-0.958	-0.670	0.841	0.063	-0.748
Beam	0-24	-0.717	-0.695	0.732	0.248	-0.962	-0.945	0.008
	24-72	-0.627	-0.683	-0.927	0.301	-0.994	-0.980	0.995
	72-120	-0.490	-0.534	-0.107	0.666	-0.736	-0.948	0.923
	120-168	-0.354	-0.431	-0.754	0.438	-0.595	-0.956	0.965
Stern-Quartering	0-24	-0.412	-0.350	-0.746	0.608	-0.993	-0.895	-0.278
	24-72	0.878	0.775	0.933	0.975	0.957	-0.748	0.842
	72-120	0.980	0.925	-0.996	0.422	0.999	-0.884	0.942
	120-168	0.989	0.971	-0.919	0.902	0.999	-0.894	0.764
Following	0-24	0.220	0.328	-0.374	0.201	-0.994	-0.873	0.897
	24-72	0.699	0.576	0.926	0.019	-0.961	-0.956	0.913
	72-120	0.844	0.818	-0.999	0.237	-0.332	-0.992	0.987
	120-168	0.566	0.624	-0.998	0.308	0.140	-0.685	-0.908

Table 8 Correlation between uncertainty metrics of attainable ship speed $V_{\text{att.}}^{(\text{NC})}$ and uncertainty metrics of significant wave height H_{S} in case of intended ship speed 12.0 kn and sea states 0-7.

		Pearson correlation coefficient $r(x, y)$						
Encounter angles	Lead time (h)	RMSE($V_{\text{att.}}^{(\text{NC})}$), RMSE($H_{\text{S}}$)	MAE($V_{\text{att.}}^{(\text{NC})}$), MAE($H_{\text{S}}$)	Bias($V_{\text{att.}}^{(\text{NC})}$), Bias($H_{\text{S}}$)	UGR($V_{\text{att.}}^{(\text{NC})}$), UGR($H_{\text{S}}$)	CRPS($V_{\text{att.}}^{(\text{NC})}$), CRPS($H_{\text{S}}$)	IoA($V_{\text{att.}}^{(\text{NC})}$), IoA($H_{\text{S}}$)	FSS($V_{\text{att.}}^{(\text{NC})}$), FSS($H_{\text{S}}$)
Head seas	0-24	0.973	0.972	-0.997	-0.135	-0.998	0.797	-0.831
	24-72	0.997	0.996	-0.999	0.964	-0.945	-0.757	-0.851
	72-120	1.000	1.000	-0.999	0.916	0.978	0.581	0.041
	120-168	0.999	0.998	-0.992	0.989	0.994	-0.999	0.989
Bow-Quartering	0-24	0.988	0.986	-0.971	0.294	-0.998	-0.023	-0.938
	24-72	0.992	0.995	-0.991	0.954	0.840	0.957	-0.826
	72-120	0.998	0.998	-0.982	0.995	0.983	-0.561	-0.176
	120-168	1.000	1.000	-0.998	0.994	0.995	0.048	-0.761
Beam	0-24	0.933	0.946	0.731	0.535	-0.995	-0.945	0.013
	24-72	0.947	0.959	-0.977	0.859	-0.829	-0.982	0.993
	72-120	0.928	0.908	-0.971	0.705	0.032	-0.947	0.925
	120-168	0.957	0.940	-0.998	0.818	0.816	-0.956	0.966
Stern-Quartering	0-24	0.973	0.967	-0.990	0.542	-1.000	-0.894	-0.250
	24-72	1.000	1.000	0.978	0.715	0.838	-0.749	0.850
	72-120	0.980	0.983	-0.971	0.448	0.944	-0.884	0.940
	120-168	0.980	0.977	-0.874	0.795	0.951	-0.891	0.761
Following	0-24	0.990	0.991	-0.950	0.120	-0.922	-0.871	0.787
	24-72	0.976	0.977	-0.997	-0.079	-0.879	-0.963	0.919
	72-120	0.969	0.979	-0.993	0.363	0.944	-0.998	0.999
	120-168	0.991	0.991	-0.974	0.991	0.958	-0.658	-0.913

JONSWAP generates more extreme, often deterministic relationships that may oversimplify complex ship-wave interactions. NavCad correlations typically fall between these extremes but show distinctive stepped patterns reflecting its quasi-static calculation approach. The analysis also reveals that probabilistic metrics (CRPS, FSS) often show inverse correlations compared to deterministic metrics (RMSE, MAE), particularly in complex sea states. This suggests that improved probabilistic weather forecast skill does not necessarily translate to better ship speed predictions, highlighting the need for specialised uncertainty quantification methods in marine applications.

These correlation patterns have direct implications for voyage planning and weather routing systems. Strong positive correlations in head seas justify simple linear uncertainty propagation methods, while weak or variable correlations in beam seas require more sophisticated Monte Carlo approaches. The speed-dependent correlation structures indicate that uncertainty models must be configured differently for slow-steaming versus normal operations. The increasing correlation strength with lead time suggests that long-range routing decisions are paradoxically more sensitive to weather forecast quality than short-range tactical adjustments, contrary to common operational assumptions.

