Performance assessment of HurryWave in the North Sea using climate datasets

Msc Thesis report

Daniel van der Hoorn (5012929)

Delft University of Technology

TU Delft

Performance assessment of HurryWave in the North Sea using climate datasets

Msc Thesis report

by

Daniel van der Hoorn (5012929)

to obtain the degree of Master of Science
at the Delft University of Technology,

Student number:	5012929	
Project duration:	April, 2025 – October, 2025	
Thesis committee:	Dr. ing. J.A. (José A.) A. Antolínez,	TU Delft, supervisor
	Dr. ir. A.J.H.M. Reniers,	TU Delft, supervisor
	P.E. (Paulina) Kindermann,	TU Delft, supervisor

Cover:	Waves at Scheveningen beach by Daniel van der Hoorn
Style:	TU Delft Report Style, with modifications by Daan Zwaneveld

TU Delft

Preface

This master thesis is the final product of all my years of education. I am very proud to submit this work relevant to my chosen hydraulic coastal engineering track. I am very thankful to all the people who helped and guided me along the way.

I would like to express my gratitude to José Antolínez, the chair of my thesis committee, who suggested this thesis topic to me and for his guidance and valuable feedback throughout this project. I am also grateful to Ad Reniers, the second member of my committee, for providing datasets and for his insightful comments and for helping me strengthen the foundation of my work.

A special thanks goes to Paulina Kindermann, my daily supervisor, for her encouragement and continuous support during every stage of this thesis. Her advice and feedback have been invaluable to both my research and my personal development as an engineer.

My appreciation also extends to Koen van Asselt for his insights into the functioning and conception of HurryWave, and to Jacco Groeneweg for sharing his expertise in wave modeling, which helped me better understand the physical and computational aspects of my work. I would also like to thank Valvanuz Fernández for helping me learn the ropes of the Snellius high-performance computing system, her guidance greatly facilitated the work with complex simulations.

Finally, I want to thank my friends and family for their unwavering support and encouragement throughout this journey. They reminded me to keep perspective, and celebrate progress along the way.

Through this process, I learned as much about research as I did about myself, especially the importance of writing things on time, not overplanning, and trusting that progress often matters more than perfect preparation.

Abstract

The North Sea is a highly dynamic environment where accurate wave modeling is essential for coastal safety, offshore operations, and long-term climate adaptation. Recently, policy in the Netherlands has shifted toward reliance on SEAS5-based datasets to estimate storm loads, creating a demand for computationally efficient wave models capable of producing long climatologies with sufficient accuracy. HurryWave, a new third-generation spectral wave model developed by Deltares, promises an order-of-magnitude faster performance than SWAN due to simplified physics and numerics. This thesis investigates whether such efficiency compromises wave modeling accuracy in the North Sea.

HurryWave was evaluated against buoy measurements and ERA5 reanalysis data from 1950–2023. Model performance was assessed for significant wave height (H_{m0}), peak period (T_p), and mean wave direction (θ_m), with additional analyses on storm events, and spectral representation. Results show that HurryWave reproduces significant wave height with skill comparable to ERA5, achieving root mean square error (RMSE) values of 0.35-0.55 m and scatter index values of 0.12-0.20 across stations in shallower water near the Netherlands. However, in deeper areas such as at North Cormorant, HurryWave without boundary conditions underestimates wave height by a large amount with a scatter index of 0.36. This is mitigated by including a larger domain or by adding boundary conditions. Additionally, using this model setup, large wave heights are underestimated, even at stations in shallower water. This is mainly due to the general underestimation of high wind speeds in the ERA5 input compared to KNMI observations. Peak period (T_p) was consistently underestimated using this setup, with RMSE values of 1.5-2.5 s and negative biases of -0.5 to -1.5 s. Wave direction was the weakest-performing parameter, with RMSE frequently exceeding 30° and large directional scatter during storms. This is due in part to both, the absence of triad modeling which would have an effect on energy distribution to lower frequencies and in transferring energy to other directions, and the inaccuracies in the ERA5 wind data. Additionally the DIA approximation used in modeling quadruplets can lead to errors in fast turning wave fields, contributing to the error in direction. During the 2013 "Sinterklaas storm", HurryWave captured overall storm evolution but underestimated offshore peak H_{m0} by up to 1.0 m and showed narrower spectral energy distributions compared to buoy observations. The analysis of other spectral wave parameters revealed that HurryWave performed better when modeling mean period than the peak period with it performing best in modeling the mean period of the negative first order spectral moment T_{m-10} with an RMSE of up to 0.24s compared to the 1.5 to 2.5s RMSE observed for T_p during the Sinterklaas storm.

The findings highlight HurryWave's strength in reproducing significant wave height, making it suitable for applications such as long-term SEAS5-based climatologies. However, limitations in extreme wave heights, peak period, wave direction, and spectral realism caution against its sole use in operational forecasting or safety assessments that depend on extreme storm accuracy. This study concludes that HurryWave is a promising complement, not replacement, to established models, aligning well with the SEAS5-driven shift toward risk-based, long-horizon coastal policy, while still requiring refinement for storm-scale hazard analysis.

Contents

List of Figures

List of Tables

1 Introduction

1.1. Research context

The North Sea is a marginal sea characterized by complex geology, strong wind patterns, and significant wave activity (Ziegler, 1975). It has served as a vital resource and hub of human activity for centuries, playing a crucial role in trade, energy production, fishing, recreation, and numerous other endeavors that have shaped the economies and cultures of surrounding nations. This expansive body of water has not only facilitated maritime commerce and resource extraction, but has also been a theater for historical events, technological advancements, and environmental challenges (Jesper H. Andersen et al., 2013).

In modern times, the North Sea and the countries surrounding it face the challenge of climate change, which leads to rising sea levels and rising water temperatures (Wageningen University & Research, 2021). Accurate prediction of changes in these elements through climate models is vital to both the ecology, and human use of the system, whether for energy extraction or coastal flood defense of the surrounding countries (Wageningen University & Research, 2021). In the shorter term, wave modeling is necessary to accurately understand and predict dynamics of both the sea and the surrounding coast (Barnard et al., 2019). This ensures navigational safety and flood risk assessment in the short term. Accurate wave modeling is also used to predict erosion of the coast and plan out sand nourishments accordingly (Curto, 2014).

The general policy regarding extreme storms on the North Sea coast is shifting towards reliance on the SEAS5 forecast dataset as the basis for estimating loads on hydraulic structures in the Netherlands (de Valk & van den Brink, 2023). This transition marks a departure from historical reliance on locally measured or deterministic reanalysis data, aiming instead to derive all necessary parameters from a globally validated dataset. This dataset incorporates synthetic time series from SEAS5 and hydrodynamic simulations, providing the equivalent of around 8000 years of data (de Valk & van den Brink, 2023). The motivation behind this policy shift is the significant improvement in forecast skill, model resolution, and data assimilation techniques available in SEAS5, which enhance long-range predictability of storm events (Johnson et al., 2019). This policy change also stems from the increasing need for consistency in managing coastal risks amid a changing climate. This consistency was a weakness of the previous WBI-2017 framework, where wind and water level extremes were treated separately, leading to potential mismatches in joint probability analysis (de Valk & van den Brink, 2023). Unlike WBI-2017, which relied heavily on relatively short observational series and deterministic reanalysis products, BOI-2023 incorporates long synthetic time series from SEAS5 to derive consistent statistics of wind, surge, and waves. This reduces uncertainty in rare-event estimates and improves coherence between variables that jointly determine coastal flood risk. Early results show that return values for wind and water levels are generally lower and more precise than previously in WBI-2017, which can lower reinforcement costs while still meeting safety standards (de Valk & van den Brink, 2023). An example of this is shown in the assessment of de Valk and van den Brink (2023) which shows that for a same return period of 1/10,000 years, the water level at Hoek van Holland in the Netherlands goes from 5.09m in the WBI-2017 policy to 4.61m in the newer BOI-2023 policy. Uniform adoption of the newer policy therefore helps Dutch authorities calibrate protection measures against the latest climate trends rather than historical averages, accounting for nonlinear effects in both waves and winds and offering better risk mitigation over decades.

Within this framework, improved wave modeling is critical. Flood risks in the Netherlands are often dominated by wave overtopping and dune erosion rather than water level alone. Fast and computationally efficient wave models, calibrated with SEAS5 outputs and validated against regional

measurements, allow better representation of nearshore processes that directly load dikes and dunes. By reducing uncertainty in wave height, direction, and period under extreme storms, such models support more reliable safety assessment, timely decision making and cost-effective infrastructure investment. Better wave modeling underpins the broader BOI-2023 shift from deterministic analyses toward integrated, risk-based flood safety policy aligned with long-term climate adaptation (de Valk & van den Brink, 2023).

1.2. Research problem

A new third-generation spectral wave model developed by Deltares named HurryWave is currently being finalized and is set to be officially released in the coming year. This model is on average "an order of magnitude faster than the SWAN model due to its reduced physics, combined with simpler wave propagation schemes" (Destination Earth, n.d.). While this computational efficiency makes it a promising tool for large-scale wave modeling applications, the extent to which its reduced physics and different numerical scheme affects model accuracy has not been extensively studied.

HurryWave has been validated against SWAN, another 3rd generation wave model in some specific use cases (van der Lugt et al., 2024), however, fully comprehensive studies comparing HurryWave to both measurements gathered on the North Sea, and larger climate datasets are yet to be conducted. The accuracy of the model in simulating key wave parameters, including significant wave height, wave period, and wave direction, needs to be evaluated to determine whether the trade-off between speed and physical complexity impacts its predictive skill. Additionally, HurryWave's performance under extreme wave conditions, such as storms in the North Sea, has not been assessed separately from long-term climate predictions. Given the energetic and dynamic nature of the North Sea, where storms and strong wind-wave interactions are common (Weisse et al., 2012), it is crucial to evaluate whether HurryWave can reliably reproduce wave behavior in both daily and extreme conditions. The study of each unimodal and multimodal wave family is also of interest to understand where the model performs well and where it shows limitations. Moreover, the analysis of the spectral shape of the waves produced by HurryWave is also of importance to see if certain wave families are accurately modeled.

1.3. Research aim and objectives

This study aims to assess the accuracy of HurryWave by comparing its outputs to historical measurements and alternative wave models such as WAM used in the ERA5 dataset (Hersbach et al., 2020). By validating the model against observational sources and refining its configuration, this research will provide critical insights into whether HurryWave can serve as a viable alternative to existing wave models in the North Sea. This main goal is divided into different objectives. These objectives are as follows.

- Assess the differences between HurryWave and other wave models in the simulation of wind-wave growth, transformation, and dissipation through a literature review of the different manuals.
- Compare HurryWave's performance against historical data in terms of accuracy with default settings.
- Assess HurryWave's performance in modeling spectra through spectral analysis.

1.4. Research scope

The scope of this research is limited to the geographical area of the North Sea as shown in figure 1.1 taken from the validation report (van der Lugt et al., 2024). This area goes from -12 degrees to 10 degrees in longitude, and 48 and 65 degrees in latitude. This research focuses exclusively on the North Sea. The analysis is conducted on deep to intermediate water wave conditions and with stationary sea levels. Long waves, or waves above a 25s period (Rijnsdorp et al., 2021) are excluded from this study. No boundary conditions are applied to the HurryWave model.

Additionally, the HurryWave model output is exclusively compared with the outputs of wave buoys and wave data from the ERA5 dataset. This comparison is made qualitatively or quantitatively using the significant wave height (H_{m0}), the peak period (T_p) and the mean wave direction (θ_m).

The significant wave height is defined as four times the square root of the zeroth spectral moment. The

zeroth spectral moment m_0 and the significant wave height H_{m0} are defined as:

$$m_0 = \int_0^\infty S(f)\, df \tag{1.1}$$

- $S(f)$: wave energy spectrum [m^2/Hz]
- f: wave frequency [Hz]
- m_0: zeroth spectral moment [m^2]

$$H_{m0} = 4\sqrt{m_0} \tag{1.2}$$

- H_{m0}: significant wave height [m]

The peak period corresponds to the period associated with the spectral peak:

$$T_p = \frac{1}{f_p} \tag{1.3}$$

- f_p: frequency at which $S(f)$ reaches its maximum [Hz]
- T_p: peak period [s]

The mean wave direction is obtained from the directional spectrum $S(f,\theta)$ and is computed following the formulation given in by the European Centre for Medium-Range Weather Forecasts (ECMWF) in Bidlot (2020):

$$\theta_m = \arctan\left(\frac{\int_0^{2\pi}\int_0^\infty S(f,\theta)\sin\theta\, df\, d\theta}{\int_0^{2\pi}\int_0^\infty S(f,\theta)\cos\theta\, df\, d\theta}\right) \tag{1.4}$$

- $S(f,\theta)$: directional wave spectrum [m^2/Hz/rad]
- θ: wave direction [rad]
- θ_m: mean wave direction [rad]

The mean wave direction is then converted to degrees. These 3 parameters are then used to compare the output of HurryWave to ERA5 and buoy measurements.

The HurryWave model is set up using the wind data from the extended ERA5 dataset, along with the bathymetry used in the previous validation (van der Lugt et al., 2024). The model is run backwards from 2023 until 1950, as that is the first instance where wind data is obtained from the ERA5 dataset (Bell et al., 2021). Spectral energy output is exclusively compared to spectral buoy data at locations where these records are available.

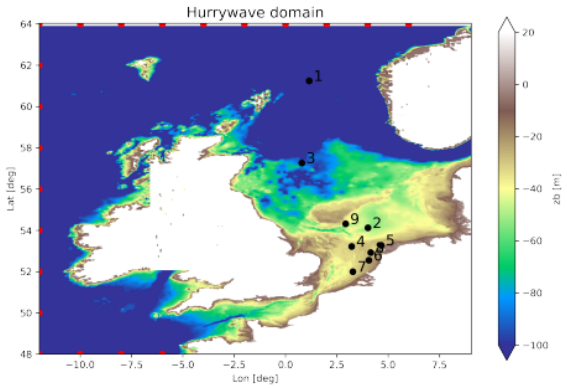

Figure 1.1: North Sea domain in HurryWave considered in this study, taken from van der Lugt et al. (2024), the bathymetry used is visible in the color scale, the stations used are numbered and the red points on the edge denote points where boundary conditions were applied in the report by van der Lugt et al. (2024), these were not applied in this model setup.

1.5. Research questions

The research question guiding this project is therefore:

> **How does HurryWave compare to historical modeled and measured data, in terms of accuracy in modeling wave parameters in the North Sea?**

The sub-questions guiding this project are as follows:

1. What are the key differences in methodology and assumptions between the HurryWave model and other wave models?
2. How does HurryWave perform against measured data and the ERA5 dataset in modeling wave height, period and direction in the North Sea?
3. How does HurryWave perform against measured data in modeling the wave spectra in the North Sea?
4. What are the advantages and limitations to the use of the HurryWave model in the North Sea?

1.6. Societal relevance

1.6.1. Stakeholder analysis

The main stakeholders directly involved in this research are TU Delft providing the research opportunity and Deltares providing the research team with the HurryWave model and expertise. However, this project may involve other stakeholders through either exchange of expertise, interest in outcomes or consequences of policies adopted using the model. These include the Royal Netherlands Meteorological Institute (KNMI), the Dutch Ministry of Infrastructure and Water Management (Rijkswaterstaat), other potential (international) model users, and inhabitants affected by policy choices made with the model. An overview of these stakeholders is shown in a power-interest grid in figure 1.2.

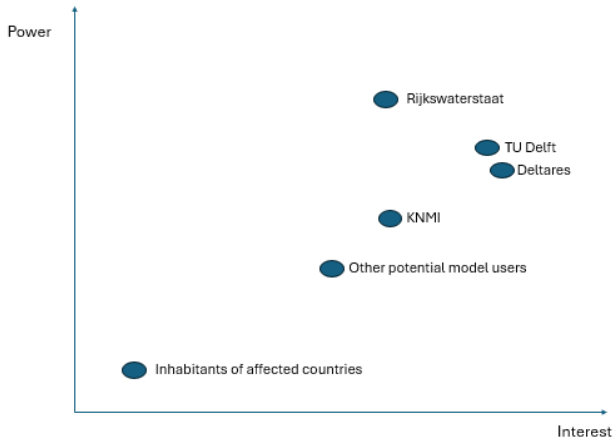

Figure 1.2: Power interest grid of the most relevant stakeholders involved in this project.

TU Delft is the university providing the opportunity to conduct this research. They are invested in progressing scientific knowledge through research of this new model. Their core values are diversity, integrity, respect, engagement, courage, and trust (Roseser & Copeland, 2020).

Deltares are the developers of the HurryWave model (van Asselt, 2025). They are invested in further validating and improving the model after an initial validation report (van der Lugt et al., 2024). Their core values are objectivity, independence, openness and transparency (Deltares, n.d.).

The Royal Netherlands Meteorological Institute (KNMI) is the national weather service in the Netherlands. Their tasks involve weather forecast and monitoring. They are also the "national research and information centre for meteorology, climate, air quality, and seismology" (“KNMI - About KNMI”, n.d.). Their interest in this project would be to more reliably and accurately predict weather and climate scenarios

using HurryWave. Their core values are "safety, accessibility, sustainability and prosperity of the Netherlands" (“KNMI - About KNMI”, n.d.).

The Dutch Ministry of Infrastructure and Water Management (Rijkswaterstaat) is "responsible for the design, construction, management and maintenance of the Netherlands’ primary infrastructure facilities" (Ministerie van Infrastructuur en Waterstaat, n.d.). Their tasks involve among others, data collection (“Wave measuring buoy | Digital North Sea”, n.d.), coastal management, and policy development. These policies rely on accurate data from wave models such as HurryWave. Their main interest in this project, like the KNMI, is accuracy and reliability of the model output. Their core values are social responsibility, integrity, trustworthiness, and transparency (Pegel, 2014).

Other potential users of the model, such as private companies, or foreign public infrastructure or weather institutions could show an interest in this project. They could also be looking at output accuracy and how the use of HurryWave could benefit their weather and climate forecasting.

Lastly, if national infrastructure or weather institutions were to use HurryWave through the results of this study, the inhabitants of the countries around the North Sea will be affected by this project. Their interests would lie around the accuracy and transparency of the model.

1.6.2. Societal impact

The first order impact of this research is the improvement of wave modeling accuracy in the North Sea. If HurryWave is indeed as reliable as existing models but computationally more efficient, it could enhance the speed at which large-scale climate datasets are assessed. This would benefit infrastructure or weather institutions such as Rijkswaterstaat and KNMI by enabling them to run more frequent simulations, leading to more informed decision-making.

A second-order impact is the potential adoption of HurryWave in operational forecasting and long-term climate projections. Within a few years, if the model is widely accepted, it could supplement models like SWAN and WAM, influencing how climate risks are assessed. This could have economic implications, such as reducing the computational costs of large-scale climate simulations, allowing more frequent updates to wave forecasts, and improving the precision of climate risk assessments.

However, there are also potential negative consequences. Over-reliance on a new model before its limitations are fully understood could lead to errors in forecasting. If the model underestimates extreme wave events, for instance, it could contribute to inadequate coastal defenses or misinformed emergency response measures.

1.6.3. Risks regarding the misuse of research outputs

While this research aims to improve wave modeling, its outputs could be misinterpreted or misapplied in ways that create adverse societal consequences. The three most relevant risks include:

Model overconfidence

One of the main risks is that stakeholders assume HurryWave’s computational efficiency comes without trade-offs, potentially leading to policy decisions based on incomplete model validation. If the model is not adequately tested under extreme wave conditions but is still used for high-stakes decision-making, it could lead to unexpected failures in flood risk assessments. To mitigate this, the study should clearly communicate the model’s strengths and limitations, ensuring it is only used in contexts where it has been proven reliable.

Data misinterpretation

If wave prediction data from HurryWave is released without sufficient documentation, decision-makers unfamiliar with wave modeling could misinterpret its outputs, especially under uncertain conditions. This could result in flawed predictions of extreme weather impacts, which could lead to poorly informed flood policy decisions. The mitigation strategy here is to provide clear guidelines on data interpretation and to ensure results are always presented alongside uncertainty estimates and comparisons with established models or measurements.

Neglect of Alternative Modeling Approaches

A shift toward HurryWave could cause institutions to reduce investments in other, more detailed wave modeling efforts. While a more efficient model is beneficial, a single-model approach increases risk as no model is perfect in every situation. This risk could extend to the population if this model under predicts the effects of storms or other extreme climate events. The research should highlight that HurryWave complements rather than replaces existing models, maintaining a multi-model approach for robustness.

1.7. Thesis structure

This thesis is structured into seven chapters.

- Chapter 1 introduces the research context, the problem addressed, the objectives, and the research questions.
- Chapter 2 provides background information on wave modeling in the North Sea, the HurryWave model, and the datasets used.
- Chapter 3 details the methodology, including data acquisition, model setup, and statistical evaluation metrics.
- Chapter 4 presents the results of the reanalysis and sensitivity tests, assessing the model's performance in different conditions.
- Chapter 5 offers a discussion of the key findings, the study's limitations, and suggestions for future work.
- Chapter 6 concludes with a summary of findings and implications.

2
Background

2.1. Wave data collection in the North Sea

Wave data in the North Sea can be collected in various ways. The most common are the use of wave buoys which are explained below.

Wave buoy measurements

Wave buoys are the primary source of data when it comes to waves in the North Sea. The most measured variables are significant wave height (H_{m0}) and directionality. This means wave buoys capture the height and direction of the wave field (Lavidas & Alday, 2022). Wave buoys in the North Sea have a high temporal resolution providing measurements every 10 to 30 minutes (Lavidas & Alday, 2022). These measurements are spread out in space in various locations, covering both deep and shallow areas in the North Sea (Lavidas & Polinder, 2019). The recording of these measurements started in 1980, though data availability ranges between 64% and 94%, depending on the buoy (Lavidas & Alday, 2022). The measurements are available through the Dutch Ministry of Infrastructure and Water Management Rijkswaterstaat ("Wave measuring buoy | Digital North Sea", n.d.).

Spectral buoy measurements

Some buoys are designed to measure wave energy and frequency in the North Sea. Spectral wave buoys record the vertical motion of the sea surface using high-resolution accelerometers and motion sensors. These produce time series data from which the wave energy spectrum is extracted through spectral analysis. These buoys, such as the Datawell Directional Waverider output spectral energy density as a function of frequency and direction, allowing the quantification of wave energy across different sea states. These buoys can also output integrated parameters such as significant wave height, mean wave period, and directional spreading (Lavidas & Polinder, 2019).

2.2. HurryWave model description and validation

The description of the HurryWave model and its comparison to other models is based on the documentation described in van Asselt (2025). Figure 2.1 presents the conceptual framework of HurryWave, which is elaborated upon in the following sections.

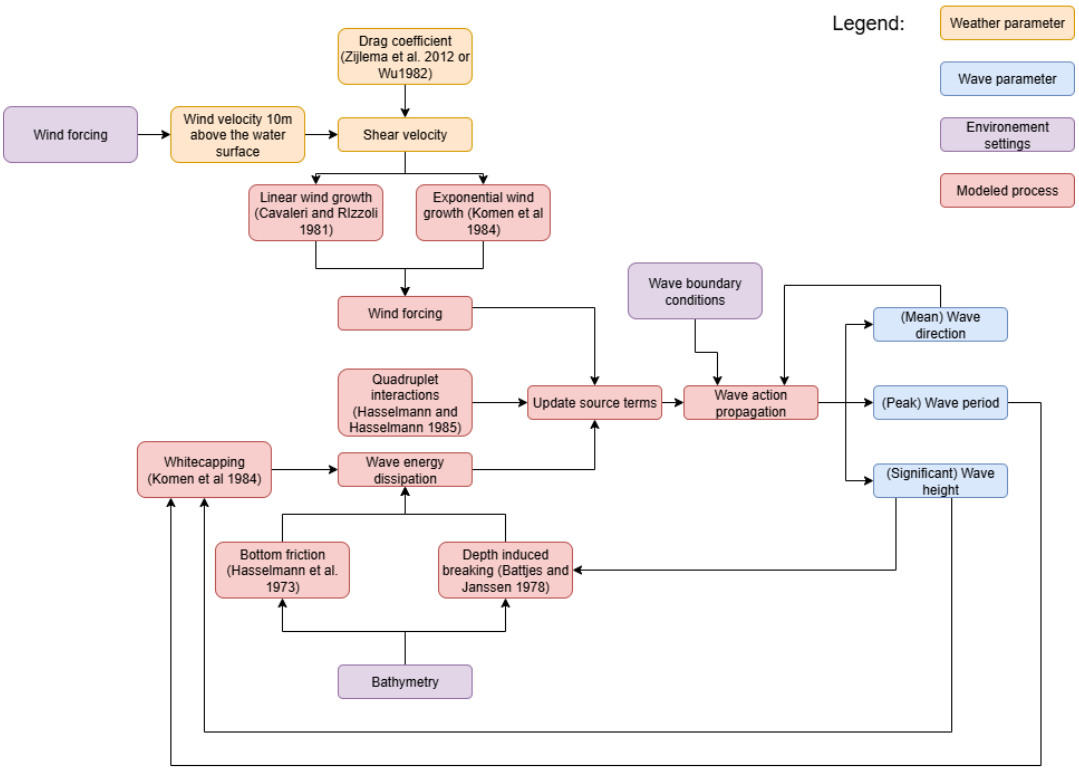

Figure 2.1: Conceptual framework of the HurryWave model based on van Asselt (2025).