3.3 The Uncertainty of ETA

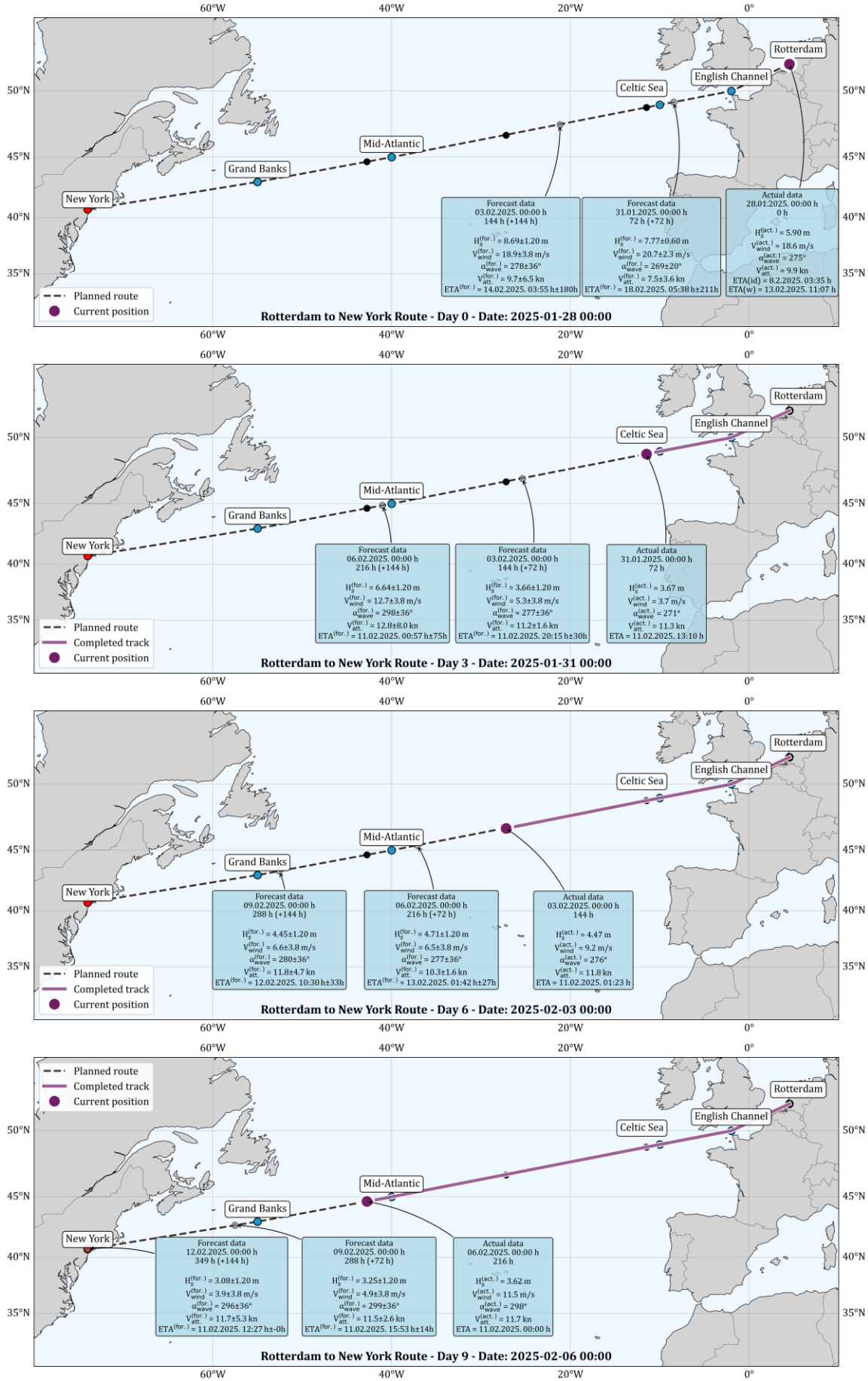
To demonstrate the practical application of the uncertainty quantification framework developed in this study, a transatlantic voyage from Rotterdam to New York was selected as a representative case study (shown in Figure 8). This route was chosen for several compelling reasons: it represents one of the most commercially significant shipping corridors globally, traverses the North Atlantic where comprehensive weather forecast data from NOAA GFS is readily available, experiences diverse meteorological conditions ranging from sheltered waters in the English Channel to severe open-ocean storms, and provides sufficient voyage duration (approximately 10-12 days) to observe the full evolution of forecast uncertainty from short-range (0-24 h) through extended-range (120-168 h) predictions. The NTPRO 5000 Pierson–Moskowitz spectrum implementation was selected for this analysis based on its demonstrated balance between accuracy and physical consistency, as evidenced by the correlation analysis in Section 3.2. Unlike the JONSWAP spectrum,

which showed extreme polarisation in uncertainty propagation, or NavCad, which exhibited stepped uncertainty patterns, the Pierson-Moskowitz model provided smooth, physically interpretable uncertainty growth that aligns with observed meteorological forecast degradation patterns. Pseudo-code of this integrated framework for attainable ship speed uncertainty quantification is given in Appendix 1. The pipeline and key computational steps comprise: (*i*) initialisation, (*ii*) weather forecast processing, (*iii*) ship performance lookup tables, (*iv*) uncertainty quantification, (*v*) correlation analysis, and (*vi*) voyage planning with uncertainty.

The start of the voyage simulation was set for 28 January 2025, at 00:00 UTC, deliberately chosen during the North Atlantic winter season when weather variability and forecast uncertainty are at their maximum, providing a robust test of the uncertainty quantification framework. The actual voyage length of 3880 nm and duration of 337.0 hours, i.e. 14 days and 1 hour, arriving on 11 February 2025, at 01:00 UTC. This represents a 20.6 % increase over the nominal 267.6-hour duration at intended speed 14.5 kn, confirming systematic weather-induced speed loss. It should be noted that the voyage simulation was set to start 10 nm from Rotterdam port and finish 10 nm before New York port in order to avoid low-speed manoeuvring near ports. In Figure 8, four sequential snapshots of the voyage progress at 3-day intervals (Days 0, 3, 6, and 9), each displaying the complex relationship between actual conditions and forecast projections. The visualisations employ an information architecture where the purple track indicates completed voyage segments, the current ship position is marked with a prominent purple circle, and information boxes connected by arrows display both actual and forecasted states with their associated uncertainties. Each information box follows a structured format presenting temporal information (date/time and hours elapsed), meteorological conditions with uncertainties ($H_s \pm \Delta H_s$, $V_{wind} \pm \Delta V_{wind}$, $\alpha_{wave} \pm \Delta\alpha_{wave}$), resulting attainable speed with uncertainty ($V_{att.} \pm \Delta V_{att.}$) and estimated time of arrival with temporal uncertainty (ETA ± ΔETA in hours). The uncertainty values (±) represent Mean Absolute Error (MAE) as derived from the uncertainty metrics framework described in Section 2.5.