HurryWave is a phase-averaged spectral wave model designed for large-scale, non-stationary wave modeling applications, particularly those influenced by storms and cyclones in deep water. The model employs an explicit first-order upwind numerical scheme, prioritizing computational efficiency over some aspects of physical complexity. Unlike models such as SWAN and Wave Watch III (WW3), which incorporate implicit solvers and detailed nearshore processes, HurryWave is optimized for rapid simulation of wave propagation in deep-water environments where wave-current interactions and shallow-water triad interactions are not explicitly resolved. The model is built upon the classical wave action balance equation shown in equation 2.1.

$$\frac{\partial N(\sigma,\theta)}{\partial t} + \nabla \cdot [c_g N(\sigma,\theta)] + \frac{\partial c_g N(\sigma,\theta)}{\partial \sigma} + \frac{\partial c_\theta N(\sigma,\theta)}{\partial \theta} = \frac{S_{\text{wind}}}{\sigma} + \frac{S_{\text{diss}}}{\sigma} + \frac{S_{\text{nl4}}}{\sigma} \tag{2.1}$$

On the left-hand side of equation 2.1, wave action propagation follows $N = E/\sigma$ in both directional (angular domain) and geographical space (physical spatial domain) , with E the wave energy, θ the direction and σ the radial frequency. In equation 2.1, c_g is the wave group celerity found for each frequency and c_θ is the celerity of refraction.

On the other side of equation 2.1, the source and sink terms can be found where S_{wind} represents the wind input, S_{diss} represents the dissipation terms and S_{nl4} represents the quadruplet interaction terms. These terms are expanded upon in the following paragraphs.

Wind input

Wind input is the main wave generation mechanism modeled by HurryWave. Wind in itself has a speed proportional to the air pressure gradient and direction related to the air pressure isobars. The wave heights and periods generated by a wind field depend on the wind velocity, and the distance of water over which the wind blows also known as fetch (Bosboom & Stive, 2023).

The wind forcing in HurryWave is given by $S_{wind} = A + BE(\sigma,\theta)$ where A and B are based on the shear velocity given by $U_*^2 = C_d * u_{10}^2$ with U_* the shear velocity, C_d the drag coefficient and u_{10} the wind

velocity 10m above the water surface. Linear wind growth A uses the formulas described by Cavaleri and Rizzoli (1981), whereas exponential wind growth B uses the formulas described by Komen et al. (1984). The drag coefficient can be specified using two methods, the formulation by Wu (1982), or the formulation by Zijlema et al. (2012). The formulation by Wu (1982) is an empirical formulation used historically. This formulation tends to produce stronger wind forcing and therefore overestimating wave heights in extreme events (Wróbel-Niedźwiecka et al., 2019). The formulation by Zijlema et al. (2012) corrects this by capping the wind speed when it is large. This produces a better wave height prediction in storms and reduces the need for tuning the dissipation to compensate the source term (Kim et al., 2015).

Dissipation

Once the wave field has been generated, the dissipation starts to balance out the forcing, leading to more efficient energy transfer, where wind forcing on larger waves balances wave breaking also called white capping (Bosboom & Stive, 2023).

The modeled dissipation is composed of three terms, dissipation caused by whitecapping, dissipation caused by depth induced breaking in shallow water, and dissipation caused by bed friction. The whitecapping formulation follows the formulation of Komen et al. (1984) using tunable coefficients which control dissipation based on wave steepness and spectral parameters. The dissipation due to depth induced breaking follows the formulations made by Battjes and Janssen (1978). Bed friction follows the JONSWAP formulation of K. Hasselmann et al. (1973).

Quadruplets

The Discrete Interaction Approximation (DIA) approach by S. Hasselmann and Hasselmann (1985) is used to model four-wave interactions. The DIA is computationally cheap compared with the full Boltzmann integral but it is known to misrepresent the exact transfer rates and spectral redistribution. A consequence of this misrepresentation is the increase of nonlinear transfers in shallow water. This is due to the method being developed for deeper water, and not accounting for changes in the spectral shape when moving into shallow water as shown in van Vledder (2001).

Other assumptions

Other assumptions of this model include constant and uniform water-levels. This is because the model is meant to be applied in deep water conditions. The effect of this assumption occurs near the shore where water level variations affect shoaling, refraction, and wave breaking. Additionally, currents are also not modeled in HurryWave which could also lead to a misrepresentation of wave height, period, and direction, specifically in areas with large currents such as tidal inlets. Currents going against the wave direction tend to increase wave height and reduce period, whereas the current going in the same direction as waves tends to reduce wave height and increase wave period. A specific model concerning infragravity waves is not yet included in HurryWave which could also affect nearshore model validity. Lastly, the omission of three wave interactions or triads also reduces accuracy at the nearshore. Triad interactions transfer energy from the spectral peak frequency to infragravity or higher harmonics in shallower waters (Eldeberky, 1996). This means model output could be inaccurate in the nearshore or in locations with a lower depth.

Validation

In terms of performance, HurryWave has been validated against analytical solutions, stationary test cases, and some real-world observations in the validation report by van der Lugt et al. (2024).

This report shows that the model successfully reproduces fundamental wave phenomena such as shoaling, refraction, and depth-induced breaking, aligning well with analytical solutions and existing full-physics models in many scenarios. Comparative studies using the SWAN model indicate that HurryWave achieves similar skill in predicting significant wave height and peak wave period for medium to coarse grid resolutions (>1 km). However, in highly detailed nearshore applications requiring fine-grid resolution (<250 m), its explicit numerical scheme results in slower convergence and reduced accuracy compared to SWAN.

HurryWave has also been tested across a range of environments, including the Dutch continental shelf and extreme weather events such as hurricanes Fiona and Idalia. These studies demonstrate its

efficiency in modeling large-scale wave fields while maintaining accuracy. In hurricane simulations, HurryWave successfully captured the temporal and spatial evolution of wave parameters, though some discrepancies in peak wave height predictions were noted, likely due to unresolved bathymetric and wind field variations. For lake and estuarine environments, such as Lake George and the Haringvliet, the model showed strong performance in depth-limited wave growth conditions, with errors comparable to SWAN when boundary conditions were carefully prescribed. This also shows that model settings and boundary conditions have a large effect on the accuracy of the results, model tuning is therefore essential.

In terms of boundaries, the model supports both parametric and full spectral boundary condition inputs, with little to no transitions observed when using nested grids in large-domain applications. Sensitivity tests have shown that grid rotation, wind drag formulations, and spectral resolution impact the results. The increased directional and frequency resolution improves agreement with analytical solutions at the cost of computational efficiency. Notably, in estuarine and tidal inlet scenarios, the absence of wave-current interactions modeling can lead to localized errors, particularly in peak period estimates.

2.3. ERA5 dataset

The ERA5 dataset is defined as a "comprehensive reanalysis, from 1940 to 5 days behind real time, which assimilates as many observations as possible in the upper air and near surface" by the European Centre for Medium-Range Weather Forecasts (ECMWF) (ECMWF, n.d.-b). The following description of the dataset and its method of generating waves is based on the documentation found in ECMWF (n.d.-a), but also on the reports provided by ECMWF which detail the process of reanalysis first from 1979 onward (Hersbach et al., 2020) and then extended to start in 1950 (Bell et al., 2021).

The ERA5 reanalysis combines advanced 4D-Var data assimilation techniques with the Integrated Forecast System (IFS CY41R2) to generate hourly estimates of various atmospheric, land, and ocean-wave parameters. ERA5 consists of a high-resolution deterministic run (HRES) with a 31 km grid spacing and a 10-member ensemble data assimilations (EDA) at 63 km resolution, providing both detailed predictions and uncertainty quantification.

Within ERA5, 10m wind speeds and directions are crucial components derived from the atmospheric model coupled to the land-surface model (HTESSEL). These surface winds are generated through the assimilation of observational data, including satellite scatterometer measurements and buoy readings, using Optimal Interpolation and an Extended Kalman Filter. This process ensures consistency between surface winds and upper-air dynamics. The resulting wind data is archived hourly as "surface or single level" parameters, with a resolution of 0.28125° (approximately 31 km) for atmospheric fields and 0.36° for wave-related calculations.

Wave modeling in ERA5 is accomplished using the WAM model, which is dynamically forced by the 10m winds from the atmospheric component. The wave model generates various parameters, including significant wave height, mean wave direction, and peak period. It operates on a reduced latitude-longitude grid with a resolution of 0.36° for the high-resolution run and 1.0° for the ensemble members. This setup allows ERA5 to capture storm-driven extremes in wave conditions.

Uncertainty quantification is a key feature of ERA5, primarily achieved through its 10-member ensemble data assimilation (EDA). The EDA provides probabilistic uncertainty estimates for both winds and waves, with the lower-resolution ensemble members (63 km for atmosphere, 1.0° for waves) being particularly useful for assessing variability in data-sparse regions. It is important to note the distinction between near-real-time data (ERA5T) and the final ERA5 product, with the former being overwritten by quality-controlled data approximately two months later. This process can lead to significant corrections, as evidenced by recent snow assimilation errors in 2021 and 2024 shown in the documentation (ECMWF, n.d.-a), underscoring the importance of using finalized datasets for applications.

Despite its comprehensive nature, ERA5 does have some limitations. Wave heights may be largely underestimated with the dataset underestimating wave power by up to 42 % in some cases as shown in Afolabi et al. (2024). Additionally, the 0.36° resolution of the wave model means that coastal bathymetry and small-scale wind features are not fully resolved, which can impact the accuracy of nearshore predictions.

2.4. Other wave models

HurryWave is also compared fundamentally to other wave models such as WAM used in ERA5, SWAN and WaveWatch III (WW3). All of these models are spectral wave models used on different scales. A comparison of the fundamental approach of each model is shown in table 2.1 based on the manuals of The SWAN team (2006), ECMWF (n.d.-a) and The WAVEWATCH III R Development Group (WW3DG) (2016).

Table 2.1: Fundamental processes and typical setup values for SWAN, WAM (ERA5), and WW3.

Feature	**SWAN**	**WAM (ERA5)**	**WW3**
Equation	Spectra action / Energy balance	Spectra action / Energy balance (deep-water focus)	Spectra action / Energy balance (coastal or deep)
Timestep	-	60 min	-
Grid / space step	-	0.25° (global)	-
Numerics	Backward Space Backward Time implicit	Implicit upwind	Explicit/implicit
Wind input	Janssen	Janssen	Janssen/Ardhuin
Dissipation	Komen	WAM Cycle 4 (Janssen)	Ardhuin
Triads	Yes (in shallow water)	No	No
Quadruplets	DIA (discrete interaction approximation)	DIA (discrete interaction approximation)	DIA / other user-defined alternatives
Bottom friction	Yes	No	Yes (optional)

2.5. Wave hindcasts using different models

Model validation often comes in the form of a hindcast, where the model simulates the past conditions to see if the data collected at the same period lines up with the model. A hindcast using HurryWave has not yet been performed, however, hindcasts in other models can give valuable insight into how wave models function.

Most North Sea wave hindcast studies adopt regional, nested wave models forced with atmospheric reanalysis data. For example, Reistad et al. (2011) utilized a nested WAM model with a coarse 50 km grid forced by ERA-40 winds, incorporating a finer 10-11 km mesh for improved local accuracy. Groll and Weisse (2017). implemented coastDat2 with WAM featuring enhanced source term parametrization and a 0.05° grid (6km), forced by dynamically downscaled NCEP/NCAR wind fields. The reanalysis on December 2013 done by Ponce de Leon et al. (2018) focuses on a comparison between SWAN and Wavewatch III in the North Sea on the basis of wave height. Some other studies looked into the validation of SWAN in different cases such as a case near the Dutch coast (Ris et al., 1999), and (Scholten, 2024). These cases correspond more to validation cases but are still useful to get an idea as to how SWAN performs, even in short periods.

Each of these studies compare integrated parameters over the wave spectrum such as significant wave height as shown in equation 1.2, mean period which is an energy weighted mean of all periods as shown in equation 2.2, and wave direction as shown in equation 1.4. These studies compare these parameters through several statistical methods such root mean square error (RMSE), scatter index (SI), statistical bias and correlation coefficient.

$$T_{\text{mean}} = \frac{\int_0^\infty TS(f)\,df}{\int_0^\infty S(f)\,df}, \tag{2.2}$$

where $T = 1/f$ is the wave period in seconds corresponding to frequency f in Hertz and $S(f)$ is the directional wave spectrum [m^2/Hz/rad].

The output of these simulations for significant wave height is shown in table 2.2, the output of these simulations for mean period is shown in table 2.3.

Table 2.2: Performance statistics of different models in the North Sea comparing significant wave height (H_s)

Model / Dataset	Bias [m]	Scatter Index	RMSE [m]
ERA-Interim (Groll & Weisse, 2017)	−0.14 to 0	0.1 to 0.3	0.2 to 0.5
WAM (Reistad et al., 2011)	0.1 to 0.25	–	0.5 to 0.7
SWAN (Scholten, 2024) and (Ris et al., 1999)	0.1 to 0.5	0.18 to 0.77	0.2 to 0.5
WW3 (Ponce de Leon et al., 2018)	−0.21 to 0.17	−0.09 to 0.16	-

ERA-interim tends to underestimate the wave height whereas it is overestimated for several other models such as WAM and SWAN.

Table 2.3: Performance statistics of different models in the North Sea comparing mean wave period

Model / Dataset	Bias [s]	Scatter Index	RMSE [s]
ERA-Interim (Groll & Weisse, 2017)	−0.7 to − 0.6	0.2 to 0.1	0.2 to 0.9
WAM (Reistad et al., 2011)	–	–	–
SWAN (Scholten, 2024) and (Ris et al., 1999)	−0.8 to − 0.2	0.11 to 0.32	0.5 to 1

The mean period shows some differences depending on models. General under and over estimations are not found as certain models overestimate such as coastdat2 (Groll & Weisse, 2017), whereas others such as WAM tend to vary based on location (Elshinnawy & Antolínez, 2023) (Reistad et al., 2011). Generally, according to Reistad et al. (2011), the mean period is underestimated, especially in higher periods. SWAN also tends to systematically under predict mean periods (Scholten, 2024) up to 0.8s.

The output of these simulations for mean wave direction is shown in table 2.4.

Table 2.4: Performance statistics of different models in the North Sea comparing mean wave direction

Model / Dataset	Bias [°]	Scatter Index	RMSE [°]
ERA-Interim (Groll & Weisse, 2017)	22 to 26	0.15 to 0.17	35 to 38
WAM (Reistad et al., 2011)	–	–	–
SWAN (Scholten, 2024) and (van Nieuwkoop & Gautier, 2015)	10 to 20	–	5 to 20

Models and datasets tend to poorly represent directions with standard errors having very high values. This can be due to the metrics as these values are very sensitive to fast changing conditions.

These comparisons were made following the papers described above. The models vary in measuring stations and simulation durations, with some SWAN runs lasting only a week. These serve as benchmark and comparisons for future results obtained with HurryWave.

2.6. Research gap

Despite the significant advancements in wave modeling, several key gaps remain in the accurate prediction and modeling of wave dynamics in the North Sea. Existing third-generation spectral wave models such as SWAN, and WAM have been validated against historical wave buoy data and satellite observations (Lavidas & Alday, 2022), (Hersbach et al., 2020). However, the newly developed HurryWave model has not yet been systematically tested against these established models in the context of large-scale climate datasets, particularly for the North Sea.

A critical limitation in wave modeling is the computational cost associated with high-resolution simulations. While HurryWave is designed to be significantly faster than SWAN due to its reduced physics and simplified wave propagation schemes (Destination Earth, n.d.), the impact of these simplifications on wave model accuracy remains unclear. Understanding how the simplifications affect wind-wave growth, wave transformation, and dissipation is important for determining the accuracy of HurryWave in operational and climate forecasting applications. Separating the wave trains into different categories and analyzing their spectra to understand the impacts of different physical phenomena, especially in multimodal settings, is yet to be done with results produced by the HurryWave model.

Given these challenges, this research aims to fill the gap by comparing HurryWave against observations from both wave buoys, and datasets derived from established models and reanalysis products such as ERA5. By evaluating its performance in modeling wave characteristics in the North Sea, this study will provide insights into the accuracy of the new HurryWave model, ultimately determining how this can be accurately applied in climate and operational forecasting contexts. Ultimately, the HurryWave model is intended for forecasting applications (van Asselt, 2025). This research would therefore contribute to assessing the accuracy of the HurryWave model and the conditions in which it performs best.

3
Research methodology

This study aims to evaluate the performance of the HurryWave model for the North Sea by comparing its results with historical wave measurements, and existing wave models. The methodology follows a structured approach involving data acquisition and preprocessing, model simulations, and statistical validation against buoy data. An overview of the methodology can be found in figure 3.1

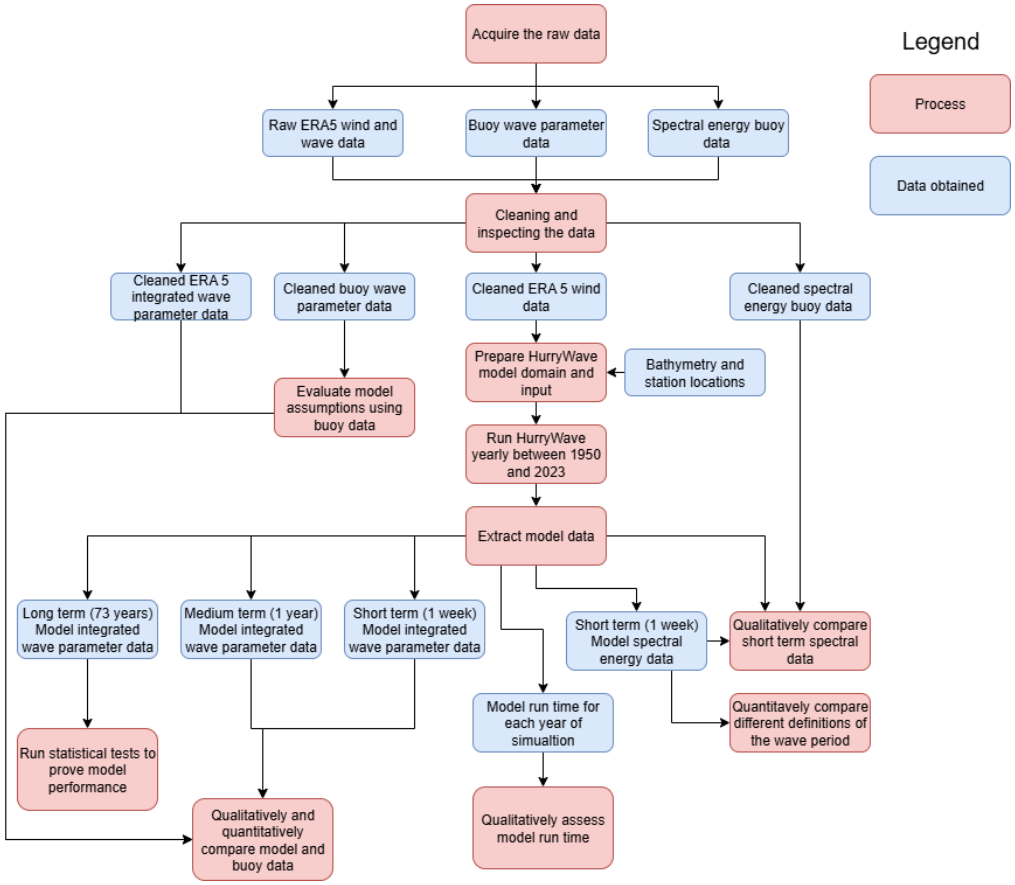

Figure 3.1: Overview of the methodological workflow followed in this study, including data collection and preprocessing, model setup, simulation, and validation. The diagram summarizes the full analysis process—from raw data acquisition (buoy, ERA5, and spectral datasets) through data cleaning and model preparation, to reanalysis simulations and performance evaluation using statistical and spectral analyses.

From figure 3.1, the methodology is structured as follows:

- The methodology used to collect and clean data is explained in section 3.1.
- The methodology used to prepare HurryWave domain and input is shown in sections 3.4 and 3.5.
- Running the model is explained in sections 3.6 and 3.7.
- The qualitative assessment of the model run time is also described in section 3.7.

- The qualitative and quantitative comparison of long, medium and short term model data to buoy data and ERA5 data is explained in subsection 3.8.1.
- The statistical tests proving model performance are laid out in subsection 3.8.2.
- The comparison of short term spectral data is explained in section 3.9.

3.1. Data collection

The first step involves acquiring and preprocessing wave data from multiple sources. The datasets are categorized into three main groups: buoy data collected from Rijkswaterstaat, long-term climate reanalysis data from ERA5, and spectral buoy data which measures the energy and direction of the waves at specific locations.

One of the datasets used to compare to the model output is the wave buoy data from Rijkswaterstaat's monitoring network. These buoys provide high-frequency measurements of wave height and period with a resolution of 10 to 30 minutes ("Wave measuring buoy | Digital North Sea", n.d.). These are matched to the output resolution of ERA5, which is 1 point per hour. The variables extracted from the wave buoy dataset include:

- Measured significant wave height (H_s) in cm
- Mean wave period of the highest 1/3 of waves ($T_{1/3}$)
- Mean wave direction (θ_m)

In this case the significant wave height (H_s) is converted to meters from centimeters and is assumed to be equivalent to the spectral significant wave height H_{m0}. This equivalency assumption means the waves are assumed to be Rayleigh distributed. This assumption is valid in deep water where the sea surface represents a stationary Gaussian process (Leo. H. Holthuijsen, 2007). Additionally, peak wave period is not available as that would require spectral analysis, something these buoys are not equipped for. Therefore the period of the 1/3 highest waves ($T_{1/3}$) is the closest approximation to peak period (T_p). This needs to be considered when observing results as "$T_{1/3}$ is approximately equal to 0.9-0.95 T_p" (Bosboom & Stive, 2023) for relatively wide spectra, which is the case in the North Sea. The parameter $T_{1/3}$ is therefore converted to T_p using a conversion factor of 0.95 so:

$$T_p = \frac{T_{1/3}}{0.95} \tag{3.1}$$

The locations of the buoy measurements are detailed in table 3.1 and a map showing their location is shown in figure 3.2.

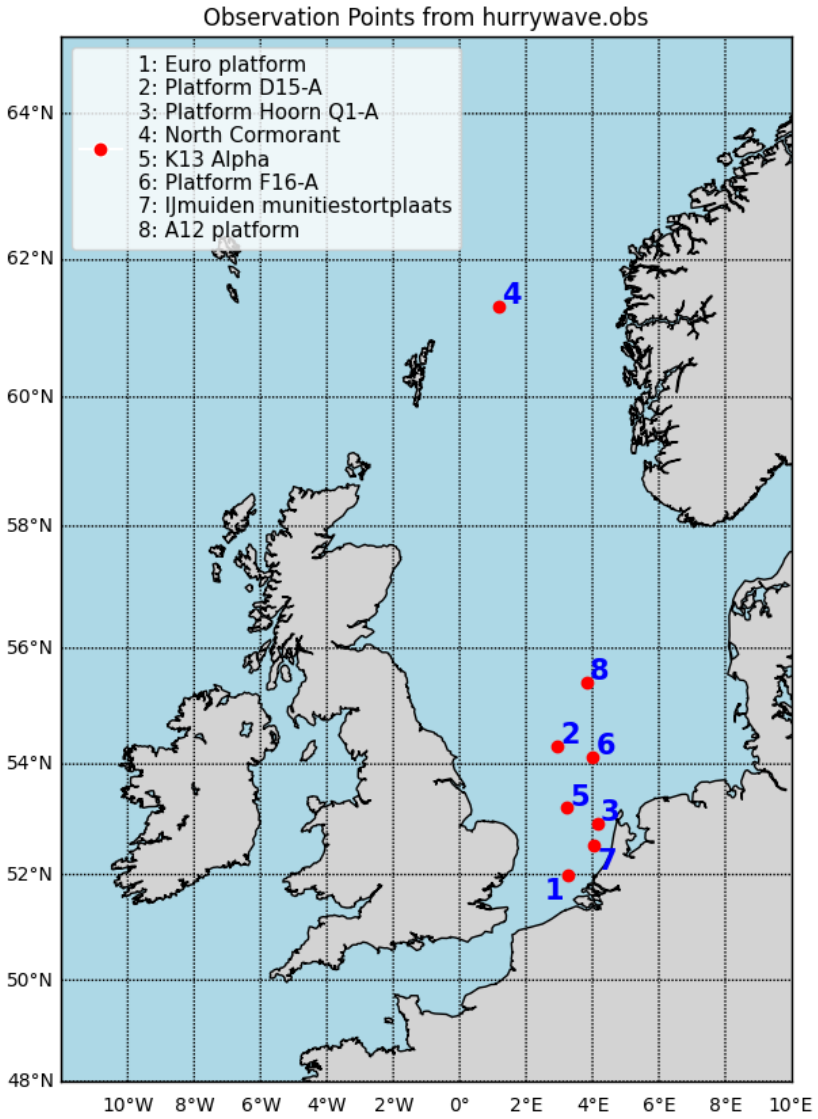

Figure 3.2: Map showing the geographical distribution of the North Sea measurement stations used in this study. Each point marks a Rijkswaterstaat buoy or platform station providing significant wave height, peak period, and mean wave direction data. The spatial coverage ensures representation of both coastal and offshore conditions used for model calibration and validation.

Station name	Latitude [°N]	Longitude [°E]	Depth [m]
Euro platform	51.997799	3.275037	33.66
K13 Alpha	53.217010	3.218932	27.96
IJmuiden munitiestortplaats	52.549214	4.056983	26.42
North Cormorant	61.338188	1.166099	168.64
Platform Hoorn Q1-A	52.925353	4.150286	32.56
Platform D15-A	54.325667	2.935750	35.06
Platform F16-A	54.116667	4.012222	52.74
A12 platform	55.417000	3.817000	34.48

Table 3.1: Station coordinates in the North Sea (using ETRS89 standard)

The amount of data available is given through the start and end date of the available recordings of each wave parameter as shown in table 3.2. Platform A12 does not contain any integrated parameter data, however, it does contain spectral data extracted separately.

Station	H_{m0}			T_p			θ_m		
	Start	**End**	**Years**	**Start**	**End**	**Years**	**Start**	**End**	**Years**
Euro platform	1982-12-31	2023-01-01	41	2007-12-31	2023-01-01	16	1985-12-31	2023-01-01	38
K13 Alpha	1978-12-29	2023-01-01	45	2007-12-31	2023-01-01	16	1985-12-31	2023-01-01	38
IJmuiden munitiestortplaats	1976-05-01	2023-01-01	47	2007-12-31	2023-01-01	16	1989-12-31	2023-01-01	34
North Cormorant	1987-12-24	2023-01-01	36	–	–	–	1987-12-24	1989-02-28	2
Platform Hoorn Q1-A	2009-12-31	2022-01-01	13	2009-12-31	2022-01-01	13	2009-12-31	2022-01-14	13
Platform D15-A	2009-12-31	2017-01-01	8	2009-12-31	2017-01-01	8	–	–	–
Platform F16-A	2009-12-31	2020-01-01	11	2009-12-31	2020-01-01	11	–	–	–
A12 platform	-	-	-	-	-	-	–	–	–

Table 3.2: Overview of available H_{m0}, T_p, and θ_m data at each station, with start/end years and record length in years.

The ERA5 dataset, provided by the European Centre for Medium-Range Weather Forecasts (ECMWF), offers hourly estimates of wave parameters such as significant wave height, peak wave period, and wave direction (ECMWF, n.d.-a). This dataset is uploaded onto the Snellius supercomputer and processed to extract relevant variables, focusing on the North Sea region in the same area as the one run by the model shown in figure 1.1. These variables include the items listed below.

- 10m u component of wind in m/s
- 10m v component of wind in m/s
- Significant wave height of ocean waves (H_{m0}) in m
- Peak period of ocean waves (T_p) in s
- Mean wave direction (θ_m) in degrees

The 10m u and v components of wind enable the creation of the meteo files which are spatially varying wind fields in the North Sea. These are extracted while creating the input files before the simulation during preprocessing. The significant wave height, peak period and mean wave direction are the target variables that the HurryWave model tries to emulate and are used as a comparison benchmark in postprocessing.

The wind data is collected for each year and downloaded using the Climate data store API to the Snellius high performance computer. This enables fast access to the data directly while preprocessing and postprocessing.

The spectral buoy data used further in section 3.9 comes from three different stations, A12 platform, Hoorn Q1-A and Euro platform, all three of which can be found on the map in figure 3.2 and table 3.1.The availability of the spectral buoy data is shown in table 3.3. The data includes the wave energy, the frequency at which it is measured, and the direction from which it comes. This dataset is provided by the authors of the paper by Rijnsdorp et al. (2021), which itself describes this dataset in length. The data providers also shared a script needed to read the data. The spectral buoy data shows energy (E) per degree per frequency, its unit is therefore $m^2/degree/Hz$.

Table 3.3: Overview of spectral buoy data used in this study.

Station	Data period
A12 platform	2009–2013
Hoorn Q1-A	2009–2013
Euro platform	2010–2015

3.2. Cleaning and inspecting the data

The data from the wave buoys are checked for outliers by imposing a maximum peak period of 25s, and a maximum wave direction of 360 degrees to avoid large outliers and faulty measurements. The peak period limit is chosen as to not include infragravity waves which are not in the scope of this research as detailed in section 1.4. This is done as these infragravity waves are not modeled by HurryWave given the minimum frequency modeled is 0.04 Hz shown further in section 3.5. A table summarizing the amount of data points from each station before and after cleaning the data for peak period and mean wave direction is detailed in tables 3.4, and 3.5.

Station	# Points before	# Outliers	# Points after
IJmuiden munitiestortplaats	2,346,635	6	2,346,629
Platform D15-A	782,150	0	782,150
K13 Alpha	2,375,595	11,757	2,363,838
North Cormorant	960,617	0	960,617
Euro platform	2,156,478	0	2,156,478
Platform Hoorn Q1-A	726,982	1,095	725,887
Platform F16-A	608,332	0	608,332

Table 3.4: Cleaning statistics for Tp at all stations

Station	# Points before	# Outliers	# Points after
IJmuiden munitiestortplaats	2,346,629	1,156	2,345,473
Platform D15-A	782,150	0	782,150
K13 Alpha	2,363,838	315	2,363,523
North Cormorant	960,617	0	960,617
Euro platform	2,156,478	3,840	2,152,638
Platform Hoorn Q1-A	725,887	0	725,887
Platform F16-A	608,332	0	608,332

Table 3.5: Cleaning statistics for mean wave direction at all stations

The data is visualized by means of a scatter plot between the significant wave height and the peak period to help identify large outliers and sea states that are not physically possible due to the waves being too steep following the Goda (2010) method. This method helps eliminate outliers by defining a maximum wave steepness. To do this, firstly, wave length is calculated using the iterative method as described in the following equations.

$$\omega^2 = gk \tanh(kd), \tag{3.2}$$

where $\omega = 2\pi/T_p$ is the angular frequency [s^{-1}], T_p [s] the peak period, $k = 2\pi/L$ the wave number [m^{-1}], d [m] the water depth and g the gravitational acceleration [m s^{-2}]. Writing the dispersion relation in terms of L yields the following equation.

$$L = \frac{gT_p^2}{2\pi} \tanh\left(\frac{2\pi d}{L}\right). \tag{3.3}$$

Because L appears on both sides of the equation, L is solved iteratively. An efficient fixed-point iteration used here is initialized with the deep-water wave length

$$L_0 = \frac{gT_p^2}{2\pi}, \tag{3.4}$$

and then updated by

$$L_{n+1} = \frac{gT_p^2}{2\pi} \tanh\left(\frac{2\pi d}{L_n}\right), \tag{3.5}$$

until $|L_{n+1} - L_n| < \varepsilon$, where ε is the tolerance indicating the maximum allowed error set here to 10^{-6}.

This calculation relies on the assumptions of linear wave theory: small-amplitude waves ($H \ll L$ and $H \ll d$), irrotational and inviscid flow, monochromatic (single-frequency) wave components, and negligible dissipation (no breaking, no bottom friction) (Leo. H. Holthuijsen, 2007).

Once this is done, the ratio of the significant wave height to the wave length can be calculated, representing a steepness of the sea state. A maximum threshold can be applied using the method described in Goda (2010).

$$\frac{H_{m0,b}}{d} = \frac{A}{d/L_0}\left\{1 - \exp\left[-1.5\pi\frac{d}{L_0}\left(1 + 11s^{4/3}\right)\right]\right\} \tag{3.6}$$

with:

- H_{m0b} = breaking significant wave height [m]
- d = water depth at the station [m]
- $A = 0.12$ = empirical constant for significant wave height from Goda (2010) [-]
- L_0 = deep water wave length [m] obtained from equation 3.4
- s = bed slope (taken to be 0 in this case) [-]

This equation gives a maximum possible wave height at the location of each station for a given deep water wave length. This can be converted back to peak wave period using equation 3.4. From this the maximum significant wave height for a given period can be visualized.

This visualization is shown at IJmuiden munitiestortplaats in figure 3.3

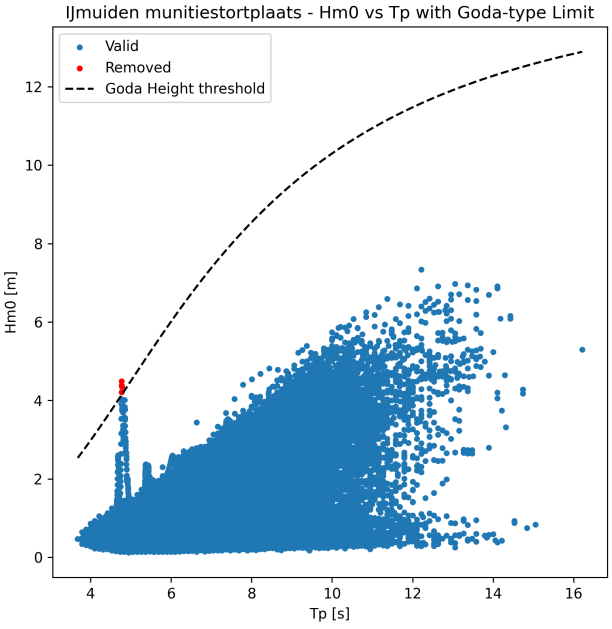

Figure 3.3: Scatter plot comparing significant wave height (H_{m0}) and peak period (T_p) for all buoy data points at IJmuiden munitiestortplaats. The dashed curve shows the theoretical steepness limit following Goda (2010). Points exceeding this limit are classified as outliers and removed. The plot illustrates the physically consistent relationship between wave height and period.

The outliers removed occurred at IJmuiden Munitiestortplaats, where 4 outlier points are removed. No other outliers are removed from the data. Other points lying outside what is expected are observed such as the ones shown in figure 3.3 with a period of around 5s with over 2m significant wave height. In other stations these points which seem to be measurement errors are observed but not removed. Another example here at platform K13 Alpha where values around 0 in significant wave height are observed. Values seeming to share the same period with different significant wave heights are also observed. Lastly, peak periods above 20 seconds are observed. These seem far from all other measurement points, however, they do fall within the limits imposed by this methodology.

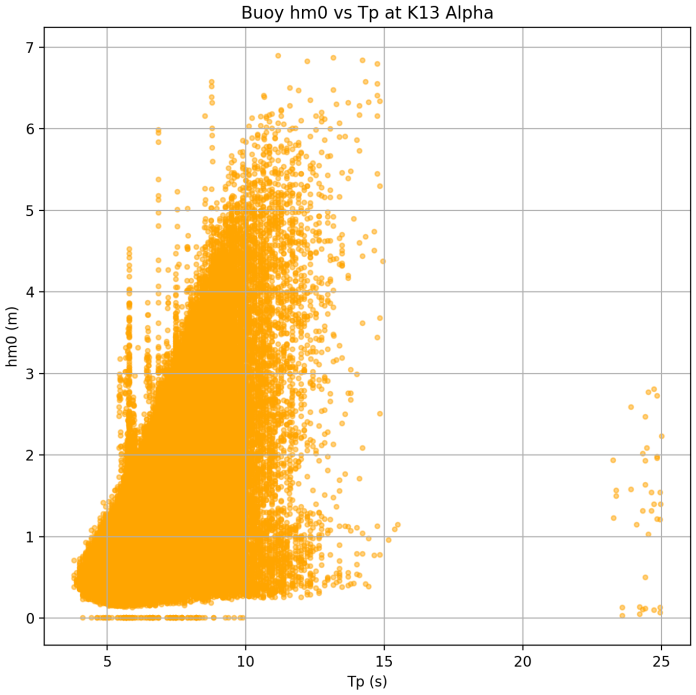

Figure 3.4: Scatter plot comparing significant wave height (H_{m0}) and peak period (T_p) for all buoy data points at K13 Alpha. Here points sharing the same period for large differences in wave heights, points having a near 0 value for H_{m0}, and points with large periods are shown as potential outliers which are not removed in this framework

A second plot looks to the comparison between wave height and wave direction. This comparison is shown in figure 3.5. Wave directions used in this document follow the meteorological convention, stating where the wave *comes from*. This means a 0 degree wave direction shows waves *coming from* the North and *going towards* the South.

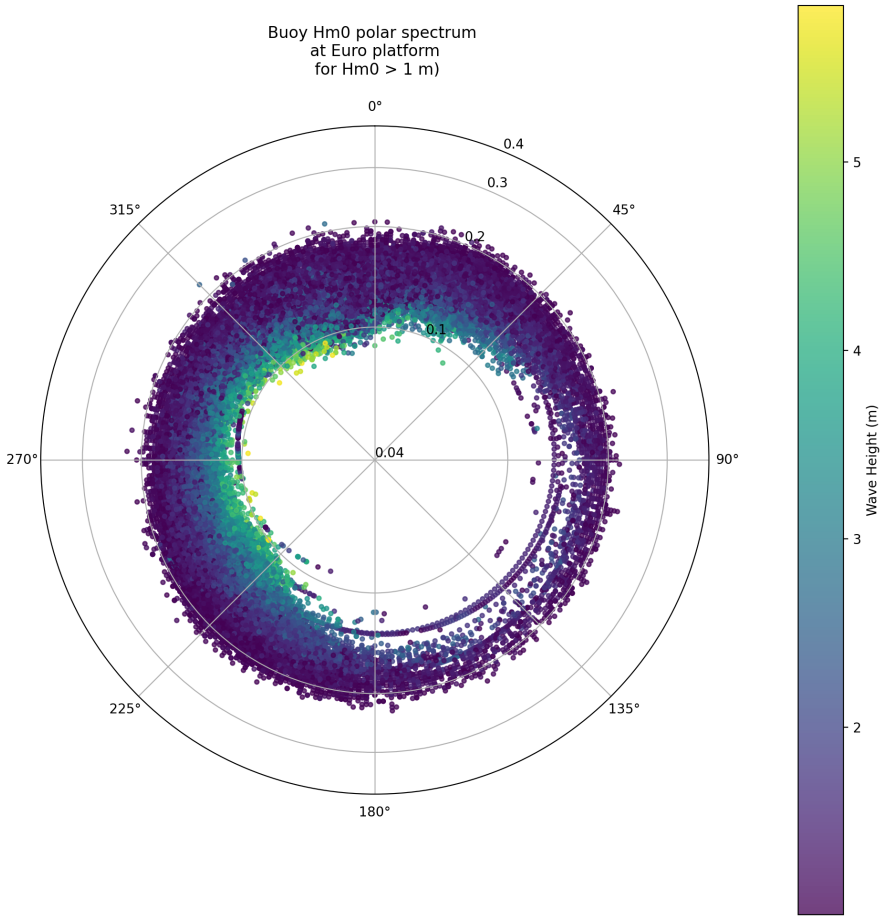

Figure 3.5: Polar plot of significant wave height (H_{m0}) as a function of peak frequency and wave direction for the Euro platform buoy. The color scale indicates wave height, revealing that the largest waves occur at lower frequencies (longer periods) and predominantly arrive from the southwest and north sectors. The concentric frequency lines emphasize the dominance of short-period wind seas (0.1–0.2 Hz) in the North Sea.

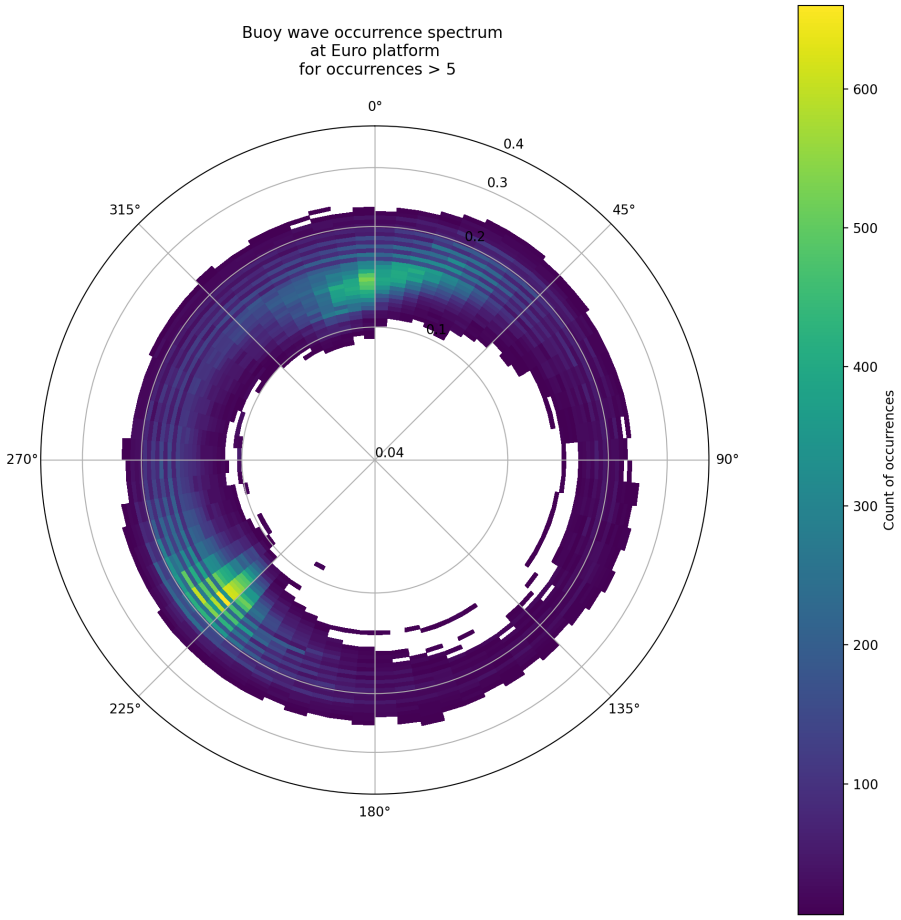

Figure 3.6: Polar plot of the occurrence density of wave events at the Euro platform, as a function of direction and frequency. Two dominant wave systems are evident: one from the southwest and another from the north. The frequency distribution highlights that the most common waves have periods between 5 and 10 seconds, corresponding to short wind-generated seas. Longer-period swells (>10 s) originate mainly from the north, consistent with remote wave generation.

The plot in figure 3.5 shows the significant wave height of waves as a function of their frequency and direction. The trend showing larger waves with larger periods (and thus smaller frequencies) shown in figure 3.3 is also shown in this plot where larger wave heights are found close to the center of the polar plot. The plot in figure 3.6 shows the occurrence of different combinations of wave direction and period. In this way, one can find the dominant wave directions and periods. This plot also shows two dominant states, one from the South West, and one from the North. The dominant frequency is shown to be between 0.1 and 0.2 Hz, meaning a dominant period between 5 and 10 seconds. This is consistent with short wind waves observed in the North Sea. Most longer waves (above 10 seconds) come from the North, traveling a larger distance to arrive to the stations placed near the Netherlands.

3.3. Evaluating model assumptions at stations in the data

The HurryWave model runs on the assumption of intermediate to deep water as shown in section 2.2. To verify this, the ratio of the depth to the wave length is calculated (d/L). To do so, first the wave length is calculated using the iterative method applied in linear wave theory as shown in the previous paragraph. A plot of the the ratio of depth to wave length is shown in figure 3.7

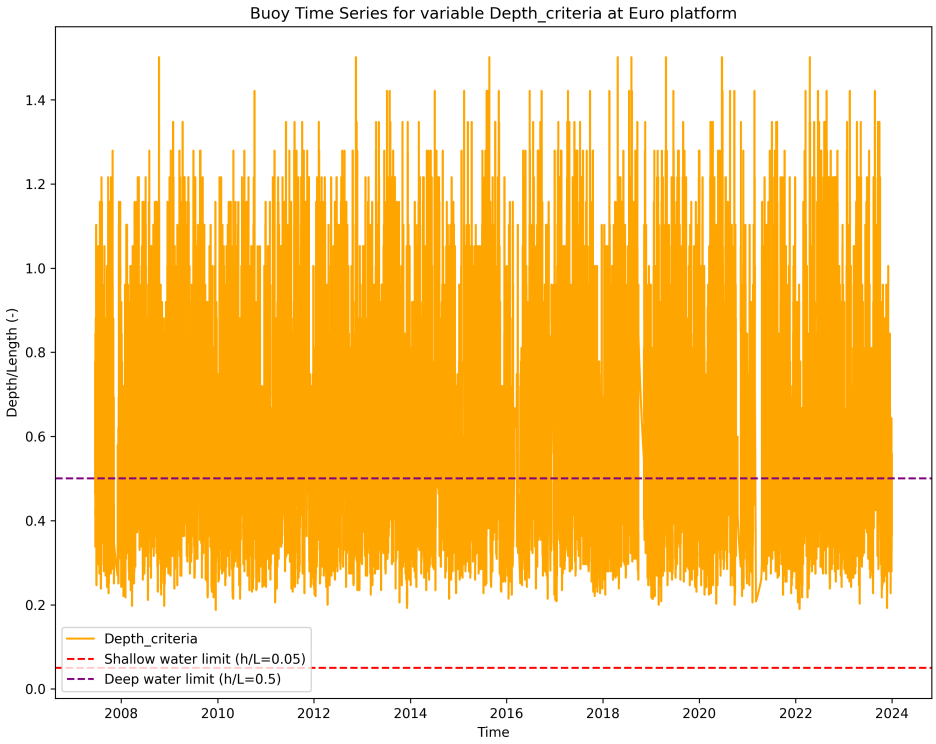

Figure 3.7: Time series of the ratio between local water depth and computed wavelength (d/L) at the Euro platform station. The shaded regions correspond to the thresholds for shallow, intermediate, and deep-water conditions defined in table 3.6. The results confirm that all observed sea states at this station occur in intermediate to deep-water regimes, validating the core assumptions of the HurryWave model.

The criteria for being in shallow, intermediate or deep water is shown in table 3.6 based on Bosboom and Stive (2023).

Depth assumption	**Value of d/L**
Shallow	$d/L < 0.05$
Intermediate	$0.05 < d/L < 0.5$
Deep	$0.5 < d/L$

Table 3.6: Depth assumptions based on the depth to wave length ratio (Bosboom & Stive, 2023)

Figure 3.7 shows that the depth to wave length ratio at Euro platform never goes below 0.05 and stays between intermediate and deep water. This is true for all stations shown in table 3.1, with deeper stations showing larger amounts of "deep water" waves.

3.4. Preparing HurryWave model domain

To run the model in the North Sea, the domain needs to be defined. The domain extends from 48 to 65 degrees in latitude and from -12 to 10 degrees in longitude. These values are conserved even when changing spatial resolution through the changing of the number of grid cells.

The bathymetry is generated using a dep file collected from the validation simulation based on the rectangular DCSMv6 grid of the operational Dutch North Sea flow model (van der Lugt et al., 2024). This bathymetry data is derived from the October 2016 version of the European Marine Observation and Data Network which was compounded from several surveys. Data missing from the surveys is filled with GEBCO 30 bathymetry (Zijl & Groenenboom, 2019). The domain used in the North Sea is shown in figure 3.8 and the bathymetry used in this domain is shown in figure 3.9.

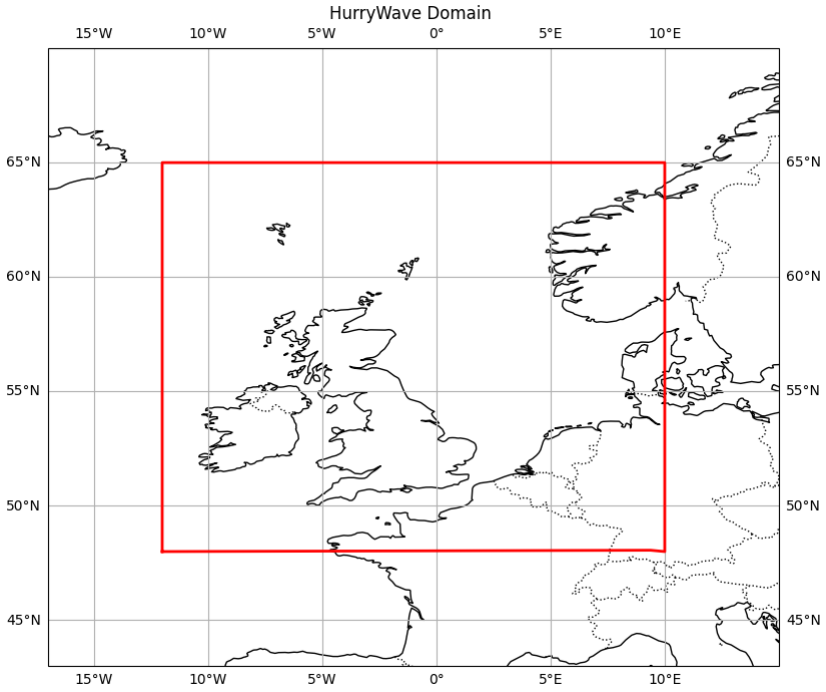

Figure 3.8: Geographic extent of the HurryWave computational domain covering the North Sea. The model grid spans latitudes 48°–65° N and longitudes –12°–10° E, encompassing the full basin from the English Channel to the Norwegian Sea. This domain aligns with ERA5 reanalysis coverage and the Rijkswaterstaat buoy network used for validation.

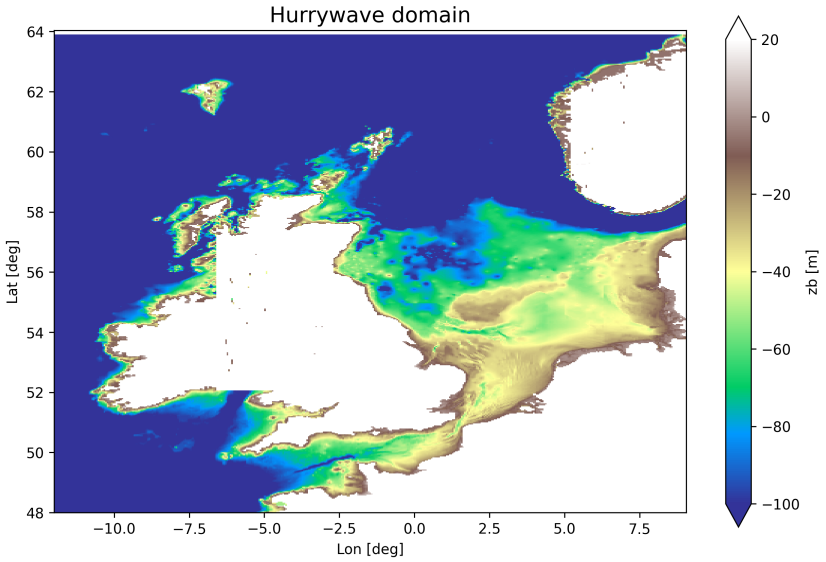

Figure 3.9: Bathymetric map used in the HurryWave model, derived from the DCSMv6 grid and the EMODnet 2016 database, with gaps filled using GEBCO 30 m data. Depths are shown in meters below sea level. The map captures the key morphological features of the North Sea, including the shallow southern shelf and deeper offshore basins, which influence wave propagation and energy dissipation.