Table 9 synthesises the actual and forecasted conditions at 3-day intervals throughout the voyage, following a systematic pattern of $(3k-3,\ 3k,\ 3k+3)$ days where $k = 1,2,...,5$. This mathematical structure creates a rolling forecast window where, for each value of k, Day $3k-3$ represents the current observation point, Day $3k$ represents the next forecast point 3 days ahead, and Day $3k+3$ represents the next forecast point 6 days ahead. For instance, when $k = 2$, the pattern yields Days 3, 6, and 9, where Day 3 contains actual observed conditions (H_s = 3.67 m, V_{wind} = 3.70 m/s, $V_{att.}$ = 11.3 kn), Day 6 shows what was forecasted 3 days ahead from Day 3, and Day 9 shows what was forecasted 6 days ahead from Day 3, as also shown in Figure 8(*b*). Uncertainty propagation from Table 9 for significant wave height (H_S), attainable ship speed ($V_{att.}$) and ETA in case of k = 1, 2, 3 and 4, is also visualised in Figure 9, which elegantly shows how uncertainty decreases as the ship sails towards the destination port. This structure elegantly captures how each location along the route receives multiple forecasts from different lead times, first appearing as $3k-3$ (actual observations), then as $3k$ (a 3-day forecast), and finally as $3k+3$ (a 6-day forecast). The data from Table 9 and for k = 1, 2, 3 and 4, corresponds to Figures 8(*a*)-8(*d*), respectively. The uncertainty values (±) associated with each forecast demonstrate how prediction confidence changes with lead time. For instance, the ETA uncertainty varies dramatically from ±211 hours, for the initial long-range forecast at Day $3k$ for k = 1, to ±14 hours for short-range forecast at Day $3k$ for k = 4.

This pattern of uncertainty evolution shown in Table 9 suggests that uncertainty growth is not simply a function of forecast lead time but is modulated by spatial variability in forecast skill and environmental predictability. The average actual attainable ship speed of 11.38 knots, calculated from actual values, with an MAE of 1.45 knots, demonstrates substantial variability around the mean performance. The 21.5 % speed reduction from the intended 14.5 knots aligns with typical winter North Atlantic conditions, where average significant wave height of 3.94 m (maximum 5.90 m) and wind speeds averaging 10.78 m/s (maximum 18.6 m/s) create persistent adverse conditions. The cumulative uncertainty growth from ±211 to ±215.28 hours over the voyage duration represents the integrated effect of speed variations, substantially lower than the initial projections but still significant for operational planning.



(*a*)

(*b*)

(*c*)

(*d*)

Fig. 8 The visualised route from Rotterdam to New York, with actual weather data and weather forecasts along the route

Table 9 Forecasted and actual weather variables values, for every 3 days of the voyage, along with the ETA uncertainty

Time frame		Actual state				Forecasted state			
Day	Date Time (d.m.y. h:m h)	H_S (m)	V_{wind} (m/s)	$V_{att.}$ (kn)	ETA (d.m.y. h:m h)	$H_S + \Delta H_S$ (m)	$V_{wind}+\Delta V_{wind}$ (m/s)	$V_{att.}+\Delta V_{att.}$ (kn)	ETA + ΔETA (d.m.y. h:m h)
0	28.1.2025. 00:00 h	5.90	18.6	9.87	8.2.2025. 03:35 h (ideal) 13.2.2025. 11:07 h (weather-adjusted)	-	-	-	-
3	31.1.2025. 00:00 h	-	-	-	-	7.77±0.60	20.70±2.30	7.50±3.60	18.02.2025. 5:38 h ±211 h
6	3.2.2025. 00:00 h	-	-	-	-	8.69±1.20	18.90±3.80	9.70±6.50	14.02.2025. 3:55 h ±180h
3	31.1.2025. 00:00 h	3.67	3.70	11.3	11.2.2025. 13:10 h	-	-	-	-
6	3.2.2025. 00:00 h	-	-	-	-	3.66±1.20	5.30±3.80	11.20±3.60	11.02.2025. 20:15 h ±30 h
9	6.2.2025. 00:00 h	-	-	-	-	6.64±1.20	12.70±3.80	12.80±8.00	11.02.2025. 00:57 h ±75 h
6	3.2.2025. 00:00 h	4.47	9.2	11.8	11.2.2025. 01:23 h	-	-	-	-
9	6.2.2025. 00:00 h	-	-	-	-	4.71±1.20	6.5±3.80	10.30±1.60	13.02.2025. 01:42 h ±27 h
12	9.2.2025. 00:00 h	-	-	-	-	4.45±1.20	6.6±3.80	11.80±4.70	12.02.2025. 10:30 h ±33 h
9	6.2.2025. 00:00 h	3.62	11.5	11.7	11.2.2025. 04:30 h	-	-	-	-
12	9.2.2025. 00:00 h	-	-	-	-	3.25±1.20	4.9±3.80	11.50±2.60	11.02.2025. 15:53 h ±14 h
15	12.2.2025. 00:00 h	-	-	-	-	3.08±1.20	3.9±3.80	11.70±5.30	11.02.2025. 12:27 h ±00 h
12	9.2.2025. 00:00 h	2.04	10.90	12.23	11.02.2025. 02:08 h	-	-	-	-
15	12.2.2025. 00:00 h	-	-	-	-	2.15±0.11	18.93±8.03	11.39±3.89	11.02.2025. 05:49 h ±18h

Attainable speed uncertainty varied non-linearly from ±1.60 to ±8.00 knots, reflecting the complex transformation of environmental uncertainties through ship performance models. The ±8.00 kn uncertainty observed at a certain point (12.80±8.00 knots at Day $3k+3$ for $k = 2$) represents the statistical MAE for that specific forecast configuration (lead time: 72-120 h, Day 3 to Day 9; sea state at that location with associated encounter wave angle conditions), without imposing physical constraints on the vessel's propulsion system. This purely statistical approach yields a theoretical speed range of 4.80 to 20.80 knots, which extends beyond the vessel's actual operating envelope of 0 to 14.5 knots (maximum intended speed).

This unbounded statistical quantification was deliberately employed in this study to capture the full magnitude of forecast uncertainty propagation. For practical operational applications, however, the uncertainty model should incorporate physical constraints through:

a. Engine power limitations: Capping maximum speed at 14.5 knots based on installed power
b. Minimum steerage speed: Setting a lower bound of 3-4 knots for maintaining directional control
c. Truncated distributions: Implementing bounded probability distributions that respect these physical limits while preserving the underlying uncertainty structure.

The large uncertainty values observed (particularly the ±8.00 knots) therefore serve as indicators of high forecast volatility rather than literal speed ranges. They highlight periods where environmental conditions are highly uncertain, signalling to operators that speed predictions during these periods have low confidence. For decision support systems, these high uncertainty periods would trigger risk-based planning protocols, even though the actual speed must remain within physical bounds. This distinction between statistical uncertainty quantification (used for analysis) and operationally bounded uncertainty (required for implementation) is essential for using these research findings in practical voyage planning tools.