Once the bathymetry is known, a mask is applied to identify which points are active and considered as water and which are inactive and considered as land. In the water the mask is identified as 1 and on land it takes the value of 0. On the boundary the mask has a value of 2. In figure 3.10 the mask is shown. Some discrepancies exist due to the shallow area between Ireland and Great Britain considered as land. The province of Zeeland in the Netherlands is also considered water due to the land being below Sea Level.

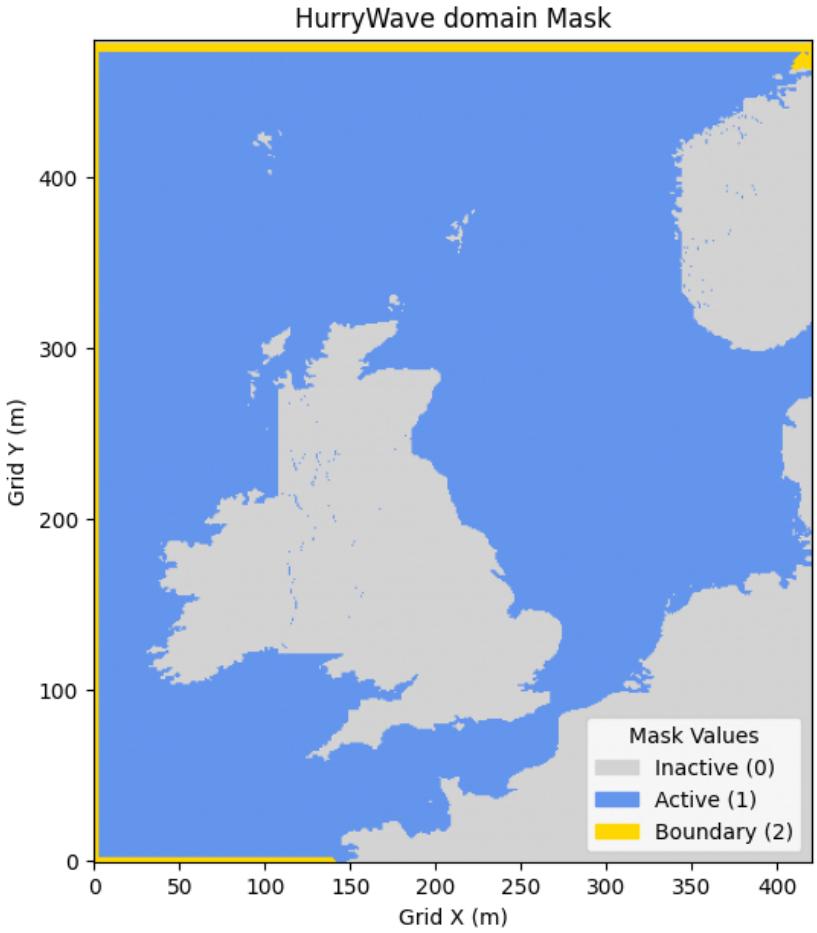

Figure 3.10: Binary mask used to define active (water) and inactive (land) grid points in the HurryWave model domain. Blue areas indicate ocean cells (value = 1), gray areas represent land (value = 0), and red lines mark the open boundaries (value = 2). The mask highlights minor discrepancies due to topography below sea level (e.g., Zeeland region) and shallow regions between Great Britain and Ireland.

The boundary is set as an open boundary and no boundary conditions are applied.

3.5. Model input parameters

The HurryWave model is initialized with a domain covering the North Sea, and input data is prepared accordingly. The model is forced using ERA5 wind fields and does not use boundary conditions. The physical parameterizations, including wind-wave interactions, dissipation mechanisms, and resolution, is set based on standard configurations used in similar wave modeling studies (van der Lugt et al., 2024). These parameters are shown in table 3.7.

Table 3.7: Default Model Parameters

Parameter	Description	Value	Unit
mmax	Number of grid cells in x-direction	421	-
nmax	Number of grid cells in y-direction	481	-
dx	Grid size in x-direction	0.05	degrees
dy	Grid size in y-direction	0.03333333	degrees
x0	X-coordinate of first grid cell corner	-12	degrees
y0	Y-coordinate of first grid cell corner	48	degrees
rotation	Grid rotation (anti-clockwise)	0.0	degrees
latitude	Latitude reference (not standard)	0.0	degrees
dt	Time-step size for the model	300	seconds
tspinup	Internal spin-up duration	1209600	seconds
t0out	Time for initial output	-999.0	seconds
dtmapout	Global map output interval	3600	seconds
dthisout	Observation points output interval	3600	seconds
dtrstout	Restart file output interval	0.0	seconds
dtsp2out	Specific obs output interval	3600.0	seconds
dtmaxout	Max level map output interval	0.0	seconds
trstout	Restart output time	-999.0	seconds
dtwnd	Wind update interval	1800.0	seconds
rhoa	Air density	1.25	kg/m^3
rhow	Water density	1024.0	kg/m^3
dmx1	Initial dispersion in x-direction	0.2	-
dmx2	Final dispersion in x-direction	1e-05	-
crsgeo	CRS flag for geodetic projection	0	-
freqmin	Minimum wave frequency	0.04	Hz
freqmax	Maximum wave frequency	0.5	Hz
nsigma	Number of frequency bins	12	-
ntheta	Number of directional bins	36	-
crs_name	CRS name	ETRS 89	-
crs_type	CRS type	geographic	-
crs_utmzone	UTM zone (if applicable)	nil	-
crs_epsg	EPSG code	4258	-
gammajsp	JONSWAP peak enhancement factor	3.3	-
spinup_meteo	Meteorological spin-up flag	1	-
quadruplets	Activate quadruplet interactions	1	-
redopt	Redundancy option	1	-
winddrag	Wind drag formulation	zijlema	-
cdcap	Cap on drag coefficient	0.0025	-
fbed	Bed friction coefficient	0.019	-
vmax_zijlema	Max velocity for Zijlema drag	50.0	m/s

A description and potential effects of the choice of parameters is given in Appendix B.

After the simulation, results are extracted in the form of a time series over the period of the simulation. The spin-up time is defined as 31 days as a safe amount of time to propagate the energy across the North Sea. An analysis of the suitability of this spin-up time is shown in section 4.2.

3.6. Reanalysis using default settings

To assess the performance of HurryWave, a reanalysis is conducted using the default HurryWave settings described in the model setup. These settings are then applied to the simulation of multiple years from 1950 to 2023. The simulations last from 01 December of the previous year to 31 December of the studied year. The extra month accounts for the spin-up time described in section 3.5. This enables the propagation of energy throughout the system before the start of result collection.

The results are analyzed at the location of the wave buoys shown in figure 3.2 in section 3.1. This investigation is split into three parts of different time length. Firstly, the analysis of a single year, to analyze patterns on a medium term and look into the performance of HurryWave in different conditions. Secondly, the analysis of multiple years, ranging from 1950 to 2023 to look at the ability of HurryWave to model long term datasets. Lastly, a case study looking into the first week of the December 2013 is analyzed to look into the performance of HurryWave in the short term. This particular week is when the Sinterklaas storm (also known as storm Xaver) occurred, which gives insight into HurryWave's performance in the short term. This short term case study serves to analyze the impact of the time scale on the observed differences.

3.7. Computational considerations

The model is run on Snellius ("System details", n.d.), a high-performance computing (HPC) system. The most efficient setup uses a single node in the gpu_h100 partition, which is equipped with NVIDIA H100 GPUs. The job is configured to run up to 16 parallel tasks, each assigned one CPU core. This configuration balances CPU and GPU usage effectively for optimal performance.

All python notebooks, preprocessing and postprocessing are run on the scratch drive of the HPC system. Some heavy postprocessing tasks also require a job submission where the CPU is used. In this case, the model is configured to run on a single node in the rome partition, which consists of AMD EPYC CPU nodes. It uses 128 parallel tasks on that node, making use of the many-core architecture to achieve efficient parallel computation.

Using the setup laid out in the first paragraph, the model performance of HurryWave is assessed in terms of computational time. This is done qualitatively by recording the time needed to run each simulation year, and looking qualitatively into the similarities and differences between simulated years.

3.8. Uncertainty quantification

3.8.1. Metrics used

To assess the performance of the model output at various measuring stations, the results are validated against observational datasets using several statistical methods. These metrics quantify the agreement between model predictions and observations for key output, such as significant wave height and peak period. The analysis is conducted for both the reanalysis run and the sensitivity experiments against both buoy data and ERA5 data collected at the closest grid point to the measurement buoy. The key quantitative performance indicators are listed in the following points.

Bias

Bias measures the mean difference between model output and observations, indicating any systematic over- or under-estimation.

- **Purpose:** Detects systematic error.
- **Interpretation:** A bias close to 0 indicates no systematic deviation.

$$\text{Bias} = \frac{1}{N}\sum_{i=1}^{N}(M_i - O_i) \tag{3.7}$$

Root Mean Square Error (RMSE)

The Root Mean Square Error (RMSE) measures the average magnitude of the errors between model predictions and observations in the same unit as the former. It gives greater weight to larger errors and is sensitive to outliers.

- **Purpose:** Evaluates the overall deviation between model predictions and observations.
- **Interpretation:** Lower values indicate better agreement; RMSE = 0 implies perfect fit.

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)^2} \tag{3.8}$$

where M_i is the model value, O_i is the observed value, and N is the number of data points.

Scatter Index (SI)

The Scatter Index (SI) is a normalized metric representing the spread of errors relative to the observed mean.

- **Purpose:** Measures relative variability between model and observations.
- **Interpretation:** Lower SI values indicate better agreement. Values below 0.2 are considered very good, and values between 0.2 and 0.4 are considered acceptable in wave modeling as shown in the validation of SWAN where the values are around 37% for the significant wave height and 20% for the mean wave period (Ris et al., 1999)

$$\text{SI} = \frac{\text{RMSE}}{\overline{O}} \tag{3.9}$$

where $\overline{O}$ is the mean of the observed values.

Pinball Loss Function

The Pinball Loss Function is used to evaluate the accuracy of quantile forecasts. It is especially useful in probabilistic modeling, capturing how well a model predicts the distribution of the target variable.

- **Purpose:** Assesses forecast accuracy at different quantile levels.
- **Interpretation:** Lower values indicate more accurate quantile predictions. A value of 0 represents a perfect forecast at the given quantile.

$$\mathcal{L}_\tau(\hat{y}, y) = \frac{1}{N}\sum_{i=1}^{N}\max\left(\tau(y_i - \hat{y}_i), (\tau - 1)(y_i - \hat{y}_i)\right) \tag{3.10}$$

where τ is the quantile level (e.g., 0.5 for the median), $\hat{y}_i$ is the predicted quantile, and y_i is the observed value.

The Pinball Loss is computed for the following quantiles:

$$\tau = 0.05,\ 0.1,\ 0.25,\ 0.5,\ 0.75,\ 0.9,\ 0.95$$

The parameter τ determines the asymmetry of the penalty: if the prediction underestimates the observation, the error is weighted by τ, whereas if it overestimates, the error is weighted by $(1 - \tau)$. This ensures that the loss function targets the τ-th quantile of the conditional distribution.

Angular error

The Angular Error is used when comparing two angles, for example when evaluating mean wave or wind directions. It ensures that the comparison accounts for the circular nature of angles by measuring the shortest angular distance between two directions.

- **Purpose:** Quantifies the difference between modeled and observed angles while accounting for circularity.
- **Interpretation:** An error of 0 indicates perfect directional agreement. The error is typically expressed in radians but can be converted to degrees. Lower values indicate closer alignment of directions.

$$\Delta\theta = \arctan2\big(\sin(\theta_M - \theta_O), \cos(\theta_M - \theta_O)\big) \tag{3.11}$$

where θ_M is the modeled angle, θ_O is the observed angle, and $\Delta\theta$ is the angular error in radians.

Once $\Delta\theta$ is computed, common error metrics such as RMSE, bias, and standard deviation can be derived by substituting the linear difference term $(M_i - O_i)$ in their standard definitions with the angular error. These metrics may be converted back to degrees or normalized by dividing by 2π to obtain a dimensionless fraction or percentage.

3.8.2. Autocorrelation-Aware Performance Evaluation

The evaluation of numerical wave model performance requires not only the calculation of accuracy metrics, but also the quantification of their uncertainty and statistical significance. Standard statistical tests assume independent and identically distributed (i.i.d.) samples, an assumption that is violated in environmental time series due to temporal autocorrelation (Ritter & Muñoz-Carpena, 2013). Neglecting this dependence leads to underestimated standard errors and inflated significance levels. The methodology adopted here therefore accounts for autocorrelation explicitly and combines multiple resampling and testing approaches. The goal of this methodology is to rigorously assess the bias and scatter index and check to see if the model can be demonstrated to be accurate based on these metrics.

Lag-1 Autocorrelation and Effective Sample Size.

The error series often exhibits serial correlation. The lag-1 autocorrelation coefficient r_1 is collected from the autocorrelation function (ACF) of the error between modeled and observed points **e**. The effective sample size N_{eff} is then computed following the method described in Bretherton et al. (1999).

$$N_{\text{eff}} = N\,\frac{1 - r_1}{1 + r_1}. \tag{3.12}$$

Here, N is the number of valid data points (unitless) and r_1 is the value of the autocorrelation at the first lag which is dimensionless. This adjustment ensures that statistical tests reflect the reduced degrees of freedom caused by autocorrelation. This number informs if the need for autocorrelation correction exists, as if the number of valid data points is much lower than the number of data points, other methods than standard statistical tests need to be used to ensure the independent identically distributed assumption.

Visualization of Autocorrelation.

For interpretability, the full ACF of both the time series and the error series are plotted up to a pre-defined number of lags. This plot provides a direct visual check on the magnitude and persistence of autocorrelation, which informs the choice of block size B and validates the need for autocorrelation correction.

Block Bootstrap hypothesis testing

Due to autocorrelation within the time series, standard statistical tests do not respect the iid assumption and are therefore invalid. Standard bootstrap methods as proposed by Efron (1979) are also inappropriate as they break the time correlation between the waves. Instead, a block bootstrap is applied, where contiguous blocks of data of length B (here, 1 year) are resampled with replacement until a series of length N is reconstructed following the methodology described in Kunsch (1989) and Ritter and Muñoz-Carpena (2013). This procedure is repeated n_{boot} times to form a bootstrap distribution of the performance metric values:

1. **Filter valid pairs**

   Remove any observation-prediction pairs containing NaN values, ensuring only comparable data points are used in metric calculations.

2. **Compute block indices**

   Let n be the number of valid pairs, and select a block size l appropriate for the observed autocorrelation, in this 1 year or 365 days or 8760 hours. Partition the time series into $n_{\text{blocks}} = \lceil n/l \rceil$ consecutive blocks.

Metric and parameter	Threshold t
Bias for both H_{m0} and T_p	0m or 0s
RMSE for H_{m0}	0.5m
RMSE for Tp	1s
SI for both H_{m0} and T_p	0.4

Table 3.8: Thresholds used in hypothesis testing

3. **Bootstrap resampling**

   For each bootstrap iteration $b = 1, \ldots, B$:

   - Randomly select n_{blocks} blocks with replacement.
   - Concatenate the selected blocks to produce a pseudo time series of length $\sim n$.
   - Compute the metric of interest $\theta_b^* = \texttt{metric\_func}(\text{obs}_b, \text{pred}_b)$ for that bootstrap sample.
   - Store the result.

   This yields a bootstrap distribution $\{\theta_b^*\}_{b=1}^B$.

4. **Confidence intervals**

   Calculate empirical $(2.5, 97.5)$ percentile intervals for the resulting metric bootstrap distribution:

   $$\text{CI}_{95\%} = [q_{0.025}(\{\theta_b^*\}), q_{0.975}(\{\theta_b^*\})] \tag{3.13}$$

5. **Hypothesis testing**

   For bias:

   $$H_0 : \text{Bias} \geq 0 \tag{3.14}$$

   Compute one-sided p-value as

   $$p_{\text{bias}} = \frac{1}{B} \sum_{b=1}^{B} I(\theta_b^* \geq 0) \tag{3.15}$$

   where $I(\cdot)$ is the indicator function.

   For RMSE and SI, use a null distribution recentered to the desired threshold:

   $$\text{null\_dist}_b = \theta_b^* - \overline{\theta^*} + t \tag{3.16}$$

   where t is the threshold. The thresholds used are shown in table 3.8. The p-value for H_0 : metric $\geq$ threshold is:

   $$p = \frac{1}{B} \sum_{b=1}^{B} I(\text{null\_dist}_b \geq \theta_{\text{obs}}) \tag{3.17}$$

   The observed value θ_{obs} comes from the original (unresampled) data.

6. **Result output and visualization**

   The outputted results are:

   - Observed metric value and 95% CI.
   - p-value for hypothesis test.
   - Bootstrap histogram with empirical CI and test statement.

   A histogram of the empirical distribution collected from the block bootstrap is made with the mean value and the confidence interval

Rationale and Importance

These measures prevent overconfident conclusions about model skill and ensure that statistical inference remains valid for temporally dependent environmental datasets.

3.9. Spectral analysis

The HurryWave model includes the ability to collect spectral energy output at certain specified locations. This output can be used to understand qualitatively the differences between model and measurements in modeling wave energy, density, frequency, provenance and spread.

3.9.1. Extracting spectral data from the model

The spectral data is provided by the model at A12 platform, Hoorn Q1-A and Euro platform as shown in 3.11.

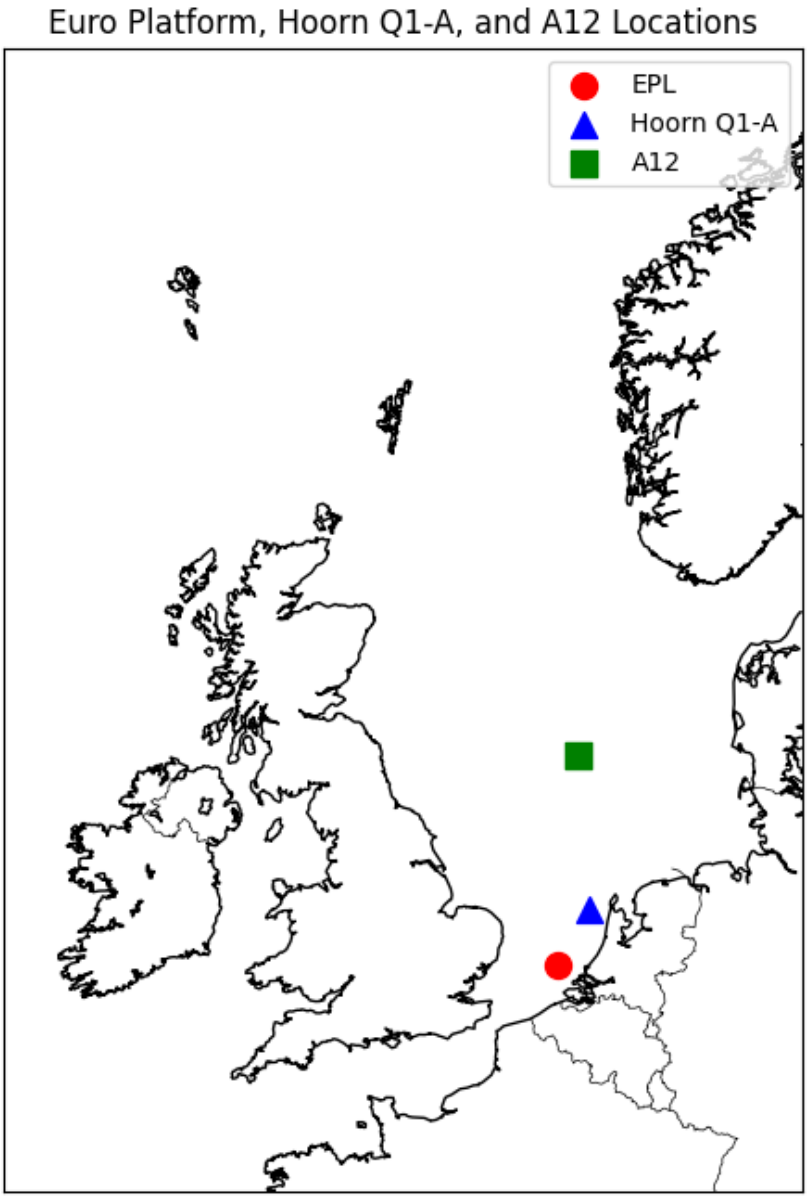

Figure 3.11: Map showing the locations of stations where both modeled and observed wave spectra are analyzed: A12 platform, Hoorn Q1-A, and Euro platform. These sites are selected because they provide high-quality directional wave spectra, enabling direct comparison of modeled and measured wave energy distributions in frequency–direction space.

These are stations where spectra measuring wave buoys exist and can be used to compare against the model output. The extraction of the wave energy spectra, from both the wave buoys and the model, is aided by the wavespectra package ("wavespectra — wavespectra 4.4.1 documentation", n.d.). This package enables the easy calculation of spectral wave parameters and plotting of the spectra themselves. Great care is taken when extracting the spectra to make sure that the units indeed are $m^2/(degree * Hz)$ so that the wavespectra package can read it appropriately ("wavespectra — wavespectra 4.4.1 documentation", n.d.). This is done by making sure the extracted spectral wave buoy data parameters match the data found in Rijnsdorp et al. (2021). Additionally the spin-up time is removed from each simulation and, angles are resolved by making the 0 value North (and not East).

3.9.2. Analyzing spectral data

The modeled spectral data is qualitatively analyzed on a short timescale to look into the short term differences between the HurryWave model and the spectral buoy data. Analyzing the Sinterklaas storm (or storm Xaver) is therefore ideal given the presence of December 2013 in all ranges of data. This will also be compared to the reanalysis short term comparison to explain differences in the integrated parameters such as significant wave height, peak period, and wave direction through the wave energy spectrum.

To further understand the temporal characteristics of the wave field, a quantitative analysis is performed on several definitions of the mean wave period at the stations described in figure 3.11: the mean period based on the energy spectrum (T_{mean}), the first-order spectral moment period (T_{m01}), the second-order spectral moment period (T_{m02}), and the inverse first negative moment period (T_{m-10}). These metrics each provide a different representation of the dominant time scales of ocean waves, depending on how energy is distributed across frequencies in the wave spectrum $S(f)$.

The mean wave period T_{mean} is a general descriptor that can be computed directly from the wave spectrum as the energy-weighted average of all periods:

$$T_{\text{mean}} = \frac{\int_0^\infty TS(f)\,df}{\int_0^\infty S(f)\,df}, \tag{3.18}$$

where $T = 1/f$ is the wave period corresponding to frequency f.

A more standardized definition uses spectral moments m_n, given by

$$m_n = \int_0^\infty f^n S(f)\,df. \tag{3.19}$$

Using these moments, different characteristic periods can be defined. The first-order moment period T_{m01} is expressed as

$$T_{m01} = \frac{m_0}{m_1}, \tag{3.20}$$

which emphasizes the contribution of lower frequencies (longer waves), since it divides total energy by the first spectral moment.

The second-order moment period T_{m02} is defined as

$$T_{m02} = \sqrt{\frac{m_0}{m_2}}, \tag{3.21}$$

which places greater weight on higher frequencies, making it more sensitive to short-period (wind sea) components of the spectrum.

Finally, the negative first moment period T_{m-10} is defined as

$$T_{m-10} = \frac{1}{m_{-1}/m_0} = \frac{m_0}{m_{-1}}, \tag{3.22}$$

where

$$m_{-1} = \int_0^\infty f^{-1} S(f)\, df. \tag{3.23}$$

This formulation enhances the influence of the lowest-frequency components, thus emphasizing long-period swells.

In summary, while all four metrics characterize the average period of sea states, they highlight different parts of the wave spectrum. T_{m01} and T_{mean} tend to represent the overall energy-weighted mean period, T_{m02} emphasizes shorter, wind-driven waves, and T_{m-10} reflects the influence of long-period swell. Comparing these measures allows for a more complete understanding of the spectral shape and energy distribution of the wave field.

4
Results and analysis

4.1. Computation time

The computation time for each yearly simulation was recorded and is plotted in figure 4.1.

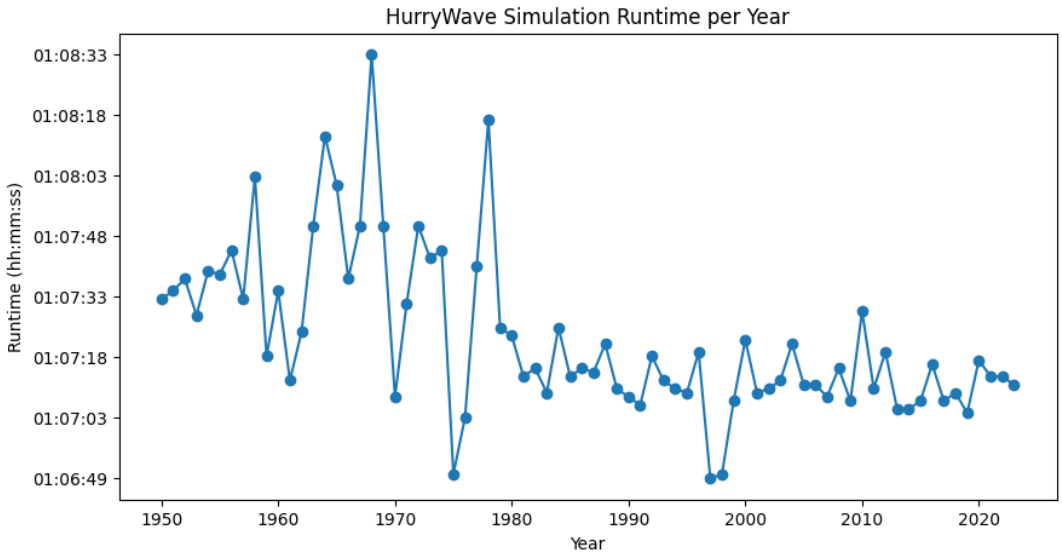

Figure 4.1: Computation time for each yearly HurryWave simulation conducted on a single NVIDIA H100 GPU on the Snellius supercomputer. Each bar represents the wall-clock runtime per simulation year under identical model configurations and grid settings. The minor variation between years (1h 06 min – 1h 09 min) reflects differences in the number of simulated days per year. These results demonstrate that multi-year hindcasts can be efficiently parallelized, enabling full-year simulations to be completed in under 1 h 10 min per GPU.

The times taken by HurryWave to run a 1 year simulation given the input parameters in table 3.7 and the computational conditions shown in paragraph 3.7 vary between 1 hour 6 minutes and 1 hour 9 minutes. These times assume the use of a GPU H100 on the Snellius supercomputer. A comparison with the speed of other models for this case is beyond the scope of this project, the time taken to run such a model heavily depends on the architecture used, the conditions simulated and the model parameters such as timestep, grid size and grid resolution. Literature exists on specific cases, however, these are not comparable to this case. The run times give a good indication that long series of data can be simulated using HurryWave, especially given that years can be simulated separately. This means, given a certain number of GPUs on a supercomputer, one can theoretically run a full simulation year under 1 hour 10 minutes on each GPU (given Snellius possesses 352 Nvidia H100 GPUs (“System details”, n.d.)), thus heavily speeding up the process of simulating long datasets. The data in figure 4.1 is difficult to interpret due to the different amount of days in between the different years. This is why an average computation time per computed day is also plotted and shown to be stable with small variations between 11.62 and 11.70 seconds as shown in figure 4.2.

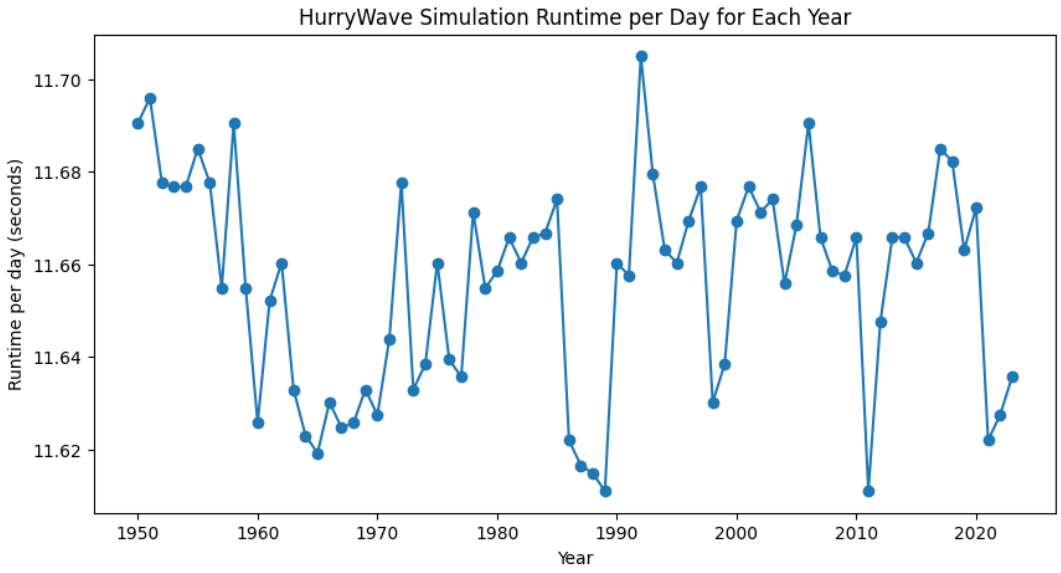

Figure 4.2: Average computation time per simulated day for each yearly HurryWave run. The figure normalizes total runtime by the number of simulated days, showing stable results between 11.62 s and 11.70 s per model day. This consistency highlights the scalability and numerical stability of the model runtime independent of annual data variations.

4.2. Input parameters sensitivity

4.2.1. Model spin-up assessment

An assessment of the spin-up time was conducted following the test runs described in table 4.1. The first run aims to show calm conditions on the North Sea with boundary values for height, period and direction. The second and third run have no boundaries, but have a constant wind field of 10 and 20 m/s respectively across the whole domain. The last run uses only constant boundary conditions with no wind field, meaning in this run the waves propagate with no external input from wind energy.

Table 4.1: Spin-up Test Configurations

Parameter	Run 1	Run 2	Run 3	Run 4
Boundary wave height [m]	1	-	-	1
Boundary peak wave period [s]	6	-	-	6
Boundary wave direction [°N]	315	-	-	315
Boundary wave directional spreading [-]	20	-	-	20
Wind field speed [m/s]	10	10	20	-
Wind field direction [°]	315	315	315	-

The results at the measuring station Euro Platform are shown in figures 4.3 and 4.4. The location of the Euro platform measuring station can be found in figure 3.2 section 3.1. These results show a simulation of 1 month with stationary wind conditions. The plots show the difference in significant wave height (H_{m0}) in figure 4.3 and peak wave period (T_p) in figure 4.4 between the current and the previous timestep. The time resolution of the output is here 1 hour meaning these plots show the difference between the current and previous hour for both variables. These conditions have either constant boundary conditions or constant wind fields, meaning the energy should spread out evenly. This means the spin-up time represents the time before the difference between two time steps becomes 0.

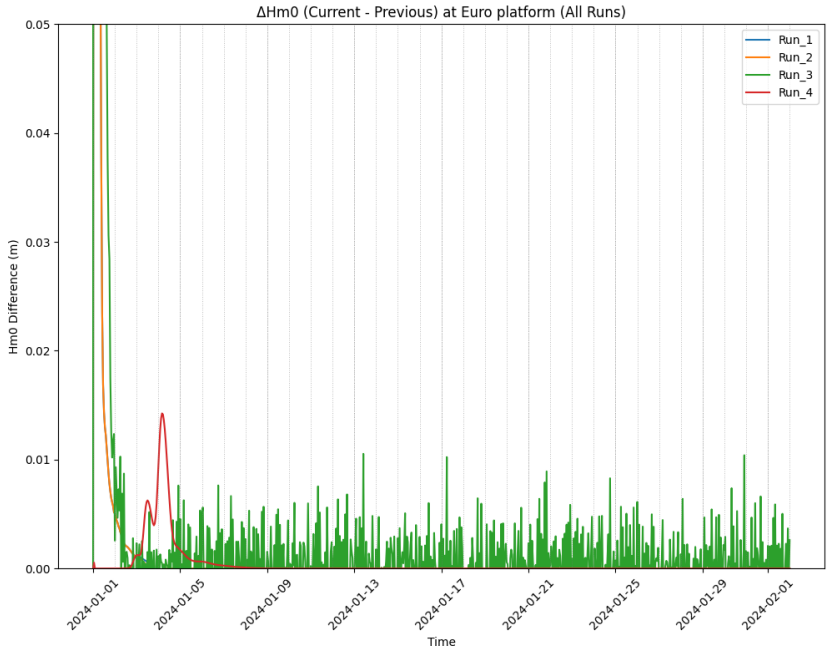

Figure 4.3: Time series of hourly differences in significant wave height (H_{m0}) at Euro Platform during the spin-up assessment. Each curve represents a different test run with varying boundary and wind conditions (see Table 4.1). The plot quantifies the model's approach to stationarity, showing that cases with active wind forcing converge within 2–3 days, while runs with only boundary forcing require up to 9 days to stabilize.

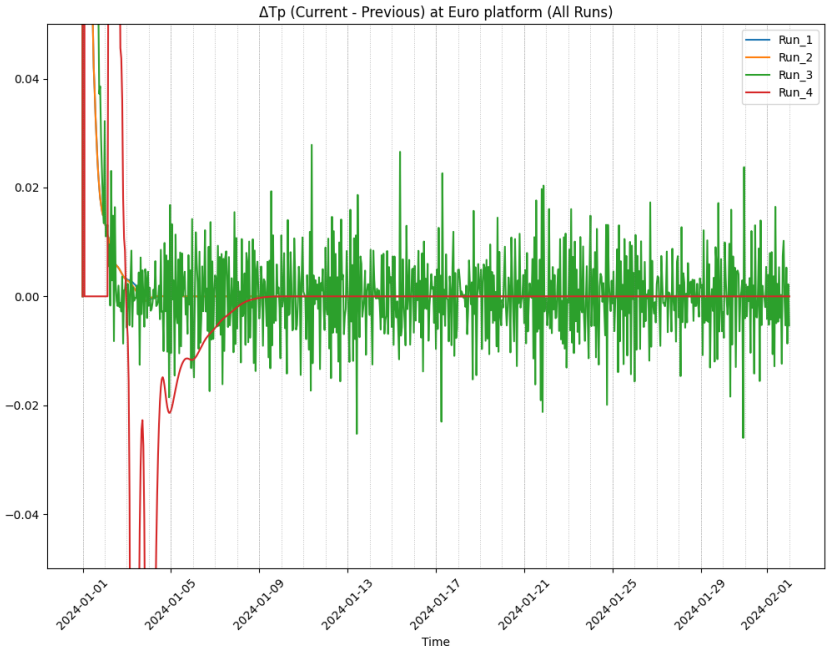

Figure 4.4: Hourly difference in peak wave period (T_p) at Euro Platform for the spin-up tests described in Table 4.1. The figure shows the model's temporal convergence behavior under different wind and boundary setups. Simulations with wind input stabilize within approximately 3–4 days, while the case driven only by boundary conditions (Run 4) requires around 13 days to reach equilibrium.

From figures 4.3 and 4.4, the clear limiting factor in determining spin-up time is found in run 4 which is the run conducted with no wind, and only boundary conditions. This run shows a slower decay in significant wave height difference, taking around 9 days to settle at 0 against 2 to 3 days for runs 1 and 2.

In terms of peak period difference, run 4 takes around 13 days to become stationary, whereas runs 1 and 2 take around 3 to 4 days. Run 3, the run conducted with no boundary conditions, and a larger wind velocity, shows constant oscillation of both H_{m0} and T_p around 0 after the initial drop meaning it is stationary after 3 days for H_{m0} and 4 days for T_p.

The observed oscillations and slower decay can be explained by the different mechanisms of wave energy input and redistribution under the applied forcing conditions. In the cases with wind forcing (runs 1–3), the local wind rapidly injects energy into the spectrum, promoting faster adjustment of both wave height and period toward equilibrium. This accelerates the spin-up process, as the wave field is continuously replenished and stabilized. By contrast, in run 4, where only boundary conditions are present, the model relies solely on wave propagation and dispersion from the boundaries into the domain. This process is inherently slower, since no additional local generation occurs, and energy gradients decay gradually through refraction, dissipation, and nonlinear interactions. The resulting memory effect of incoming swell leads to the longer adjustment time and the smoother, slower convergence observed. The oscillations in run 3 could be attributed to the dominance of local wind seas without boundary input, which could cause small fluctuations around equilibrium without significantly delaying stationarity.

These results show the largest spin-up time of 13 days occurs when simulating the peak period in a situation with no wind field. The most representative runs in this study are runs 2 and 3 which have a spin-up period of around one to two days. The decision to take 31 days, more than double this value is therefore a very safe estimate. The low computational time of an extra 2 weeks with the configuration shown in section 4.1 is small enough to not worry about optimization of this spin-up time. However, if one were to run a much longer hindcast, it could be useful to look into optimizing this value to save time and resources in the computation.

4.2.2. Sensitivity analysis at Euro Platform

A sensitivity analysis was conducted on several key input parameters at Euro Platform, with the goal of showing the sensitivity of the results to these fundamental input parameters. A total of 10 runs with a length of 1 month (31 days) were conducted for both July and December of 2023 arbitrarily. This is due to July being a relatively calm month and December being a month containing several storms including storm Pia between 19 and 23 December. The spin-up period of 31 days is used to start the model for each run before analyzing results. The buoy data used is cleaned as described in section 3.2. A table with all the runs conducted is shown in table 4.2.

Table 4.2: Overview of Parameter Changes per Run

Run #	Variable Changed	New Value	Unit	Default Value
0000	Default	-	-	-
0001	`dt`	10	seconds	300
0002	`dx`	0.25	degrees	0.05
0002	`dy`	0.25	degrees	0.03333333
0003	`freqmin`	0.01	Hz	0.04
0003	`freqmax`	1	Hz	0.5
0004	`nsigma`	24	-	12
0005	`ntheta`	72	-	36
0006	`vmax_zijlema`	30	m/s	50
0007	`winddrag`	wu	-	zijlema
0008	`fbed`	0.001	-	0.019
0009	`fbed`	0.1	-	0.019
0010	`gammajsp`	3	-	3.3

Table 4.2 shows an analysis conducted on both numerical and physical parameters. The runs are listed in the following points.

- The first run aims to reduce the timestep to see if the large timestep in the default run shows a large difference in simulating significant wave height, peak period, and mean direction.
- The second run aims to investigate how different a 0.25 degree square would reduce result accuracy.

- The third and fourth runs test the effect of a larger spectrum with the same number of bins (Less frequencies per bin), or the same spectrum with the a larger amount of bins (more frequencies per bin) respectively.
- The fifth run tests a larger amount of directional bins, meaning a theoretically larger precision in the modeled direction.
- Runs 6 and 7 test the effect of different wind drag formulation with the maximum wind speed for the Zijlema formulation reduced to 30m/s and the usage of the Wu formulation with no cap on the wind speed respectively. The expected outcome of these runs would be smaller waves when reducing wind, and larger waves induced by the Wu formulation (Wróbel-Niedźwiecka et al., 2019) as shown in section 2.2.
- Additionally, the effect of the bed friction is tested with these values decreased and increased significantly in runs 8 and 9 respectively. Run 8 is therefore expected to have larger waves as the dissipation is decreased, and run 9 is expected to have smaller waves as the dissipation is increased.
- Lastly, the gamma value of the JONSWAP spectra is lowered to simulate a condition with a more spread out spectral peak.

Sensitivity test for July 2023

The results of these runs tested using the input of July 2023 are shown for Euro Platform for significant wave height, peak period, and mean wave direction in figures 4.5, 4.6, and 4.7

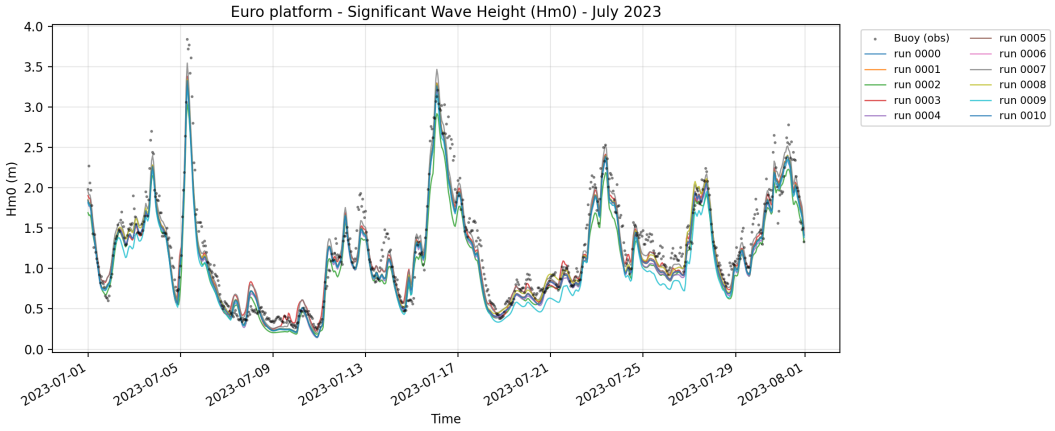

Figure 4.5: Sensitivity analysis of significant wave height (H_{m0}) at Euro Platform for July 2023, comparing 10 different input parameter configurations. Each line represents a separate model run (see Table 4.2). The figure shows that most parameter variations have limited influence on H_{m0}, with the coarser grid (Run 2) and altered friction (Runs 8–9) producing the largest deviations.

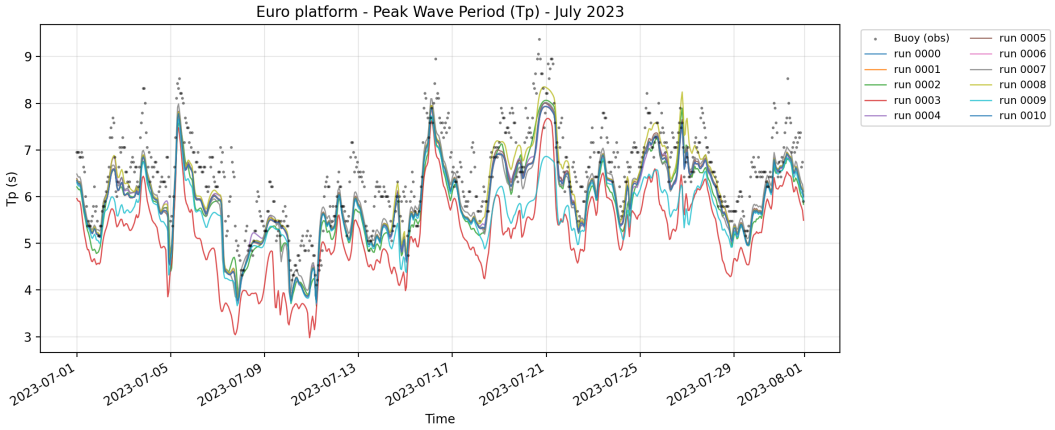

Figure 4.6: Time series comparison of modeled peak wave period (T_p) at Euro Platform under 10 parameter configurations for July 2023. Variations in spectral range (Run 3) produce the largest differences, indicating that the choice of frequency bounds significantly affects modeled wave periods. Other settings, such as timestep or directional bins, show smaller effects on T_p.

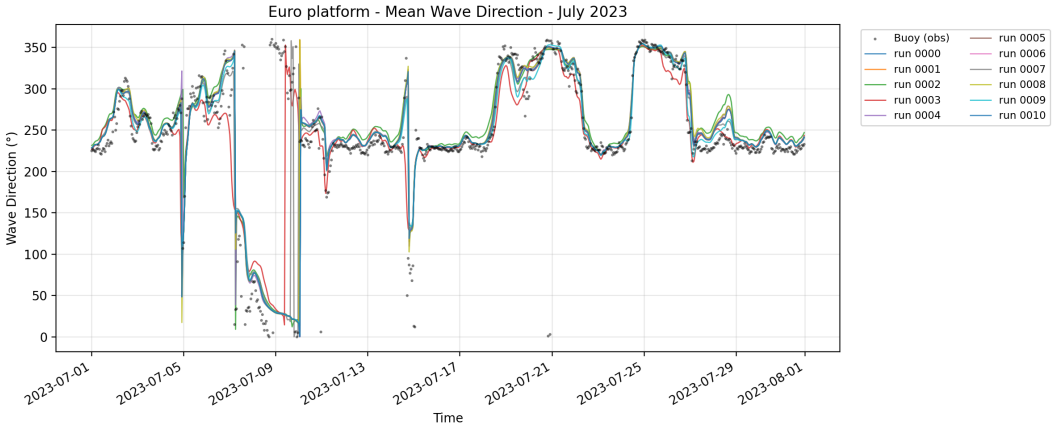

Figure 4.7: Comparison of modeled mean wave direction at Euro Platform for July 2023 across all sensitivity runs. The plot reveals limited directional deviations between configurations, confirming that spatial resolution and wind drag have a moderate influence, whereas frequency range changes (Run 3) have a more pronounced effect on modeled mean direction.

A table summarizing key statistics for each run is shown in table 4.3 for significant wave height (H_{m0}), peak period (T_p) and mean wave direction (θ_m).

Table 4.3: Comparison of RMSE values for Euro platform between sensitivity runs in July 2023

Run	RMSE of H_{m0} [m]	RMSE of T_p [s]	RMSE of θ_m [°]
0000	0.212	0.726	26.82
0001	0.214	0.727	26.93
0002	0.264	0.777	29.65
0003	0.199	1.381	30.23
0004	0.228	0.716	26.73
0005	0.208	0.705	27.04
0006	0.212	0.726	26.82
0007	0.181	0.692	26.72
0008	0.204	0.661	27.05
0009	0.252	1.002	27.30
0010	0.212	0.726	26.82

Looking at figures 4.5, 4.6, 4.7, and table 4.3, each run tends to stay around the same values with a few exceptions. The peak period and mean wave direction are the wave parameters which are the most affected by these changes. Most notably, run 3 with more frequencies shows large differences in wave height period and direction. A potential cause for this could be the inclusion of longer waves which move the peak period to lower values as shown in figure 4.6. Another difference to the expected outcome is the change of timestep in run 1, with a reduced timestep showing the same or even lower accuracy than the default timestep. The larger amount of directional bins in run 5 also do not help against error in modeling wave direction. The other runs behave as expected, with a coarser grid being less accurate in run 2, friction playing the expected role of increasing and decreasing wave height and period for runs 8 and 9, and wind drag formulation behaving as expected. It should be noted that the Wu formulation which gives larger waves is more accurate against the buoy measurements as the model tends to under predict wave heights and periods.

Sensitivity test for December 2023

After looking at the results for July 2023, the results of these runs tested using the input of December 2023 are shown for Euro Platform for significant wave height, peak period, and mean wave direction in figures 4.8, 4.9, and 4.10

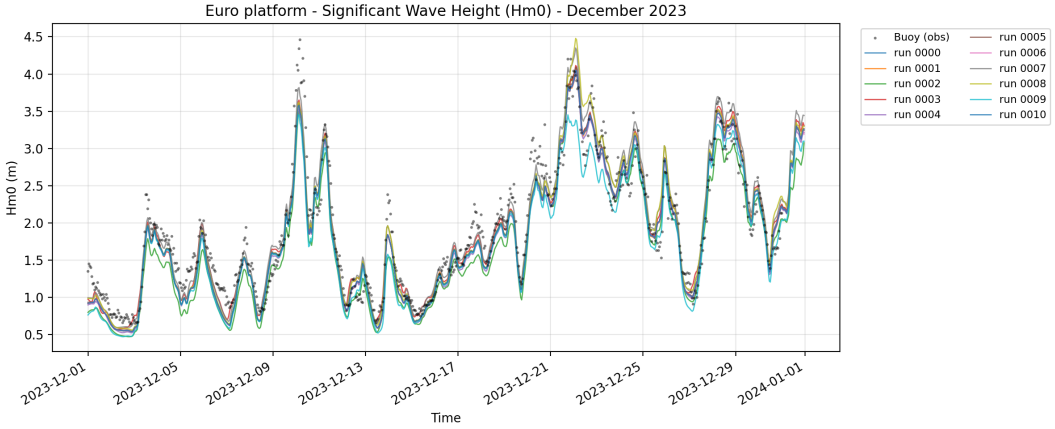

Figure 4.8: Modeled significant wave height (H_{m0}) at Euro Platform during December 2023 for all sensitivity runs. The figure highlights the model's robustness under energetic winter conditions, showing that coarse grid resolution (Run 2) and reduced bed friction (Run 9) lead to noticeable deviations, while most other parameter changes cause minor variations.

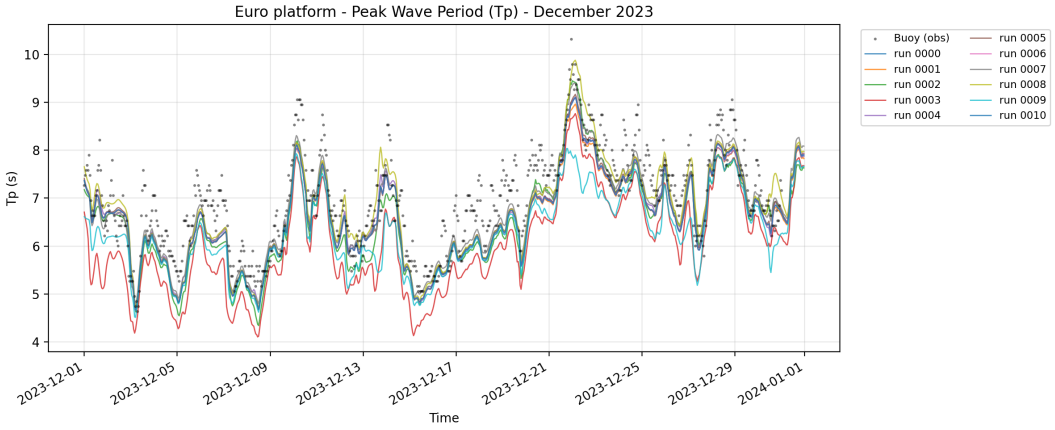

Figure 4.9: Modeled peak period (T_p) at Euro Platform for December 2023 under different parameter configurations. The broader frequency range (Run 3) along with the reduced bed friction (run 9) show a decrease in peak period

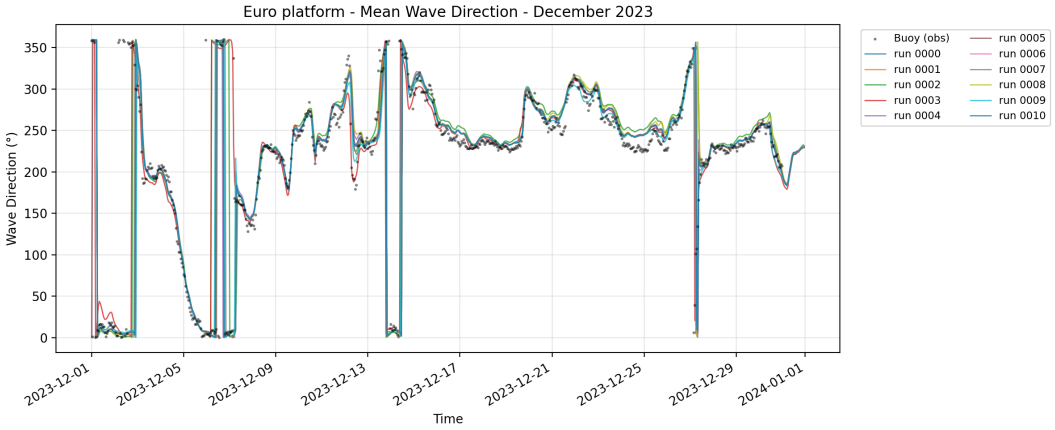

Figure 4.10: Modeled mean wave direction at Euro Platform in December 2023 for 10 sensitivity tests. Results show similar behavior across runs, with small directional shifts between parameter sets.