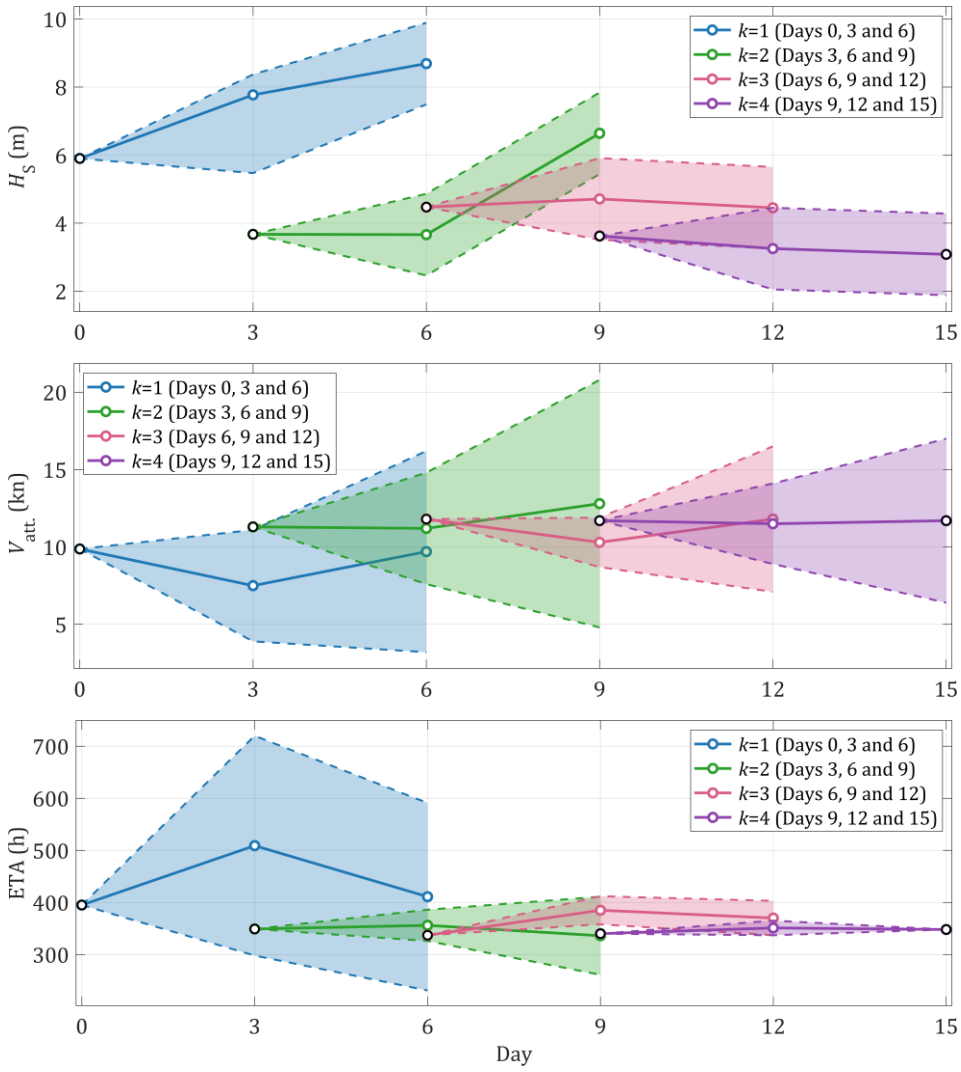


Fig. 9 Visualisation of uncertainty propagation for Hs, Vatt. and ETA in case k = 1, 2, 3 and 4

The probability distributions presented in Figure 10 show important insights about the uncertainty quantification framework when examined against the actual voyage planning outcome. The Pierson-Moskowitz model, which was selected for this voyage planning analysis, predicted a mean duration of 347.3 hours with a 95 % confidence interval spanning from 302 to 375 hours. While the actual voyage duration of 336.0 hours falls within this distribution, approximately 11 hours below the predicted mean, the framework's absolute calibration presents notable challenges that warrant careful examination.

The discrepancy between the predicted and actual arrival times exposes fundamental limitations in maritime uncertainty quantification. Despite the Pierson-Moskowitz model's reasonable central tendency prediction, the uncertainty bounds appear to be mis-calibrated in their temporal translation. This miscalibration likely stems from several interconnected factors. The models demonstrate a tendency toward conservative speed estimates, suggesting a systematic overweighting of adverse weather probabilities in the uncertainty framework. Furthermore, the transformation process from meteorological variables to ship speed predictions may not adequately capture all sources of variability, particularly the complex non-linear interactions between environmental conditions and vessel performance. The assumption of temporal independence between voyage segments also merits scrutiny, as correlations in weather patterns and cumulative effects on vessel performance could lead to overestimation of aggregate uncertainty.

Comparing the three models, the Pierson-Moskowitz implementation produces a slightly wider distribution than both JONSWAP and NavCad, with its 95 % confidence interval spanning approximately 73 hours. This broader uncertainty range reflects the model's more conservative approach to capturing fully developed sea states characteristic of North Atlantic conditions.

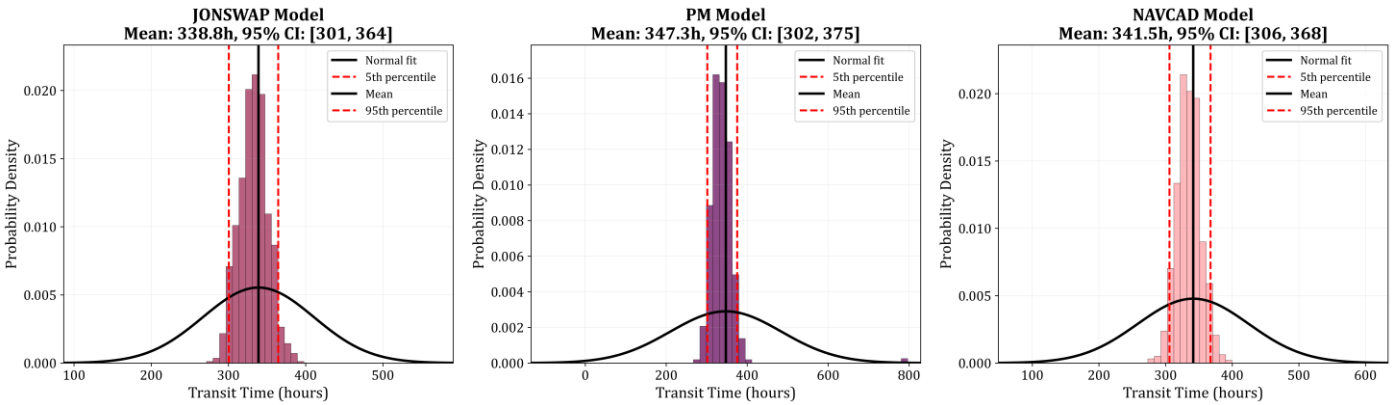


Fig. 10 The probability distributions of voyage time and ETA for the chosen case study