A table summarizing key statistics for each run is shown in table 4.4.

Table 4.4: Comparison of RMSE values for Euro platform between sensitivity runs in December 2023

Run	RMSE of H_{m0} [m]	RMSE of T_p [s]	RMSE of Mean wave direction [°]
0000	0.252	0.582	17.56
0001	0.253	0.595	17.76
0002	0.314	0.645	19.73
0003	0.243	1.029	12.48
0004	0.259	0.573	18.40
0005	0.250	0.571	17.69
0006	0.252	0.583	17.56
0007	0.245	0.506	15.33
0008	0.262	0.543	19.13
0009	0.310	0.843	14.67
0010	0.252	0.583	17.56

The results shown in these figures for December show the same general ideas as July, with some exceptions. Run 3, with the larger frequency range shows a better model of the angles instead of worse as was in July. Once more the largest differences occur when looking at the peak period.

This section shows that model input remains important with most parameters having a more or less large effect on the model outcome. The spin-up period of 31 days is appropriate here with a large safety factor. These simulations also show some numerical instabilities of around 0.01 m for significant wave height and 0.02 s for peak period when large winds affect the whole area. These values are very small, however they must be taken into account when evaluating the precision of the model.

The most important parameter according to this analysis is frequency range, determining which frequencies are included or not in the model can change the results significantly. The frequency range applied in the report by van der Lugt et al. (2024) is used and kept between 0.04 and 0.5 Hz as lower frequency waves would be considered infragravity waves (Rijnsdorp et al., 2021) and are not included in this study as shown in section 1.4. The next most important parameter is bed friction with it playing a role in dissipation. Lastly, spatial resolution is more important than time resolution, as a coarser grid shows less accurate modeling of wave movement and energy transfer.

4.3. Model performance in reanalysis

A reanalysis was conducted analyzing the performance of the model between 1985 and 2023. The analysis uses the methodology described in section 3.1 to obtain the meteo files needed to run the model.

The grid and bathymetry are initialized using the method described in section 3.4. The model uses the input described in table 3.7 in section 3.5.

4.3.1. Analysis on a single year

In order to understand the output of the model, a single year, in this case, 2023 was modeled and analyzed for the integrated wave parameters of significant wave height H_{m0}, peak period T_p and wave direction. These parameters enable to check whether the model works properly for a single year. The year 2023 is a good benchmark as it includes relatively calm periods in July as well as stormy periods in December. For this year, the model was run and compared to the data available from both Rijkswaterstaat buoys and ERA5 reanalysis obtained using the methods described in section 3.1.

Comparison of significant wave height and peak period

Figures 4.11 and 4.12 shows the result of the 2023 HurryWave model run at the Euro Platform station against measured buoy data. The ERA5 data obtained at the location of the measuring station, is also plotted alongside.

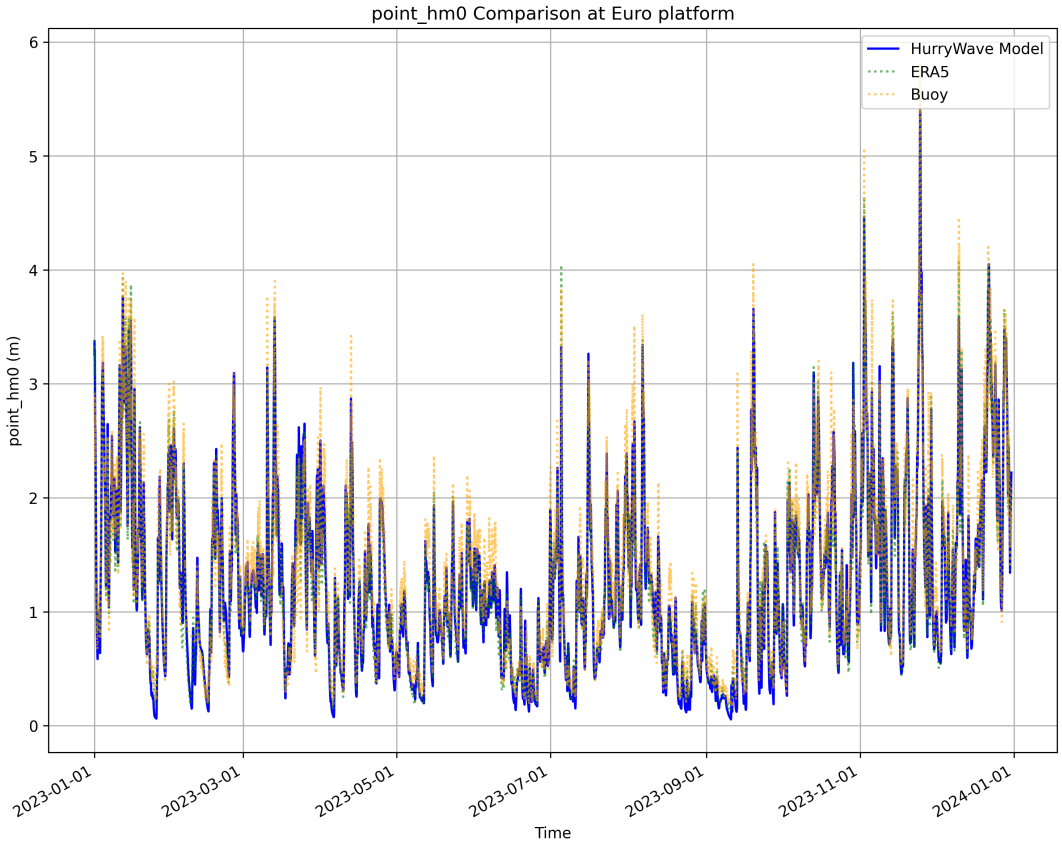

Figure 4.11: Time series comparison of significant wave height (H_{m0}) at Euro Platform for 2023, showing HurryWave model output, ERA5 reanalysis, and buoy observations. The HurryWave model accurately reproduces observed variability, with slight underestimation of extreme values.

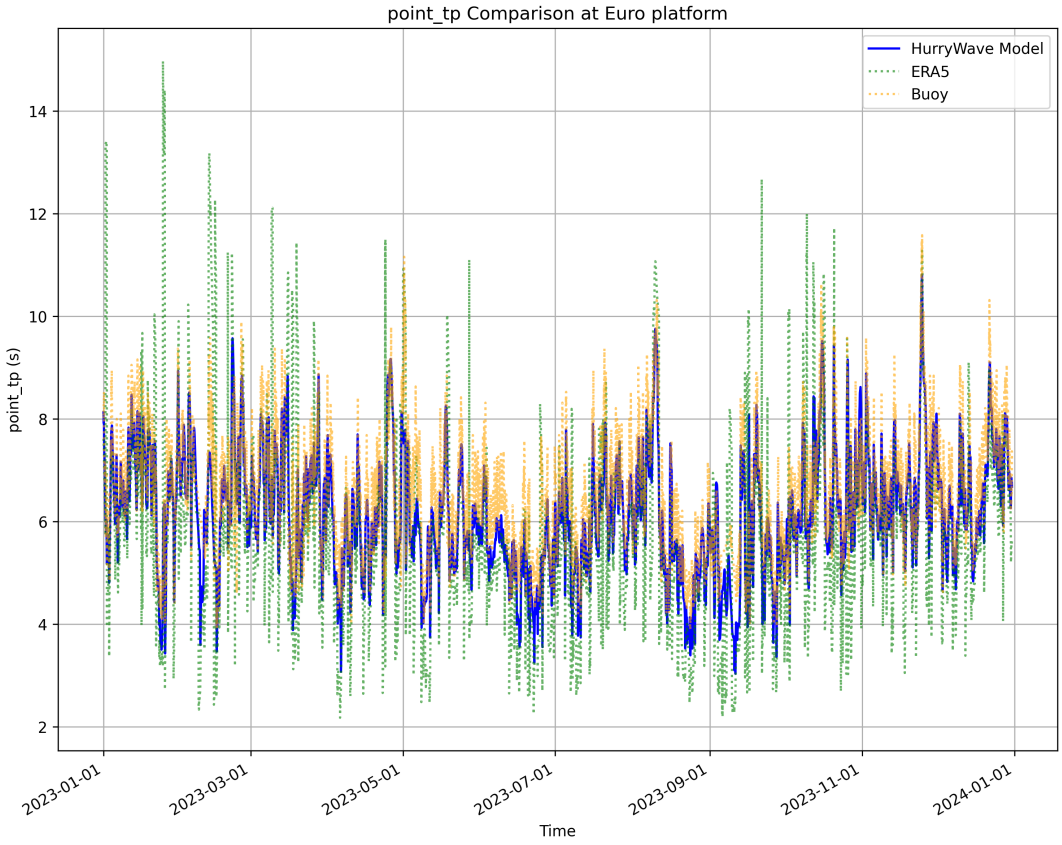

Figure 4.12: Time series comparison of peak period (T_p) at Euro Platform for 2023, showing HurryWave model output, ERA5 reanalysis, and buoy observations. The HurryWave model generally underestimates peak period. ERA5 shows high variability in peak period

Looking at the significant wave height in figure 4.11, the model output follows the buoy data and the output of the WAM model from ERA5, although the HurryWave model underestimates the largest waves.

Looking at the peak period in figure 4.12, the HurryWave model output qualitatively lines up better with the peak period of the buoy data than ERA5. The output of the WAM model used by ERA5 shows large variability in modeling peak period.

These trends are also reflected in the scatter plot comparing the HurryWave model and the WAM model used in ERA5 to buoy observations in 2023 shown in figure 4.13.

A sorted plot of both the significant wave height and peak period is shown in figure 4.13, where values are sorted and compared against ERA5 data, or buoy values. This plot shows grid lines equal to the 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, and 0.95 quantiles in order.

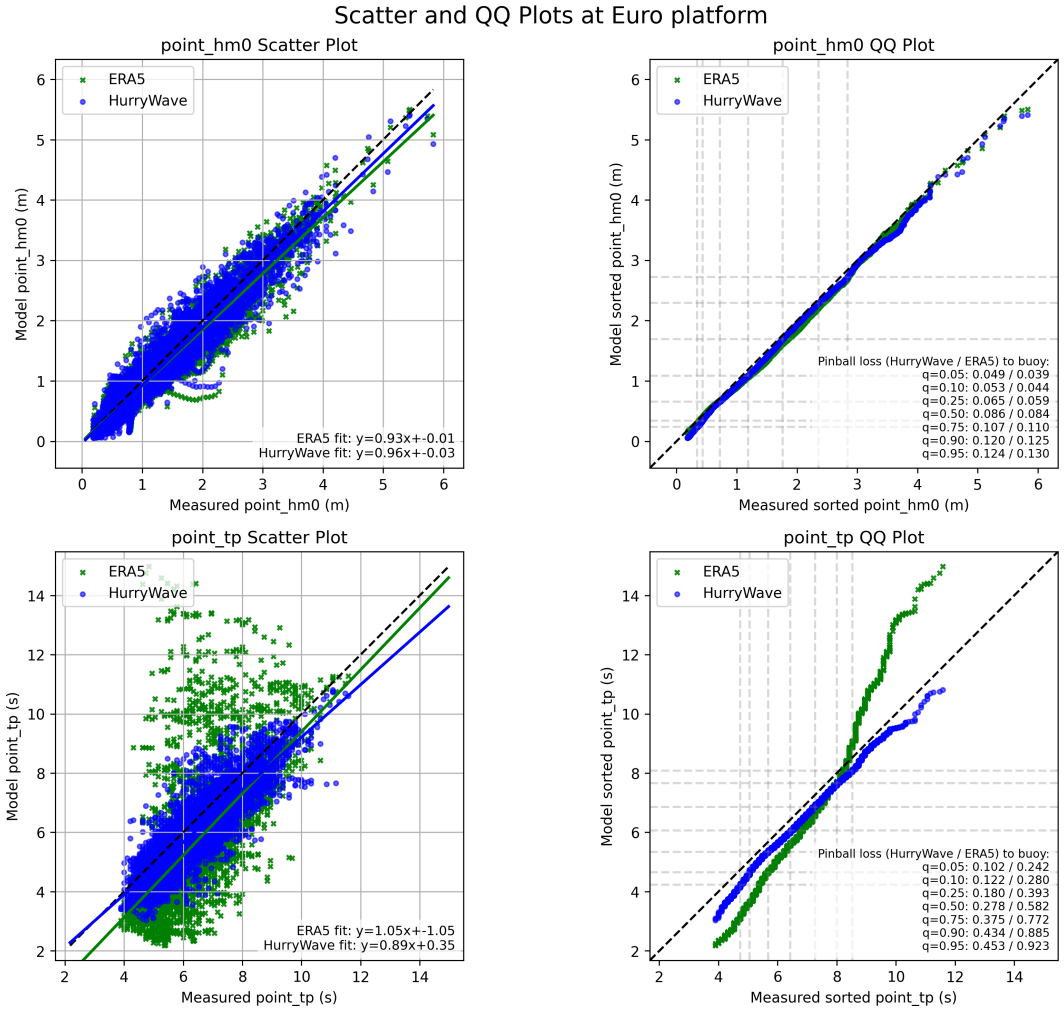

Figure 4.13: Scatter plot (left) and quantile–quantile plot (right) comparing HurryWave and ERA5 model results to buoy observations of H_{m0} and T_p at Euro Platform in 2023. The scatter plot reveals the underestimation of large wave heights by HurryWave and the general underestimation of the peak period. It also shows the large variations in ERA5 peak period

This graph in figure 4.13 shows that the significant wave height is well modeled even in the upper quantiles. The peak period shows some deviations from the measured buoy data, however, the output data of WAM in ERA5 deviates significantly more, especially in the upper quantiles. This is shown through the pinball loss function using the 0.95 quantile where it is 0.453 using HurryWave and 0.923 using WAM. WAM also underestimates lower quantiles and overestimates upper quantiles, whereas HurryWave shows an underestimation. These are the quantiles where the model needs to be reliable as they are those where extremes occur. These trends continue at all stations where measurements are available in 2023, including IJmuiden munitiestortplaats, and K13 Alpha. These places are all, as shown in figure 3.2, stations near the Dutch coast. A table showing the statistics at all locations where wave measurements are available in 2023 is shown for H_{m0} using both the HurryWave model, and the WAM model used in ERA5 in table 4.5, and for T_p in table 4.6.

Table 4.5: Performance metrics for significant wave height H_{m0} in 2023. Comparison between HurryWave and ERA5.

Station	Metric	HurryWave	ERA5
Euro Platform	RMSE (m)	0.225	0.230
	SI (-)	0.171	0.174
	Bias (m)	-0.084	-0.101
	Pinball loss $q = 0.90$ (m)	0.120	0.125
	Pinball loss $q = 0.95$ (m)	0.124	0.130
K13 Alpha	RMSE (m)	0.255	0.250
	SI (-)	0.169	0.165
	Bias (m)	-0.132	-0.109
	Pinball loss $q = 0.90$ (m)	0.153	0.135
	Pinball loss $q = 0.95$ (m)	0.159	0.140
IJmuiden Munitiestortplaats	RMSE (m)	0.226	0.262
	SI (-)	0.163	0.188
	Bias (m)	-0.094	-0.139
	Pinball loss $q = 0.90$ (m)	0.124	0.151
	Pinball loss $q = 0.95$ (m)	0.128	0.158
North Cormorant	RMSE (m)	0.979	0.314
	SI (-)	0.379	0.122
	Bias (m)	-0.767	0.010
	Pinball loss $q = 0.90$ (m)	0.696	0.110
	Pinball loss $q = 0.95$ (m)	0.735	0.110

Table 4.6: Performance metrics for peak wave period T_p in 2023. Comparison between HurryWave and ERA5.

Station	Metric	HurryWave	ERA5
Euro Platform	RMSE (s)	0.702	1.487
	SI (-)	0.108	0.228
	Bias (s)	-0.390	-0.757
	Pinball loss $q = 0.90$ (s)	0.434	0.885
	Pinball loss $q = 0.95$ (s)	0.453	0.923
K13 Alpha	RMSE (s)	1.048	1.477
	SI (-)	0.147	0.207
	Bias (s)	-0.736	-0.907
	Pinball loss $q = 0.90$ (s)	0.689	0.959
	Pinball loss $q = 0.95$ (s)	0.726	1.005
IJmuiden Munitiestortplaats	RMSE (s)	0.911	1.431
	SI (-)	0.133	0.209
	Bias (s)	-0.528	-0.766
	Pinball loss $q = 0.90$ (s)	0.543	0.861
	Pinball loss $q = 0.95$ (s)	0.570	0.900

Tables 4.5 and 4.6 show HurryWave performs better than WAM quantitatively in 2023 for all stations except North Cormorant. The North Cormorant station is found at a high latitude and is situated in deeper water than all other stations. The statistics confirm the good fit on this year and this station as the scatter index comparing the HurryWave model significant wave height to that measured by the buoy is 0.171, whereas the WAM model used in ERA5 shows a scatter index of 0.174. The low scatter index in the HurryWave model is enough to consider the HurryWave model acceptable for modeling significant wave height in this year (Ris et al., 1999). In terms of bias, the significant wave height is underestimated by HurryWave with a bias value of -0.084m. This bias value is also lower than the one seen in the WAM model, demonstrating the general underestimation of the significant wave height in 2023 is lower for HurryWave than for WAM.

The peak period is also quantitavely better modeled in HurryWave than in ERA5 with a scatter index of 0.108 comparing the buoy data to the HurryWave model and a scatter index of 0.228 when comparing to

ERA5. The HurryWave model seems to underestimate the peak wave period when comparing to the buoy data with a bias value of -0.390s.

To better understand the large scatter index at North Cormorant, the timeseries of the significant wave height is plotted for this station in figure 4.14.

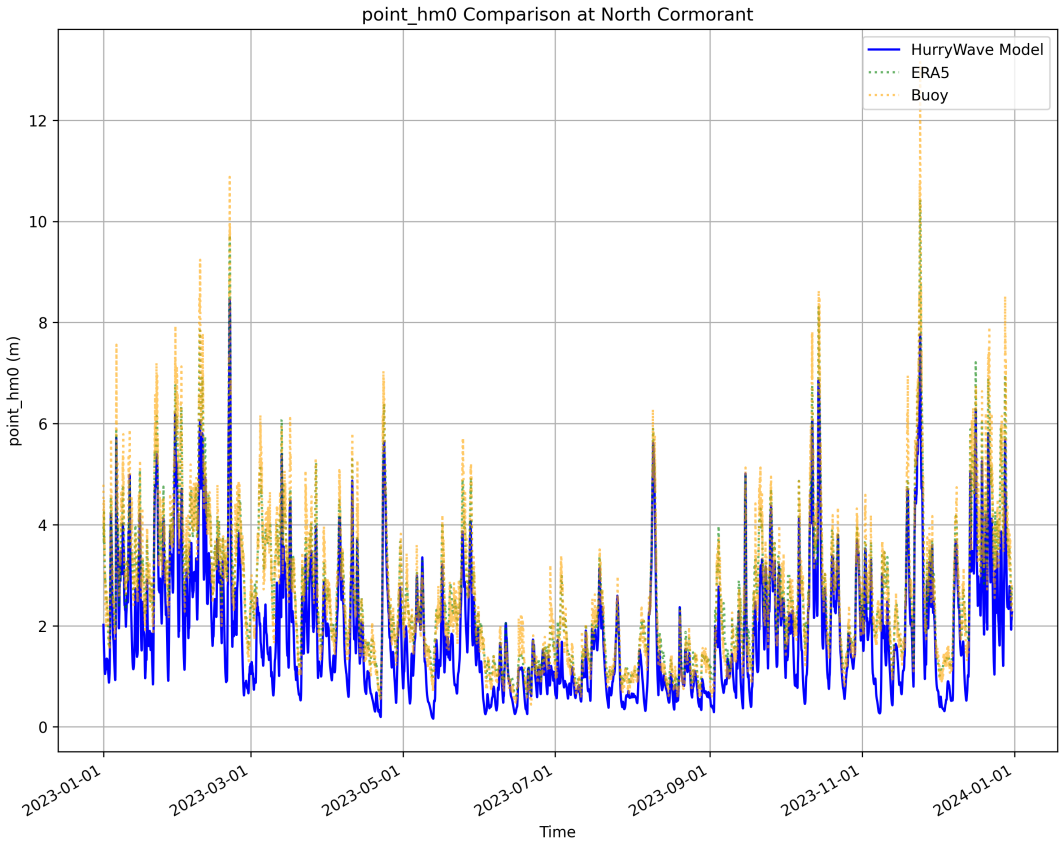

Figure 4.14: Time series of significant wave height (H_{m0}) at North Cormorant for 2023, comparing HurryWave and ERA5 results to buoy data. The plot shows underestimation of large waves by HurryWave in deep water, consistent with higher scatter and bias values, reflecting challenges in deeper waters using this configuration.

Looking at the significant wave height, there is a large uncertainty when comparing the model to the buoy measurement with a scatter index of 0.377. The main issue seems to be underestimation of the peaks as suggested by the bias value of -0.763m.

The peak period does not contain measurements for this location meaning no comparison was made for peak period at this location.

A scatter and a sorted plot of the significant wave height are shown in figure 4.15. This plot shows grid lines equal to the 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, and 0.95 quantiles in order.

This quantile plot in figure 4.15 shows the underestimation increases with the increase of the significant wave height. These results show that HurryWave performs better near the Dutch coast than in the Northern part of the North Sea where the waves contain more energy and water depth is larger. The other stations found in figure 3.2 were also modeled during this period and the results can be found in Appendix D.

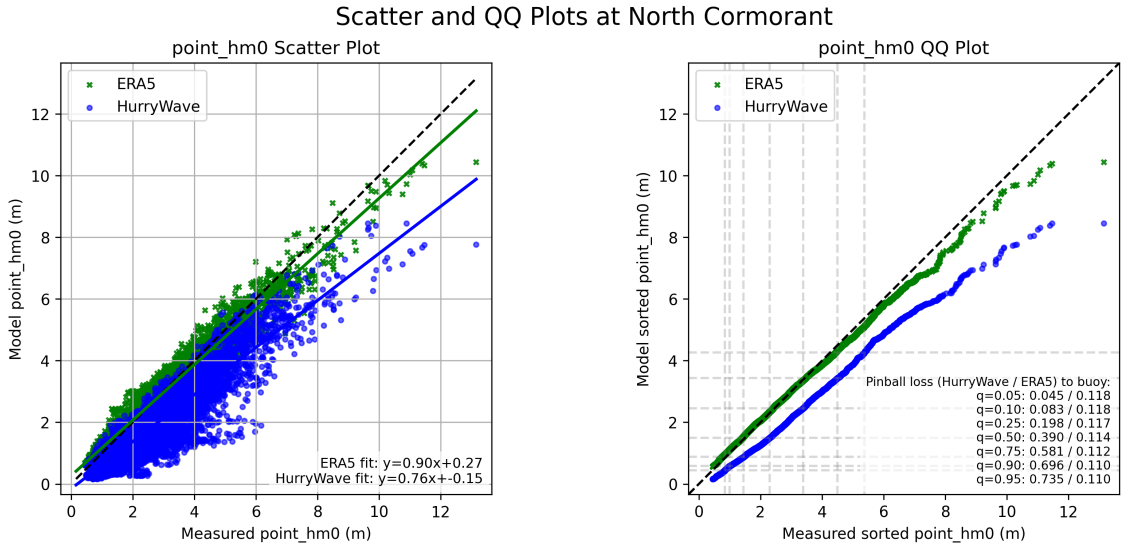

Figure 4.15: Scatter (left) and quantile–quantile (right) plots comparing HurryWave and ERA5 model outputs against buoydata at North Cormorant for 2023. The systematic underestimation at higher quantiles emphasizes the reduced model accuracy in deep-water regions using this setup compared to nearshore sites.

Analysis of wave direction

To learn more about the performance of HurryWave in plotting wave direction, the mean wave direction was analyzed comparing both the HurryWave model and the ERA5 model to measured buoy data. This comparison was performed through the use of several metrics such as angular error mean (or bias), angular error standard deviation, and angular RMSE. These were also normalized to allow comparison to other papers as shown in subsection 3.8.1. A table with these statistics is shown in table 4.7.

Table 4.7: Performance metrics for mean wave direction in 2023. Comparison between HurryWave and ERA5.

Station	Metric	HurryWave	ERA5
Euro Platform	Mean error (°)	4.77	-0.70
	Normalized mean error (%)	1.32	-0.19
	Std error (°)	24.31	22.98
	Normalized std error (%)	6.75	6.38
	RMSE error (°)	24.77	22.99
	Normalized RMSE error (%)	6.88	6.39
K13 Alpha	Mean error (°)	4.30	-0.49
	Normalized mean error (%)	1.20	-0.14
	Std error (°)	24.16	21.36
	Normalized std error (%)	6.71	5.93
	RMSE error (°)	24.54	21.37
	Normalized RMSE error (%)	6.82	5.94
IJmuiden Munitiestortplaats	Mean error (°)	4.02	-2.30
	Normalized mean error (%)	1.12	-0.64
	Std error (°)	23.88	23.51
	Normalized std error (%)	6.63	6.53
	RMSE error (°)	24.21	23.63
	Normalized RMSE error (%)	6.73	6.56

Table 4.7 shows HurryWave seems to have a larger mean error compared to ERA5, when both are compared to buoy data. The standard deviation on the error of both outputs is also high showing large variations. Most of the error is concentrated around 0 degrees for both HurryWave and ERA5 showing the models tend to accurately predict wave directions, with some outliers being present. A scatter plot

showing the wave direction error is shown in figure 4.16, and a histogram of the wave direction error is shown in figure 4.17. The other stations found in figure 3.2 were also modeled during this period and the results can be found in Appendix D.

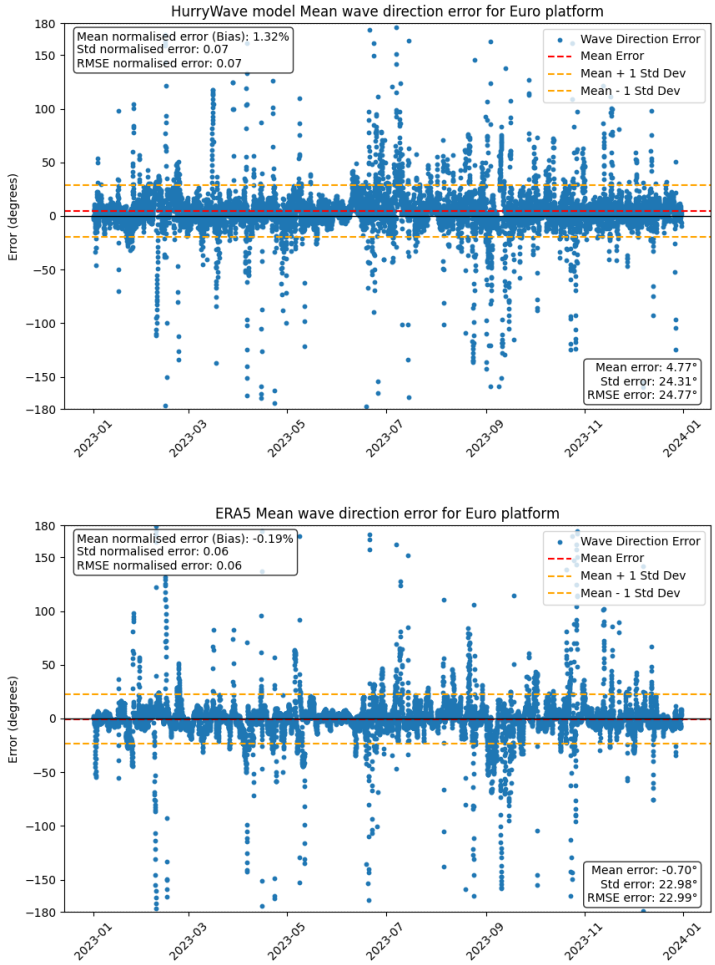

Figure 4.16: Scatter plots of mean wave direction errors at Euro Platform for 2023. The top panel shows HurryWave model results and the bottom panel ERA5. Each point represents the deviation from buoy observations.

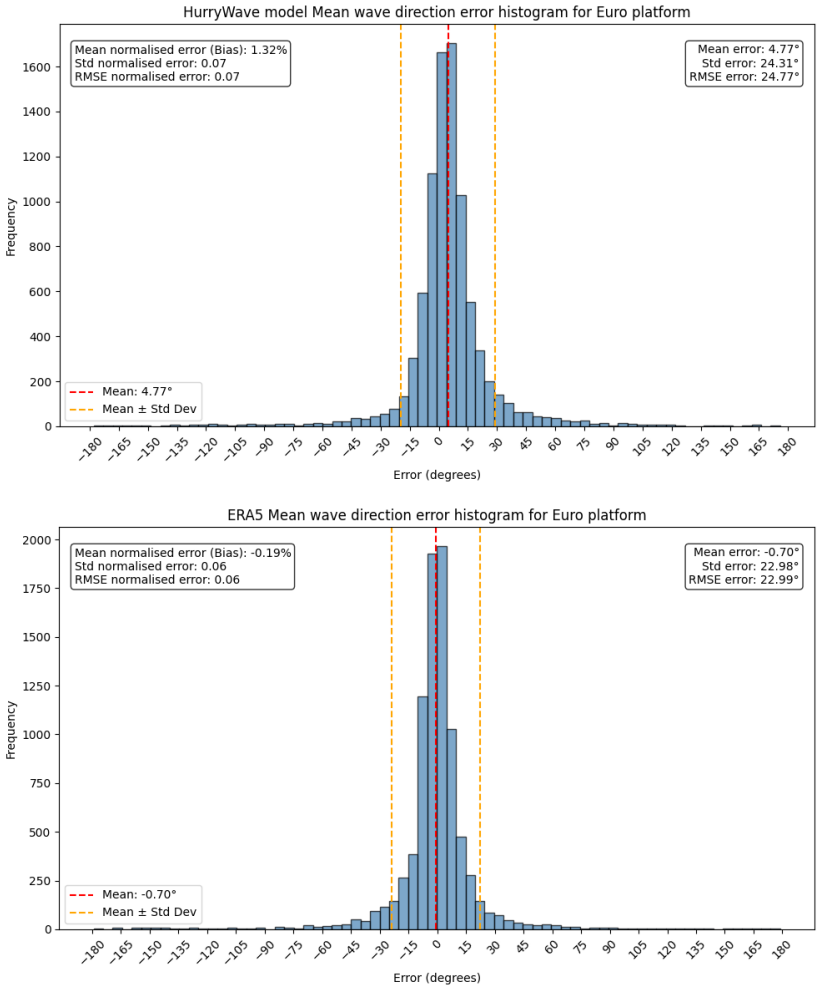

Figure 4.17: Histograms of mean wave direction errors for HurryWave (top) and ERA5 (bottom) at Euro Platform during 2023. Both distributions are centered near zero with moderate dispersion (±25°), showing comparable accuracy.

4.3.2. Study of a single storm

After analyzing a full year, a zoom in on a single storm showcases the short term performance of HurryWave in a high energy scenario. The short term focuses on the so called Sinterklaas storm (or storm Xaver) from 5 to 6 December 2013. To focus on this, the first week of December 2013 is plotted.

The significant wave height and peak period at Euro platform during this storm are plotted in figures 4.18 and 4.19 respectively.

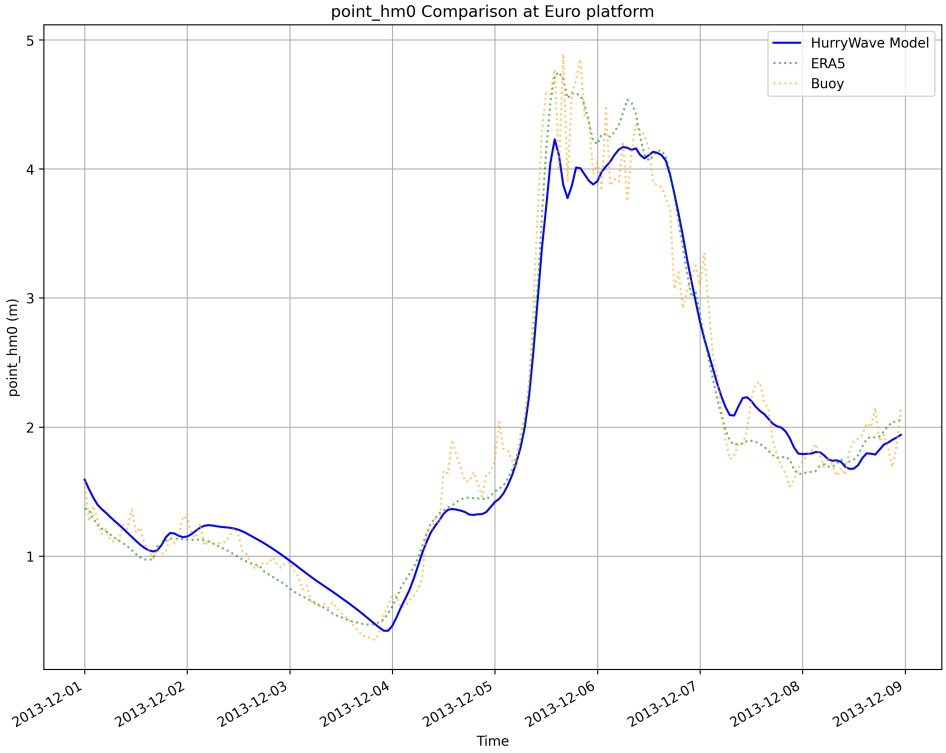

Figure 4.18: Time series of significant wave height (H_{m0}) at Euro Platform from 1–8 December 2013, capturing the Sinterklaas (Xaver) storm event. This plot highlights the underestimation of significant wave height in modeling large waves

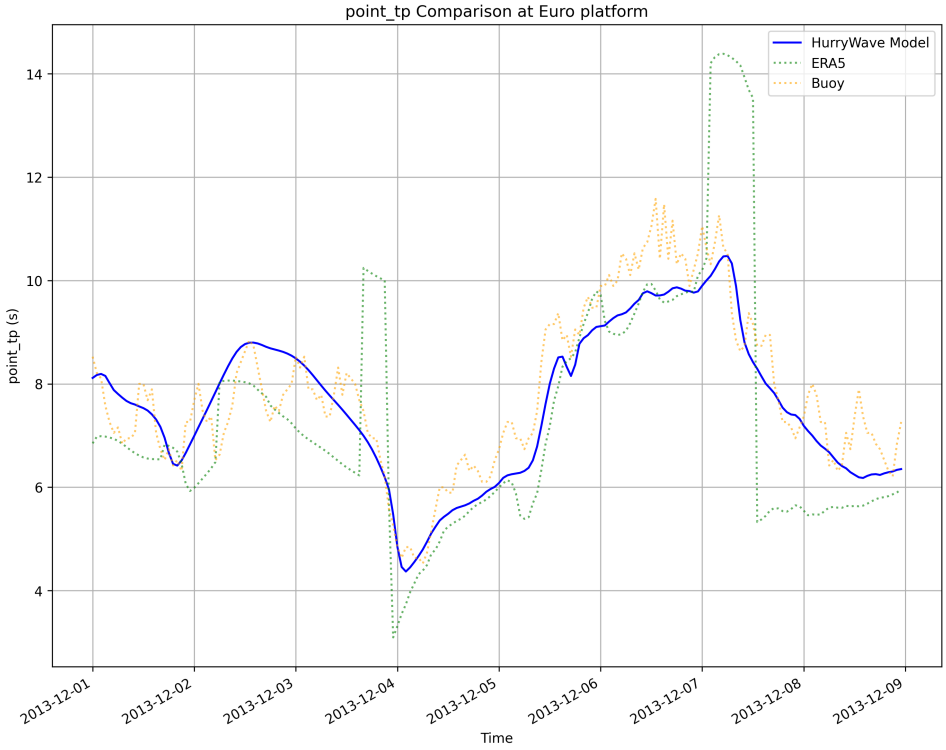

Figure 4.19: Time series of peak period (T_p) at Euro Platform from 1–8 December 2013, capturing the Sinterklaas (Xaver) storm event. This plot highlights the general underestimation of peak period in modeling large waves

This storm shows large variation between stations. At stations D15A, K13 alpha, F16-A, and Hoorn Q1-A, both the significant wave height and peak wave period are underestimated during the storm as shown in the significant wave height time series in figure 4.20 and the peak period time series in figure 4.21. The WAM model used in ERA5 does not show this behavior only slightly underestimating the wave period. The scatter plot along with the sorted quantile plot in figure 4.22 at platform D15-A confirm this observation.

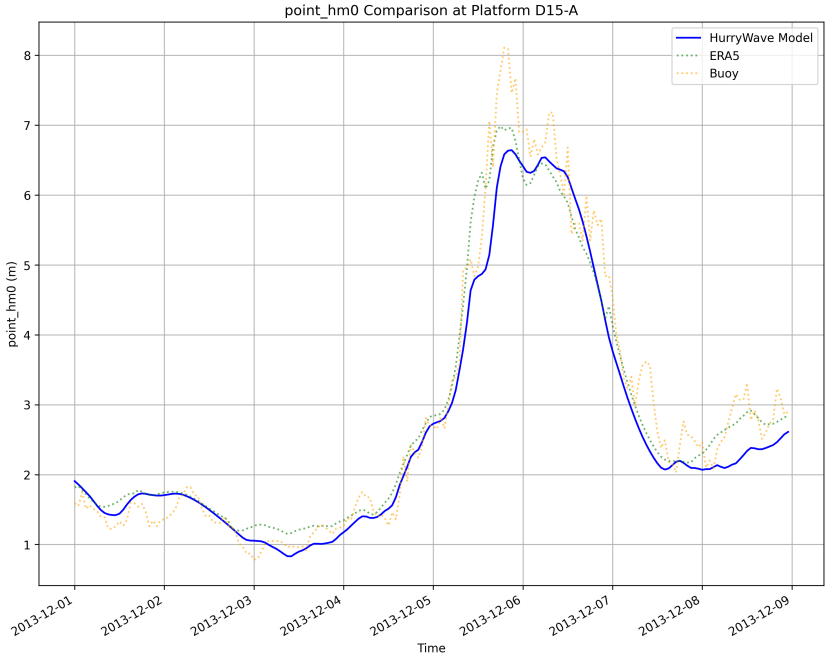

Figure 4.20: Comparison of HurryWave, ERA5, and buoy data for significant wave height at platform D15-A during the Sinterklaas storm. The figure shows the underestimation of both significant wave height and peak period by HurryWave during a storm

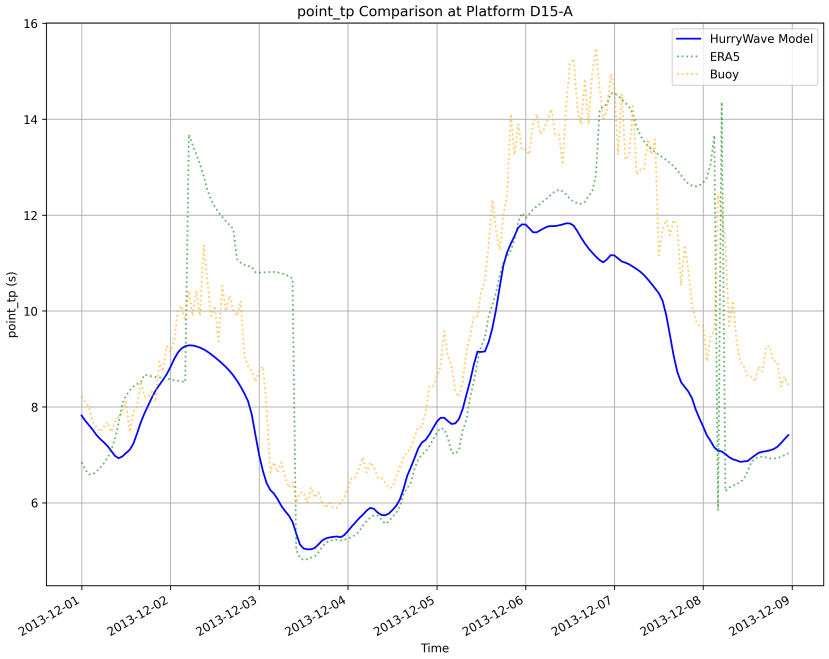

Figure 4.21: Comparison of HurryWave, ERA5, and buoy data for peak wave period at platform D15-A during the Sinterklaas storm. The figure shows that HurryWave underestimates T_p during peak storm energy

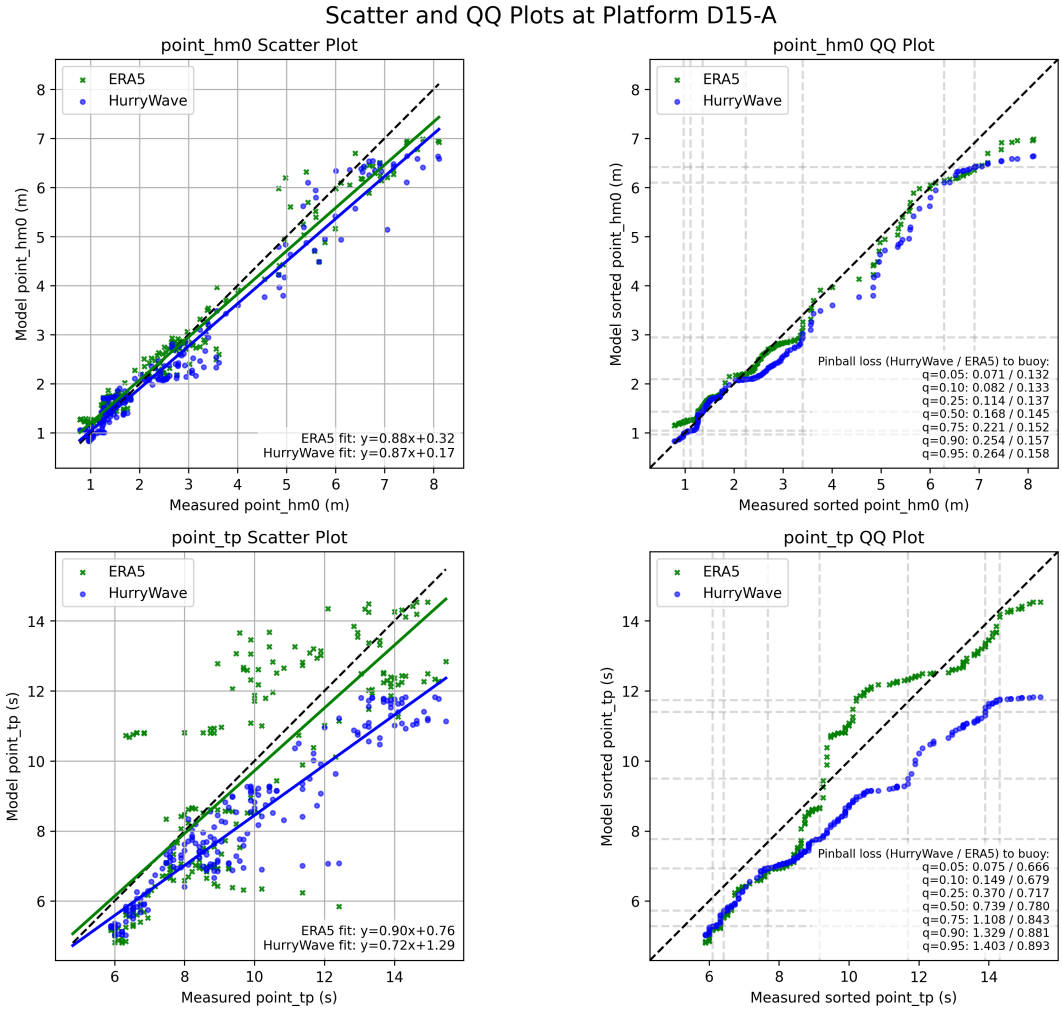

Figure 4.22: Scatter (left) and quantile–quantile (right) plots comparing HurryWave and ERA5 results to buoy data at D15-A during the Sinterklaas storm. The plots highlight the model's underestimation of long-period waves and high-energy conditions compared to observations.

At the North Cormorant station, the HurryWave model once again underestimates wave heights as shown in previous paragraphs. This is shown in the scatter plot and the sorted quantile plot for the significant wave height in figure 4.23. These figures show underestimations even at the lower quantiles.

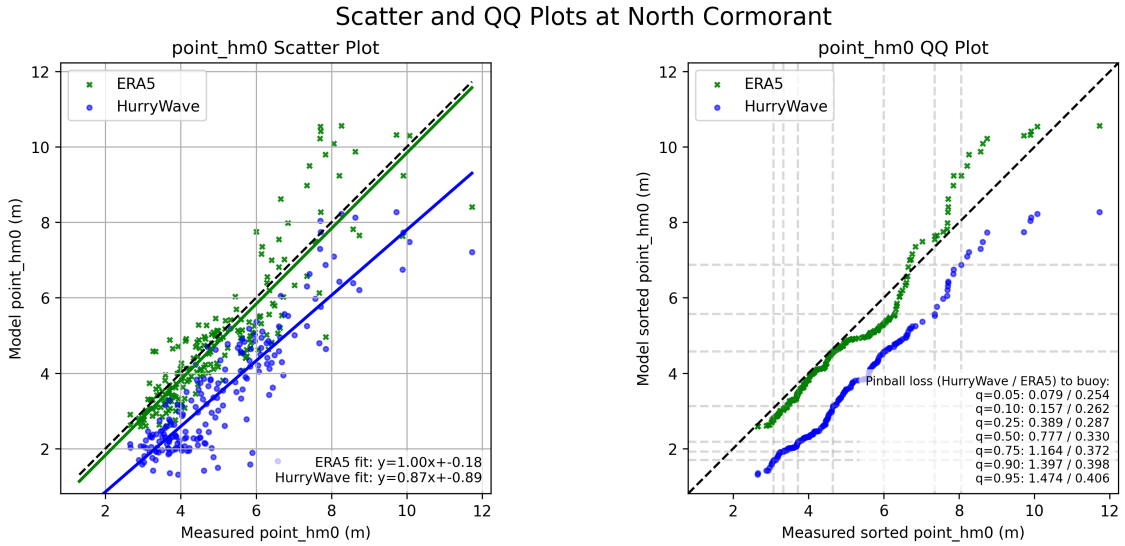

Figure 4.23: Scatter and quantile plots of significant wave height (H_{m0}) at North Cormorant for 1–8 December 2013. The figure shows underestimation across all quantiles by HurryWave relative to buoy data, reflecting model limitations in deep-water, high-energy environments

A table recording the performance metrics over this one week simulation is shown for significant wave height in table 4.8, and for peak period in table 4.9. These tables reflect the qualitative assessment made above. The model quality simply looking at scatter index always remains acceptable, below 0.4 for all stations even North Cormorant. However, the bias is large, especially for peak period where the differences between HurryWave and WAM are shown. This showcases the difficulties of modeling high energy scenarios using HurryWave in this configuration. The other stations found in figure 3.2 were also modeled during this storm and the results can be found in Appendix E.

Table 4.8: Performance metrics for significant wave height H_{m0} between 1 and 8 December 2013. Comparison between HurryWave and ERA5.

Station	Metric	HurryWave	ERA5
Euro Platform	RMSE (m)	0.267	0.223
	Scatter Index (-)	0.140	0.117
	Bias (m)	-0.025	-0.028
	Pinball loss ($p = 0.90$) (m)	0.108	0.091
	Pinball loss ($p = 0.95$) (m)	0.109	0.092
K13 Alpha	RMSE (m)	0.482	0.534
	Scatter Index (-)	0.187	0.207
	Bias (m)	-0.289	-0.357
	Pinball loss ($p = 0.90$) (m)	0.270	0.331
	Pinball loss ($p = 0.95$) (m)	0.284	0.348
IJmuiden Munitiestortplaats	RMSE (m)	0.390	0.451
	Scatter Index (-)	0.168	0.194
	Bias (m)	-0.104	-0.275
	Pinball loss ($p = 0.90$) (m)	0.178	0.267
	Pinball loss ($p = 0.95$) (m)	0.183	0.280
North Cormorant	RMSE (m)	1.748	0.898
	Scatter Index (-)	0.346	0.178
	Bias (m)	-1.550	-0.169
	Pinball loss ($p = 0.90$) (m)	1.397	0.398
	Pinball loss ($p = 0.95$) (m)	1.475	0.406
D15-A	RMSE (m)	0.478	0.394
	Scatter Index (-)	0.168	0.138
	Bias (m)	-0.215	-0.030
	Pinball loss ($p = 0.90$) (m)	0.254	0.157
	Pinball loss ($p = 0.95$) (m)	0.264	0.158
F16-A	RMSE (m)	0.545	0.473
	Scatter Index (-)	0.172	0.149
	Bias (m)	-0.366	-0.252
	Pinball loss ($p = 0.90$) (m)	0.342	0.262
	Pinball loss ($p = 0.95$) (m)	0.360	0.275
Hoorn Q1-A	RMSE (m)	0.488	0.571
	Scatter Index (-)	0.183	0.215
	Bias (m)	-0.292	-0.329
	Pinball loss ($p = 0.90$) (m)	0.273	0.307
	Pinball loss ($p = 0.95$) (m)	0.288	0.323

Table 4.9: Performance metrics for peak period T_p between 1 and 8 December 2013. Comparison between HurryWave and ERA5.

Station	Metric	HurryWave	ERA5
Euro Platform	RMSE (s)	0.686	1.687
	Scatter Index (-)	0.087	0.215
	Bias (s)	-0.253	-0.516
	Pinball loss ($p = 0.90$) (s)	0.387	0.847
	Pinball loss ($p = 0.95$) (s)	0.400	0.873
K13 Alpha	RMSE (s)	1.337	1.950
	Scatter Index (-)	0.148	0.216
	Bias (s)	-1.058	-0.764
	Pinball loss ($p = 0.90$) (s)	0.958	1.136
	Pinball loss ($p = 0.95$) (s)	1.011	1.174
IJmuiden Munitiestortplaats	RMSE (s)	1.219	1.736
	Scatter Index (-)	0.136	0.193
	Bias (s)	-0.876	-0.767
	Pinball loss ($p = 0.90$) (s)	0.834	1.063
	Pinball loss ($p = 0.95$) (s)	0.878	1.101
D15-A	RMSE (s)	1.790	1.893
	Scatter Index (-)	0.184	0.194
	Bias (s)	-1.476	-0.252
	Pinball loss ($p = 0.90$) (s)	1.329	0.881
	Pinball loss ($p = 0.95$) (s)	1.403	0.893
F16-A	RMSE (s)	1.739	1.832
	Scatter Index (-)	0.177	0.186
	Bias (s)	-1.534	-0.422
	Pinball loss ($p = 0.90$) (s)	1.380	0.943
	Pinball loss ($p = 0.95$) (s)	1.457	0.964
Hoorn Q1-A	RMSE (s)	1.369	1.891
	Scatter Index (-)	0.149	0.206
	Bias (s)	-1.067	-0.578
	Pinball loss ($p = 0.90$) (s)	0.973	1.055
	Pinball loss ($p = 0.95$) (s)	1.027	1.083

The analysis of the wave direction was also performed on this 1 week period. The wave direction over this period is plotted at IJmuiden munitiestortplaats in figure 4.24. The other stations found in figure 3.2 were also modeled during this storm and the results can be found in Appendix E.