The convergence of all three models toward similar mean values (ranging from 338.8 to 347.3 hours) suggests that extended voyage predictions tend toward climatological averages, though this convergence may inadvertently mask the true probability of extreme events that could significantly impact voyage duration.

These findings emphasise that while the framework successfully captures relative uncertainty patterns and their evolution throughout the voyage, achieving accurate absolute calibration remains an ongoing challenge in marine operations. The practical implication is that operational systems should prioritise relative risk assessment over firm adherence to statistical confidence bounds. The implications of our findings also align with risk factors identified in [38], where it was found that weather-related delays and perils of the sea constitute high-risk areas for general cargo ship operators. Our uncertainty quantification framework provides a quantitative basis for managing these risks through improved voyage planning. Continuous recalibration, based on accumulated voyage data and observed outcomes, will be crucial for enhancing the reliability of future predictions and improving the framework's operational efficiency in real-world voyage planning applications.

4. Conclusions

This study has developed and validated a comprehensive data-driven framework for quantifying attainable ship speed uncertainty under stochastic weather conditions. Through 2,028 simulations and a transatlantic case study, we established that wave height uncertainties show the strongest correlations with speed prediction errors (0.65-0.97), confirming their dominant role in ship speed loss mechanisms. The comparative assessment of three modelling approaches revealed that the NTPRO 5000 Pierson-Moskowitz implementation provides the most balanced uncertainty propagation, with physically interpretable correlation patterns that vary smoothly with environmental conditions. However, the Rotterdam to New York case study exposed critical calibration challenges: the 95% confidence interval failed to capture the actual arrival time (the ship arriving 35 hours after the lower bound for P-M spectrum ETA estimation), and unbounded statistical uncertainties reached ±8.00 knots, exceeding the vessel's 14.5-knot maximum speed. These findings highlight the distinction between statistical uncertainty quantification for analysis and operationally bounded uncertainty for practical implementation.

The framework demonstrates significant operational value despite these challenges. The systematic evolution of ETA uncertainty, from ±211 hours at departure to ±27 hours mid-voyage, provides mariners with quantitative risk assessments for voyage planning. The non-monotonic uncertainty pattern observed, where uncertainty peaked during mid-Atlantic transit rather than increasing linearly with time, reveals that forecast skill varies spatially and depends on prevailing weather systems. All three models predicted voyage durations 0.5-3 % longer than nominal, confirming systematic weather-induced speed loss in winter North Atlantic conditions. For practical applications, the framework proves more suitable for relative risk assessment than absolute uncertainty bounds, suggesting that operational systems should focus on identifying periods of high uncertainty for enhanced monitoring rather than relying solely on statistical confidence intervals.

Several limitations constrain the current framework's applicability. The exclusion of ocean currents, reliance on normal distributions that may inadequately capture extreme events, and validation limited to a single vessel type all require addressing in future implementations. Additionally, the advantage of the Pierson-Moskowitz spectrum over JONSWAP for uncertainty propagation observed in this study is specific to North Atlantic conditions and may not hold for other ocean regions where different wave generation mechanisms and fetch conditions prevail. Most critically, the unbounded statistical approach must be replaced with physically constrained models that respect engine power limitations (maximum 14.5 knots) and minimum steerage requirements (3-4 knots). The systematic over-prediction of adverse conditions observed in the case study also indicates the need for bias correction mechanisms in operational deployment. One limitation of this study is also that while the theoretical framework treats peak wave period as an independent variable affecting attainable ship speed, the practical implementation constrains it to be a dependent quantity coupled to significant wave height through the Pierson-Moskowitz and JONSWAP spectral formulations. This simplification, though physically justified for the North Atlantic routes where wind seas predominate and swell typically aligns with local wind waves, may not adequately capture speed reductions in regions with complex wave climates featuring significant swell-wind sea misalignment or where period-height relationships deviate from standard spectral assumptions.

Future research should prioritise three key areas. First, implementing bounded uncertainty distributions through truncated normal or beta distributions would ensure physically realistic speed predictions while preserving uncertainty information. Second, integrating ensemble weather forecasts and machine learning approaches could capture non-linear relationships between weather patterns and ship performance, potentially using LSTM networks or Transformer architectures with interpretability mechanisms. Third, developing real-time adaptation capabilities through online learning would enable continuous refinement of uncertainty estimates based on observed voyage data, addressing the calibration issues identified in this study.

In conclusion, this research provides a robust foundation for uncertainty-aware maritime operations, despite revealing significant challenges in absolute uncertainty quantification. The transition from deterministic to probabilistic voyage planning represents a fundamental shift in maritime operational philosophy, one that acknowledges the inherent uncertainties in ocean navigation while providing actionable insights for their management. As the shipping industry advances toward autonomous operations and faces increasingly stringent environmental regulations, the framework developed here offers essential tools for risk-based decision-making. The key insight is not that we can perfectly predict uncertainty, but that we can systematically quantify and communicate it, enabling more informed decisions even when our confidence bounds prove imperfect.

ACKNOWLEDGMENTS

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

REFERENCES

[1] Prpić-Oršić, J., Kenji, S., Valčić, M., Faltinsen, O.M., 2020. Uncertainties of ship speed loss evaluation under real weather conditions. *Journal of Offshore Mechanics and Arctic Engineering*, 142(3), 031106. https://doi.org/10.1115/1.4045790

[2] Dalheim, L., Steen, S., 2020. Added resistance and speed loss of a ship found using onboard monitoring data. *Journal of Ship Research*, 64(2), 99-117. https://doi.org/10.5957/jsr.2020.64.2.99

[3] Toman, I., Vidan, P., Barić, M., Grbić, L., 2020. Realism of Transas NTPRO 5000 radar simulation in search and rescue training exercises. *Transactions on Maritime Science*, 9(1), 99-105. https://doi.org/10.7225/toms.v09.n01.009

[4] 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. https://doi.org/10.3390/w15203584

[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. https://doi.org/10.3390/jmse13061185

[6] Wu, M., Stefanakos, C., Gao, Z., Haver, S., 2019. Prediction of short-term wind and wave conditions for marine operations using a multi-step-ahead decomposition-ANFIS model and quantification of its uncertainty. *Ocean Engineering*, 188, 106300. https://doi.org/10.1016/j.oceaneng.2019.106300

[7] Ormevik, J.H., 2023. How weather conditions greatly influence emissions from offshore logistics. SINTEF Blog - Energy. https://blog.sintef.com/energy/how-weather-conditions-greatly-influence-emissions-from-offshore-logistics/ (accessed 7th July 2025)

[8] Wang, J., Bielicki, S., Kluwe, F., Orihara, H., Xin, G., Kume, K., Oh, S., Liu, S., Feng, P., 2021. Validation study on a new semi-empirical method for the prediction of added resistance in waves of arbitrary heading in analyzing ship speed trial results. *Ocean Engineering*, 240, 109959. https://doi.org/10.1016/j.oceaneng.2021.109959

[9] 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. https://doi.org/10.1016/j.oceaneng.2020.107494

[10] 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. https://doi.org/10.1016/j.oceaneng.2017.06.051

[11] Kim, Y.-R., Esmailian, E., Steen, S., 2022. A meta-model for added resistance in waves. *Ocean Engineering*, 266, 112749. https://doi.org/10.1016/j.oceaneng.2022.112749

[12] Korkmaz, K.B., Werner, S., Bensow, R., 2021. Verification and validation of CFD based form factors as a combined CFD/EFD method. *Journal of Marine Science and Engineering*, 9(1), 75. https://doi.org/10.3390/jmse9010075

[13] Liu, S., Papanikolaou, A., 2020. Regression analysis of experimental data for added resistance in waves of arbitrary heading and development of a semi-empirical formula. *Ocean Engineering*, 206, 107357. https://doi.org/10.1016/j.oceaneng.2020.107357

[14] 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. https://doi.org/10.1016/j.oceaneng.2019.106758

[15] Nas, S., Zorba, Y., Ucan, E., 2014. The Mooring Pattern Study for Q-Flex Type LNG Carriers Scheduled for Berthing at Ege Gaz Aliaga LNG Terminal. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation*, 8(4), 543-548. https://doi.org/10.12716/1001.08.04.08

[16] Aydogdu, Y.V., 2022. Utilization of full-mission ship-handling simulators for navigational risk assessment: A case study of large vessel passage through the Istanbul Strait. *Journal of Marine Science and Engineering*, 10(5), 659. https://doi.org/10.3390/jmse10050659

[17] Nishizaki, C., Okazaki, T., Yabuki, H., Yoshimura, Y., 2019. Simulation study on the Influence of EEDI Requirements to Shiphandling in Heavy Weather. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation*, 13(4), 855-860. https://doi.org/10.12716/1001.13.04.19

[18] 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. https://doi.org/10.21278/brod67206

[19] Liu, L., Chen, M., Wang, X., Zhang, Z., Yu, J., Feng, D., 2021. CFD prediction of a full-scale ship parametric roll in head waves. *Ocean Engineering*, 233, 109180. https://doi.org/10.1016/j.oceaneng.2021.109180

[20] Gatin, I., Boxall, D., 2021. Calculating Speed Loss Due to Swell using CFD. *6th Hull Performance & Insight Conference*, Pontignano, Italy.

[21] Martić, I., Anušić, B., Degiuli, N., Grlj, C. G., 2024. Numerically Investigating the Effect of Trim on the Resistance of a Container Ship in Confined and Shallow Water. *Applied Sciences*, 14(15), 6570. https://doi.org/10.3390/app14156570

[22] Ntouras, D., Papadakis, G., Belibassakis, K., 2022. Ship bow wings with application to trim and resistance control in calm water and in waves. *Journal of Marine Science and Engineering*, 10(4), 492. https://doi.org/10.3390/jmse10040492

[23] Czaplewski, K., Swierczynski, S., Zwolan, P., 2021. Increasing the reliability of simulation tests in navigation and maneuvering simulators using the k-ε model based on the RANS method. *Sensors*, 21(15), 4995. https://doi.org/10.3390/s21154995

[24] Shin, G.-H., Yang, H., 2025. Deep reinforcement learning for integrated vessel path planning with safe anchorage allocation. *Brodogradnja*, 76(3), 76305. https://doi.org/10.21278/brod76305

[25] Zhang, Y., Zhang, J., Guo, Z., Zhang, L., Shang, Y., Chen, W., 2025. The study on dynamic modeling and path planning of a manta ray-inspired underwater glider. *Brodogradnja*, 76(2), 76208. https://doi.org/10.2139/ssrn.4706205

[26] 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. https://doi.org/10.21278/brod76206

[27] Alexiou, K., Pariotis, E.G., Leligou, H.C., Zannis, T.C., 2022. Towards data-driven models in the prediction of ship performance (speed-power) in actual seas: A comparative study between modern approaches. *Energies*, 15(16), 6094. https://doi.org/10.3390/en15166094

[28] Mittendorf, M., Nielsen, U.D., Bingham, H.B., 2022. Data-driven prediction of added-wave resistance on ships in oblique waves – A comparison between tree-based ensemble methods and artificial neural networks. *Applied Ocean Research*, 118, 102964. https://doi.org/10.1016/j.apor.2021.102964

[29] Mittendorf, M., Nielsen, U.D., Bingham, H.B., Dietz, J., 2023. Assessment of added resistance estimates based on monitoring data from a fleet of container vessels. *Ocean Engineering*, 272, 113892. https://doi.org/10.1016/j.oceaneng.2023.113892

[30] 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. https://doi.org/10.21278/brod76301

[31] Moreira, L., Vettor, R., Guedes Soares, C., 2021. Neural network approach for predicting ship speed and fuel consumption. *Journal of Marine Science and Engineering*, 9(2), 119. https://doi.org/10.3390/jmse9020119

[32] Gan, W., Ma, D., Duan, Y., 2025. Comparisons of machine learning methods in ship speed prediction based on shipboard observation. *Journal of Marine Science and Engineering*, 13(6), 1011. https://doi.org/10.3390/jmse13061011

[33] Guo, Y., Wang, Y., Zhou, J., Wang, J., 2023. Domain-adapted feature transfer: A generalized framework for short-term vessel speed prediction. *Ocean Engineering*, 280, 114536. https://doi.org/10.1016/j.oceaneng.2023.114536

[34] Tarelko, W., Rudzki, K., 2020. Applying artificial neural networks for modelling ship speed and fuel consumption. *Neural Computing and Applications*, 32, 17379-17395. https://doi.org/10.1007/s00521-020-05111-2

[35] Bassam, A.M., Phillips, A.B., Turnock, S.R., Wilson, P.A., 2022. Ship speed prediction based on machine learning for efficient shipping operation. *Ocean Engineering*, 245, 110449. https://doi.org/10.1016/j.oceaneng.2021.110449

[36] Coraddu, A., Oneto, L., Baldi, F., Cipollini, F., Atlar, M., Savio, S., 2019. Data-driven ship digital twin for estimating the speed loss caused by the marine fouling. *Ocean Engineering*, 186, 106063. https://doi.org/10.1016/j.oceaneng.2019.05.045

[37] Tarovik, O., Eremenko, D., Topaj, A., 2024. A benchmark study on ship speed prediction models in Arctic conditions: machine learning, process-based and hybrid approaches. *Ocean Engineering*, 311, 118884. https://doi.org/10.1016/j.oceaneng.2024.118884

[38] Ding, J.-F., Tseng, W.-J., Sung, Y.-J., 2024. An evaluation of operational risks for general cargo ship operators. Brodogradnja, 75(1), 75101. https://doi.org/10.21278/brod75101

[39] Vettor, R., Guedes Soares, C., 2022. Reflecting the uncertainties of ensemble weather forecasts on the predictions of ship fuel consumption. *Ocean Engineering*, 250, 111009. https://doi.org/10.1016/j.oceaneng.2022.111009

[40] Jeuring, J., Samuelsen, E. M., Lamers, M., Müller, M., Hjøllo, B. Å., Bertino, L., Hagen, B., 2024. Map-Based Ensemble Forecasts for Maritime Operations: An Interactive Usability Assessment with Decision Scenarios. *Weather, Climate, and Society*, 16(1), 235-256. https://doi.org/10.1175/WCAS-D-23-0076.1

[41] 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. https://doi.org/10.1016/j.ejor.2022.08.006

[42] Wu, M., Stefanakos, C., Gao, Z., Haver, S., 2019. Prediction of short-term wind and wave conditions for marine operations using a multi-step-ahead decomposition-ANFIS model and quantification of its uncertainty. *Ocean Engineering*, 188, 106300. https://doi.org/10.1016/j.oceaneng.2019.106300

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

[44] 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. https://doi.org/10.3390/jmse9121434

[45] Chen, Y., Zhang, C., Guo, Y., Wang, Y., Lang, X., Zhang, M., Mao, W., 2025. State-of-the-art optimization algorithms in weather routing – Ship decision support systems: Challenge, taxonomy, and review. *Ocean Engineering*, 331, 121198. https://doi.org/10.1016/j.oceaneng.2025.121198

[46] Gao, J., Zhang, Y., 2024. Ship collision avoidance decision-making research in coastal waters considering uncertainty of target ships. *Brodogradnja*, 75(2), 75203. https://doi.org/10.21278/brod75203

[47] Zhou, P., Zhou, Z., Wang, Y., Wang, H., 2022. Ship weather routing based on hybrid genetic algorithm under complicated sea conditions. *Journal of Ocean University of China*, 22(1), 28-42. https://doi.org/10.1007/s11802-023-5002-1

[48] Dębski, R., Dreżewski, R., 2024. Multi-objective ship route optimisation using estimation of distribution algorithm. *Applied Sciences*, 14(13), 5919. https://doi.org/10.3390/app14135919

[49] Wei, Q., Liu, Y., Dong, Y., Li, T., Li, W., 2023. A digital twin framework for real-time ship routing considering decarbonization regulatory compliance. *Ocean Engineering*, 278, 114407. https://doi.org/10.1016/j.oceaneng.2023.114407

[50] Latinopoulos, C., Zavvos, E., Kaklis, D., Leemen, V., Halatsis, A., 2025. Marine Voyage Optimization and Weather Routing with Deep Reinforcement Learning. *Journal of Marine Science and Engineering*, 13(5), 902. https://doi.org/10.3390/jmse13050902

[51] Liu, Q., Wang, Y., Zhang, R., Yan, H., Xu, J., Guo, Y., 2023. Arctic weather routing: a review of ship performance models and ice routing algorithms. *Frontiers in Marine Science*, 10, 1190164. https://doi.org/10.3389/fmars.2023.1190164

[52] Orlandi, A., Cappugi, A., Mari, R., Pasi, F., Ortolani, A., 2021. Meteorological Navigation by Integrating Metocean Forecast Data and Ship Performance Models into an ECDIS-like e-Navigation Prototype Interface. *Journal of Marine Science and Engineering*, 9(5), 502. https://doi.org/10.3390/jmse9050502

[53] 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. https://doi.org/10.1016/j.marpolbul.2025.117593

[54] HydroComp, 2023. NavCad - The premier hydrodynamic and propulsion system simulation software. https://www.hydrocompinc.com/solutions/navcad/ (accessed 7th July 2025)

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

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

[57] Wärtsilä, 2023. Navigation simulator NTPRO 5000. https://www.wartsila.com/marine/products/simulation-and-training/navigational-simulators/navigation-simulator-ntpro-5000 (accessed 7th July 2025)

[58] Fossen, T.I., 2011. *Handbook of Marine Craft Hydrodynamics and Motion Control*. John Wiley & Sons, Ltd., Hoboken, USA. https://doi.org/10.1002/9781119994138

[59] NDBC, 2025. Real-Time System Availability (RSA) - Frequently Asked Questions. NOAA National Data Buoy Center. https://www.ndbc.noaa.gov/faq/rsa.shtml (accessed 14th September 2025)

[60] NOAA, 2025. WAVEWATCH III Production Hindcasts and Forecasts. National Centers for Environmental Prediction, NOAA/National Weather Service. https://polar.ncep.noaa.gov/waves/wavewatch/ (accessed 14th September 2025)

[61] Salles, T., 2025. SAR Wave Mode Products - Coastal Dynamics and Evolution. University of Sydney, School of Geosciences. https://tristansalles.github.io/Coast/waveclimate/sar_waves.html (accessed 14th September 2025)

[62] NOAA-DA, 2024. 10-Year Strategy for Data Assimilation at NOAA. Environmental Prediction Innovation Center, NOAA. https://epic.noaa.gov/10-year-strategy-for-data-assimilation/ (accessed 14th September 2025)

[63] ECMWF, 2010. Jason-2 OGDR Wind and Wave Products: Random Error Estimation. European Centre for Medium-Range Weather Forecasts, Technical Memorandum. https://www.ecmwf.int/sites/default/files/elibrary/2010/7628-jason-2-ogdr-wind-and-wave-products-random-error-estimation.pdf (accessed 14th September 2025)

[64] Buizza, R., Leutbecher, M., 2015. The forecast skill horizon. *Quarterly Journal of the Royal Meteorological Society*, 141, 3366-3382. https://doi.org/10.1002/qj.2619

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

[66] Willmott, C.J., Robeson, S.M., Matsuura, K., 2011. A refined index of model performance. *International Journal of Climatology*, 32(13), 2088-2094. https://doi.org/10.1002/joc.2419

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

[68] Hersbach, H., 2000. Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting*, 15, 559-570. https://doi.org/10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2

[69] 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. https://doi.org/10.1111/j.1467-9868.2007.00587.x

[70] Rodwell, M.J., Wernli, H., 2023. Uncertainty growth and forecast reliability during extratropical cyclogenesis. *Weather and Climate Dynamics*, 4(3), 591-615. https://doi.org/10.5194/wcd-4-591-2023

[71] Kodaira, T., Sasmal, K., Miratsu, R., Fukui, T., Zhu, T., Waseda, T., 2023. Uncertainty in wave hindcasts in the North Atlantic Ocean. *Marine Structures*, 89, 103370. https://doi.org/10.1016/j.marstruc.2023.103370

Appendix 1. Pseudo-code of the integrated framework for attainable ship speed uncertainty quantification

```
INPUT: Weather forecasts (NOAA GFS), Ship parameters, Route waypoints
OUTPUT: Attainable ship speed uncertainties, ETA with confidence intervals
BEGIN
  // ========== INITIALIZATION ==========
  Grid ← 2619 points in North Atlantic [20°N-65°N, 70°W-10°W] at 50 nm spacing
  Lead_times ← {[0-24h], [24-72h], [72-120h], [120-168h]}
  Sea_states ← {[0-2.5m], [2.5-4m], [4-6m], [6-9m]}
  Encounter_angles ← {Head[0-30°], Bow[30-60°], Beam[60-120°], Stern[120-150°], Following[150-180°]}
  // ========== WEATHER FORECAST PROCESSING ==========
  FOR each lead_time IN Lead_times DO
    FOR each grid_point IN Grid DO
      // Extract meteorological variables
      Weather[grid_point, lead_time] ← {Hs, Tp, βwaves, Vwind, βwind}
      // Calculate encounter angle for 25 ship headings (0-360° at 15° intervals)
```

```
        αwaves
    END FOR
END FOR
    // ========== SHIP PERFORMANCE LOOKUP TABLES ==========
// Generate via simulations: 2028 scenarios for NTPRO, 1014 for NavCad
FOR Vref IN {12.0, 13.5, 14.5} knots DO
    FOR Hs IN {0, 1, 2, ..., 12} meters DO
        FOR αwaves IN {0°, 15°, 30°, ..., 180°} DO
            Vatt_JONSWAP[Vref, Hs, αwaves] ← NTPRO_5000(JONSWAP_spectrum)
            Vatt_PM[Vref, Hs, αwaves] ← NTPRO_5000(Pierson-Moskowitz_spectrum)
            Vatt_NavCad[Vref, Hs, αwaves] ← NavCad(quasi-static_method)
        END FOR
    END FOR
END FOR
    // ========== UNCERTAINTY QUANTIFICATION ==========
FOR each lead_time IN Lead_times DO
    FOR each sea_state IN Sea_states DO
        // Meteorological uncertainty metrics
        n ← number_of_forecast_observation_pairs
        // Calculate uncertainty metrics
        RMSE, MAE, Bias, IoA, CRPS, FSS, UGR, CMAE: (5)-(14)
        // Propagate to attainable ship speed
        Vatt_forecast ← LookupTable[Hs_forecast, αwaves_forecast, Vref]
        Vatt_actual ← LookupTable[Hs_actual, αwaves_actual, Vref]
        // Speed uncertainty metrics (same formulas applied to Vatt)
        Speed_Metrics[lead_time, sea_state] ← {RMSE, MAE, Bias, IoA, CRPS, FSS}
    END FOR
END FOR
    // ========== CORRELATION ANALYSIS ==========
FOR each metric IN {RMSE, MAE, Bias, IoA, CRPS, FSS} DO
    // Pearson correlation between weather and speed uncertainties
    rxy: (45), (46)
END FOR
    // ========== VOYAGE PLANNING WITH UNCERTAINTY ==========
current_position ← start_waypoint
current_time ← departure_time
cumulative_uncertainty ← 0
    WHILE distance_to_destination > 0 DO
    // Get weather at nearest grid point
    weather ← Weather[nearest_grid_point, current_time]
    lead_time ← current_time - departure_time
    // Lookup attainable speed and uncertainty
    Vatt_mean ← LookupTable[weather.Hs, weather.αwaves, Vref]
    ε ← Speed_Metrics[lead_time, sea_state_class]
    // Sample actual speed from distribution
    Vatt_actual: (23)
    // 95% confidence intervals
    CI: (25)
    // Update position and propagate uncertainty
    current_position ← current_position + Vatt_actual × Δt
    current_time ← current_time + Δt
    cumulative_uncertainty
```

```
    END WHILE
        // Final ETA with uncertainty bounds
    ETA ← current_time
    ETA_uncertainty ← cumulative_uncertainty
    ETA_CI ← [ETA ± 1.96×ETA_uncertainty]
        RETURN {Speed_Metrics, Correlations_rxy, ETA, ETA_CI}
END
```