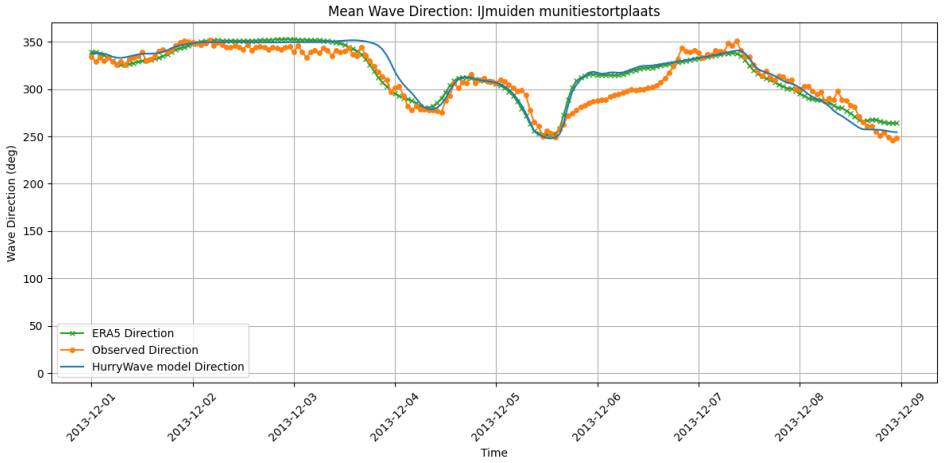

Figure 4.24: Mean wave direction at IJmuiden Munitiestortplaats during the Sinterklaas storm (1–8 December 2013), comparing HurryWave, ERA5, and buoy data. The figure shows consistent trends between models and observations, with a slight lag in HurryWave's directional shift before the storm and an overestimation of northerly directions during peak storm conditions.

Figure 4.24 shows that both the HurryWave model, and the WAM model perform well in modeling general wave direction during the Sinterklaas storm.One of the differences shown here is the delay in direction change between 3 and 4 December before the storm. The HurryWave model seems to lag behind the buoy data, a behavior not shown by the WAM model used in ERA5. On the contrary, during the storm between 5 and 6 December, both models tend to overestimate the angle, showing the storm coming from a more Northerly direction compared to the buoy's Westerly direction. This is something not shown at the other stations such as Euro platform and K13 Alpha. Platform Hoorn Q1-A shows on the contrary an underestimation of the angle, where the buoy recorded a more Northerly mean wave direction. The table with the directional statistics for this one week simulation is shown in table 4.10.

Table 4.10: Performance metrics for mean wave direction between 1 and 8 December 2013. Comparison between HurryWave and ERA5.

Station	Metric	HurryWave	ERA5
Euro Platform	Mean error (°)	2.34	2.84
	Normalized mean error (%)	0.65	0.79
	Std error (°)	10.05	8.53
	Normalized std error (%)	2.79	2.37
	RMSE error (°)	10.32	8.99
	Normalized RMSE error (%)	2.87	2.50
K13 Alpha	Mean error (°)	1.76	1.51
	Normalized mean error (%)	0.49	0.42
	Std error (°)	11.51	5.82
	Normalized std error (%)	3.20	1.62
	RMSE error (°)	11.64	6.02
	Normalized RMSE error (%)	3.23	1.67
IJmuiden Munitiestortplaats	Mean error (°)	4.62	1.72
	Normalized mean error (%)	1.28	0.48
	Std error (°)	11.29	7.38
	Normalized std error (%)	3.14	2.05
	RMSE error (°)	12.20	7.58
	Normalized RMSE error (%)	3.39	2.11
Hoorn Q1-A	Mean error (°)	-5.77	3.11
	Normalized mean error (%)	-1.60	0.86
	Std error (°)	10.78	8.09
	Normalized std error (%)	2.99	2.25
	RMSE error (°)	12.22	8.67
	Normalized RMSE error (%)	3.40	2.41

4.3.3. Analysis on multiple years

After the analysis on a single year, and on a single storm, the analysis was also conducted on multiple years of data to observe the performance of the HurryWave model on a longer time series. The simulations are run from 1950 to 2023, spanning 73 years of hourly data. In these years, the model outputs were concatenated into a single dataset to be compared with a single dataset obtained from the hourly buoy measurements. After matching the times between the model output and the buoy measurements, the output was plotted and compared using the statistics metrics described in 3.8.1.

The output of the long term reanalysis at Euro platform is shown in figures 4.25 for significant wave height and in figure 4.26 for peak wave period. This plot showcases the lack of wave period data. It is shown here at Euro platform, however, all stations lack wave period data before 2000. The figure also shows a lack of significant wave height data as the plot only starts in 1983 when buoy data is available.

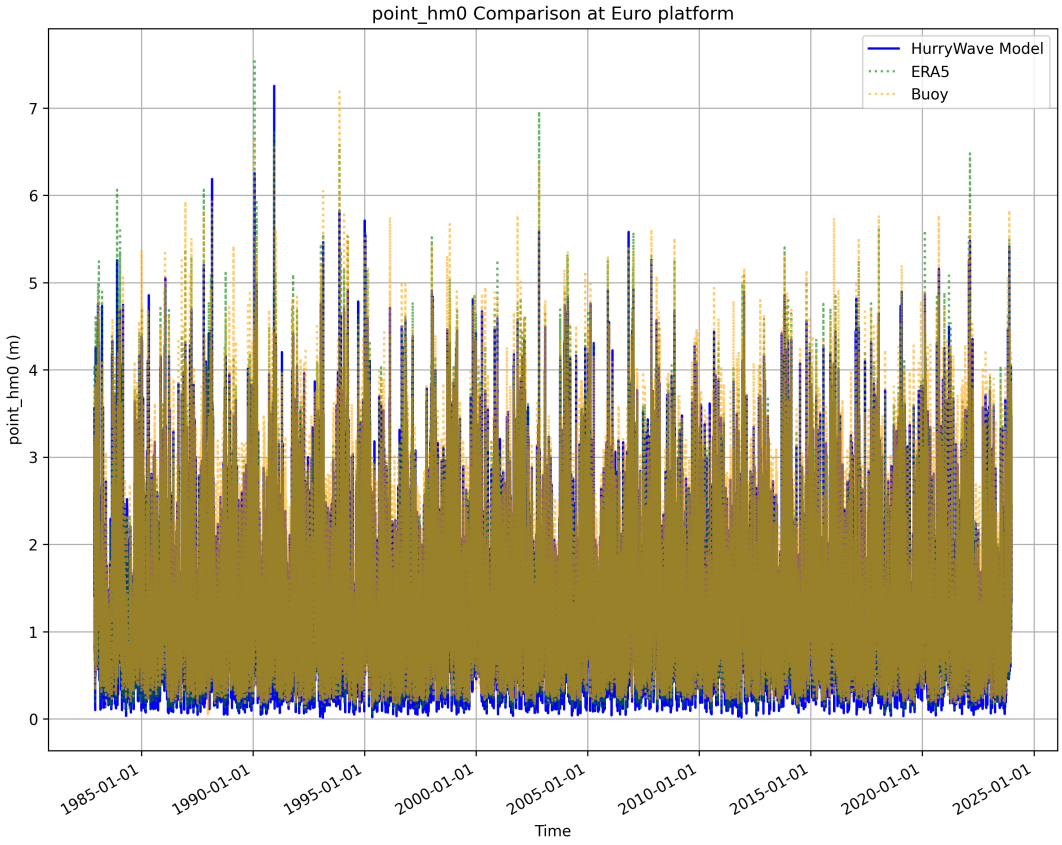

Figure 4.25: Time series comparison of significant wave height (H_{m0}) at Euro Platform for the full reanalysis period (1950–2023). The figure shows the concatenated HurryWave model output against buoy observations. The absence of data before 1983 highlights the temporal limitations of the observational record.

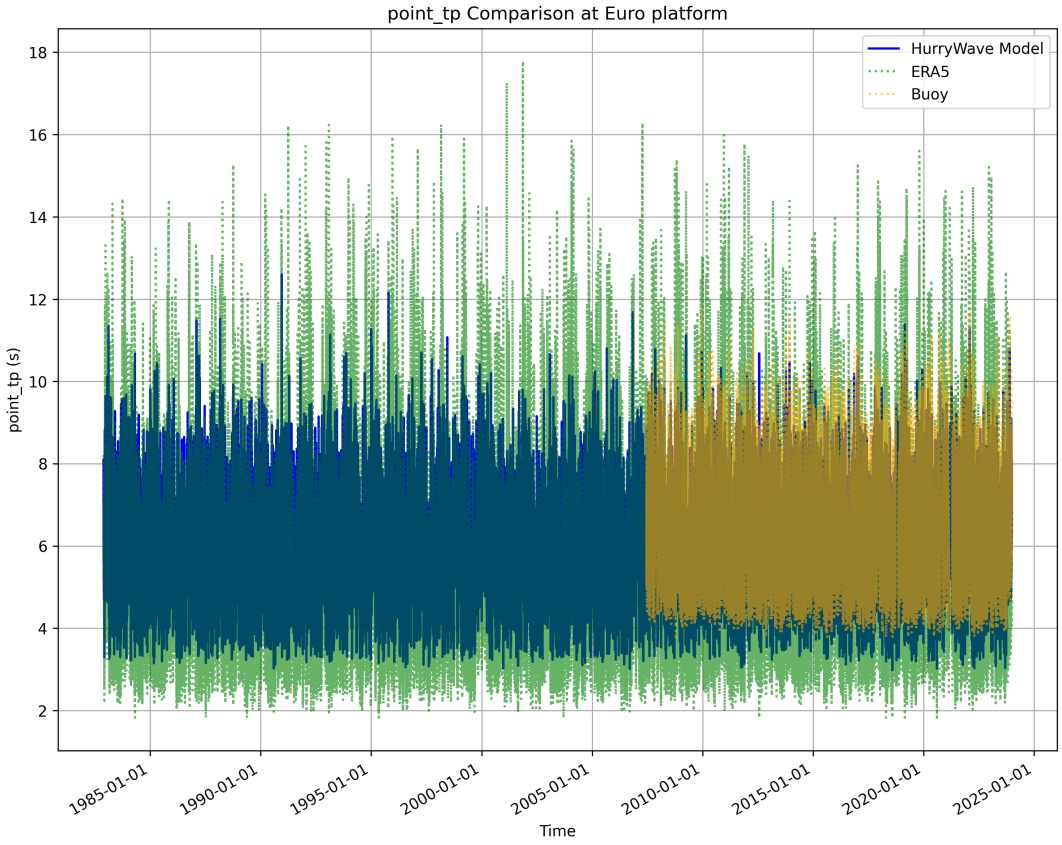

Figure 4.26: Time series comparison of peak wave period (T_p) at Euro Platform for the full reanalysis period (1950–2023). The figure shows the concatenated HurryWave model output against buoy observations. The absence of data before 2000 highlights the temporal limitations of the observational record.

A scatter plot and sorted quantile plot of both the significant wave height and peak period is shown in figure 4.27 for Euro platform, where values obtained from the HurryWave model as well as the WAM model from ERA5 are shown. This comparison was also done for other stations and can be found in Appendix F.

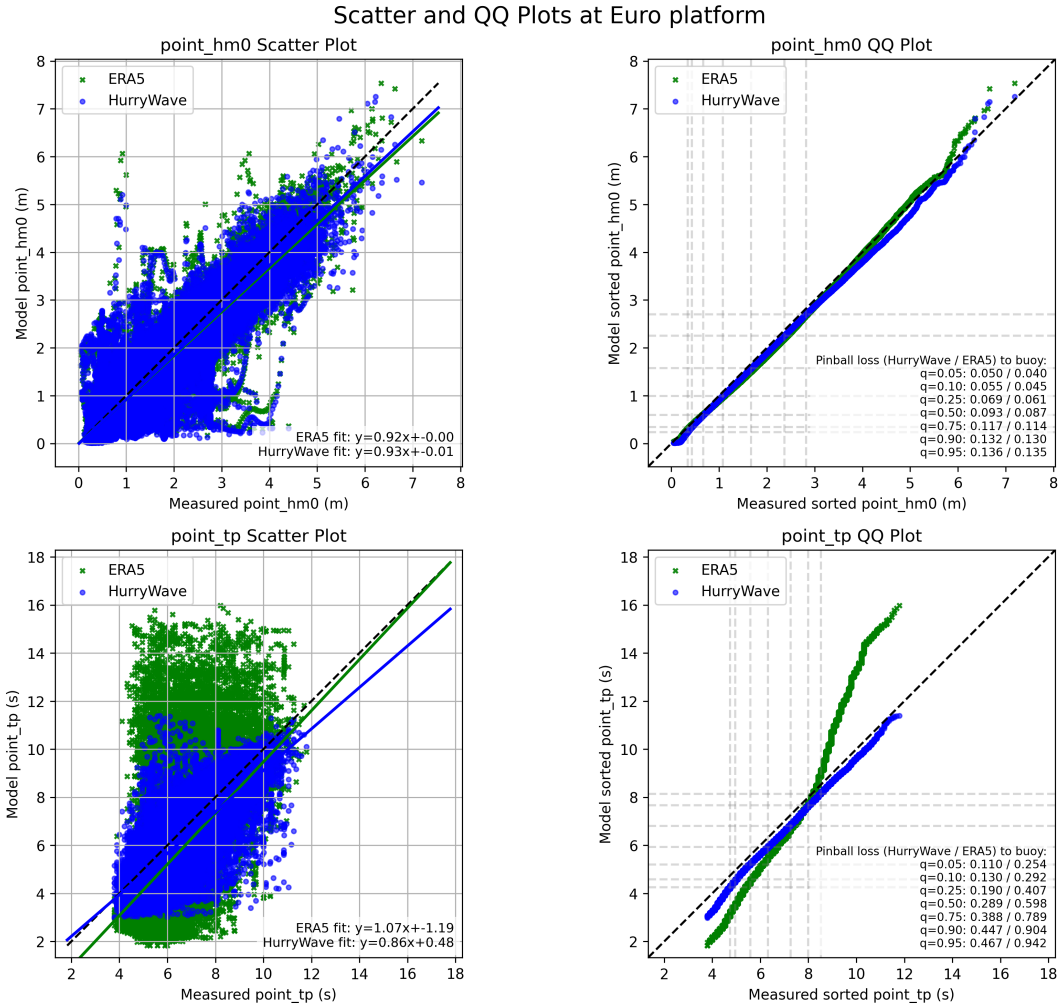

Figure 4.27: Comparison of HurryWave and ERA5 results against buoy observations at Euro Platform for the period 1950–2023. The left panel shows scatter plots of modeled versus observed H_{m0} and T_p, while the right panel presents sorted quantile plots. The results reveal a similar pattern to the 2023 single-year analysis: significant wave height is well captured, whereas peak period is systematically underestimated at higher quantiles. ERA5 underestimates low quantiles and overestimates high ones, indicating greater variability but reduced central accuracy compared to HurryWave.

Figure 4.27 shows the same trends as when simulating a single year as shown in figure 4.13 with the significant wave height being modeled accurately and the peak period being underestimated at high quantiles. ERA5 also shows issues when modeling peak period, in this case underestimating low quantiles and overestimating large quantiles with a large variation. The table showing the quantitative statistics for this 73 year simulation is shown for significant wave height in table 4.11, and for peak period in table 4.12

Table 4.11: Performance metrics for significant wave height H_{m0} between 1950 and 2023. Comparison between HurryWave and ERA5.

Station	Metric	HurryWave	ERA5
Euro Platform	RMSE (m)	0.265	0.259
	SI (-)	0.210	0.205
	Bias (m)	-0.096	-0.106
	Pinball loss $q = 0.90$ (m)	0.132	0.130
	Pinball loss $q = 0.95$ (m)	0.136	0.135
K13 Alpha	RMSE (m)	0.297	0.290
	SI (-)	0.201	0.197
	Bias (m)	-0.115	-0.093
	Pinball loss $q = 0.90$ (m)	0.149	0.133
	Pinball loss $q = 0.95$ (m)	0.155	0.138
IJmuiden Munitiestortplaats	RMSE (m)	0.322	0.325
	SI (-)	0.247	0.249
	Bias (m)	-0.075	-0.107
	Pinball loss $q = 0.90$ (m)	0.131	0.144
	Pinball loss $q = 0.95$ (m)	0.135	0.149
North Cormorant	RMSE (m)	0.972	0.432
	SI (-)	0.366	0.163
	Bias (m)	-0.728	0.024
	Pinball loss $q = 0.90$ (m)	0.677	0.127
	Pinball loss $q = 0.95$ (m)	0.714	0.126
D15-A	RMSE (m)	0.427	0.445
	SI (-)	0.312	0.325
	Bias (m)	0.083	0.204
	Pinball loss $q = 0.90$ (m)	0.092	0.051
	Pinball loss $q = 0.95$ (m)	0.088	0.041
F16-A	RMSE (m)	0.353	0.335
	SI (-)	0.232	0.220
	Bias (m)	-0.031	0.042
	Pinball loss $q = 0.90$ (m)	0.120	0.078
	Pinball loss $q = 0.95$ (m)	0.121	0.076
Hoorn Q1-A	RMSE (m)	0.347	0.344
	SI (-)	0.252	0.250
	Bias (m)	-0.105	-0.061
	Pinball loss $q = 0.90$ (m)	0.146	0.122
	Pinball loss $q = 0.95$ (m)	0.151	0.125

Table 4.12: Performance metrics for peak period T_p between 1985 and 2023. Comparison between HurryWave and ERA5.

Station	Metric	HurryWave	ERA5
Euro Platform	RMSE (s)	0.753	1.533
	SI (-)	0.117	0.238
	Bias (s)	-0.397	-0.765
	Pinball loss $q = 0.90$ (s)	0.448	0.904
	Pinball loss $q = 0.95$ (s)	0.467	0.942
K13 Alpha	RMSE (s)	1.089	1.453
	SI (-)	0.155	0.207
	Bias (s)	-0.703	-0.877
	Pinball loss $q = 0.90$ (s)	0.672	0.928
	Pinball loss $q = 0.95$ (s)	0.707	0.972
IJmuiden Munitiestortplaats	RMSE (s)	0.907	1.407
	SI (-)	0.135	0.210
	Bias (s)	-0.517	-0.735
	Pinball loss $q = 0.90$ (s)	0.535	0.839
	Pinball loss $q = 0.95$ (s)	0.560	0.876
D15-A	RMSE (s)	1.729	1.838
	SI (-)	0.225	0.240
	Bias (s)	-1.197	-1.016
	Pinball loss $q = 0.90$ (s)	1.091	1.085
	Pinball loss $q = 0.95$ (s)	1.151	1.136
F16-A	RMSE (s)	1.207	1.389
	SI (-)	0.164	0.188
	Bias (s)	-0.866	-0.801
	Pinball loss $q = 0.90$ (s)	0.804	0.881
	Pinball loss $q = 0.95$ (s)	0.847	0.921
Hoorn Q1-A	RMSE (s)	1.004	1.411
	SI (-)	0.147	0.207
	Bias (s)	-0.573	-0.735
	Pinball loss $q = 0.90$ (s)	0.569	0.833
	Pinball loss $q = 0.95$ (s)	0.598	0.870

The performance comparison in tables 4.11 and 4.12 shows that HurryWave and ERA5 yield similar accuracy for significant wave height (H_{m0}), with both models achieving low RMSE, SI, and bias values at most stations. Differences are small, though ERA5 performs better at North Cormorant, while HurryWave slightly outperforms ERA5 at some Dutch coastal stations (e.g., D15-A and IJmuiden) in terms of bias and extreme quantile representation. For peak period (T_p), however, HurryWave consistently outperforms ERA5 across all stations, with notably lower RMSE, scatter index, and pinball loss values. This indicates that while both datasets are comparable for wave height, HurryWave provides a clear advantage in reproducing peak periods, especially in the higher quantiles relevant for storm conditions.

The mean wave direction also shows the same trend as previously shown in the 1 year experiment, where large errors are found between the modeled values and the buoy data as shown in figure 4.28. The error in direction histogram for all compared years for all stations can be found in Appendix F.

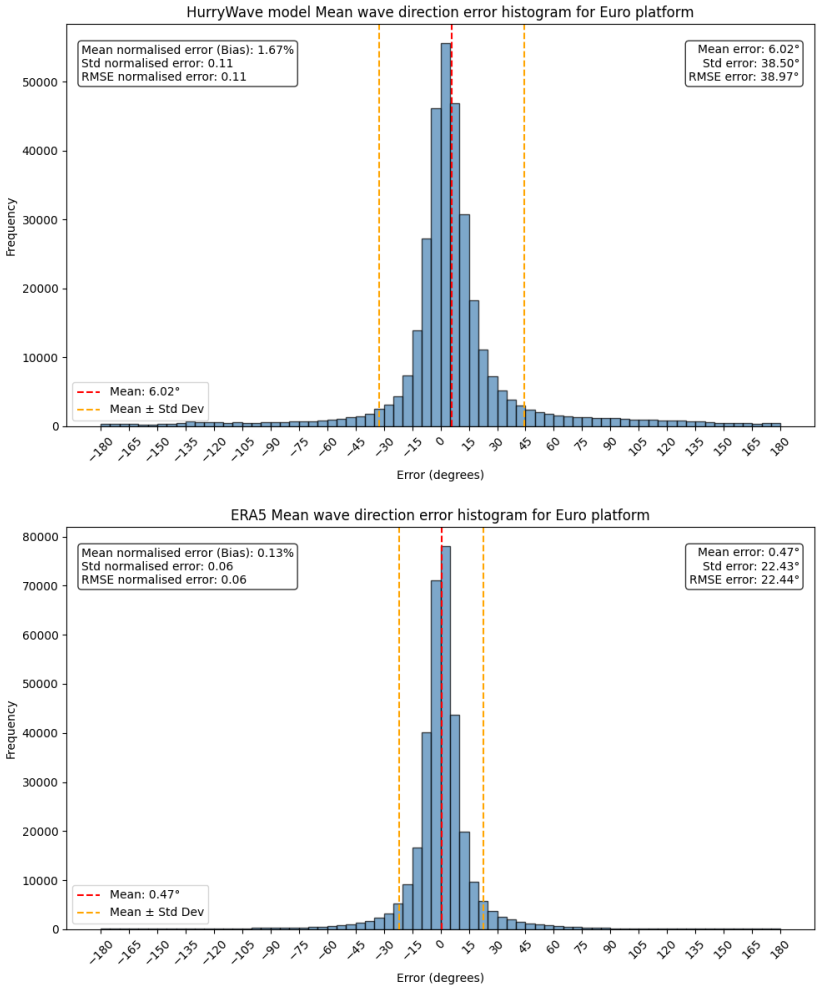

Figure 4.28: Histograms of mean wave direction error at Euro Platform for all years between 1950 and 2023. The top panel shows HurryWave model results and the bottom panel ERA5. The comparison illustrates that while both datasets exhibit broad distributions of directional error, ERA5's histogram is more concentrated around zero, reflecting higher directional accuracy. HurryWave displays a wider spread (±40°), indicating greater variability in modeled directions.

The table showing the quantitative analysis of wave direction error is shown in table 4.13. For mean wave direction, ERA5 consistently outperforms HurryWave across all stations. ERA5 shows lower mean and normalized errors, as well as reduced standard deviation and RMSE values, indicating a better alignment with observations. HurryWave exhibits larger directional biases (typically around 5–7°) and higher variability, especially at offshore locations such as North Cormorant, where errors are most pronounced. Overall, while HurryWave provides reasonable directional estimates, ERA5 demonstrates substantially higher skill in reproducing wave direction.

Table 4.13: Performance metrics for mean wave direction between 1950 and 2023. Comparison between HurryWave and ERA5.

Station	Metric	HurryWave	ERA5
Euro Platform	Mean error (°)	6.02	0.47
	Normalized mean error (%)	1.72	0.13
	Std error (°)	38.50	22.43
	Normalized std error (-)	0.11	0.06
	RMSE error (°)	39.97	22.43
	Normalized RMSE error (-)	0.11	0.06
K13 Alpha	Mean error (°)	4.95	-0.09
	Normalized mean error (%)	1.38	-0.03
	Std error (°)	34.51	19.98
	Normalized std error (-)	0.10	0.06
	RMSE error (°)	34.86	19.98
	Normalized RMSE error (-)	0.10	0.06
IJmuiden Munitiestortplaats	Mean error (°)	6.57	-0.81
	Normalized mean error (%)	1.83	-0.22
	Std error (°)	35.33	22.13
	Normalized std error (-)	0.10	0.06
	RMSE error (°)	35.93	22.15
	Normalized RMSE error (-)	0.10	0.06
North Cormorant	Mean error (°)	-8.45	-10.13
	Normalized mean error (%)	-2.35	-2.82
	Std error (°)	57.09	38.51
	Normalized std error (-)	0.16	0.11
	RMSE error (°)	57.71	39.82
	Normalized RMSE error (-)	0.16	0.11
Hoorn Q1-A	Mean error (°)	2.09	0.25
	Normalized mean error (%)	0.58	0.07
	Std error (°)	40.36	21.60
	Normalized std error (-)	0.11	0.06
	RMSE error (°)	40.41	21.60
	Normalized RMSE error (-)	0.11	0.06

4.3.4. Reanalysis discussion

The reanalysis across different temporal scales reveals a consistent picture of HurryWave's capabilities and limitations. For the single-year simulation (2023), the model demonstrates strong agreement with buoy observations in terms of significant wave height, with scatter indices comparable to ERA5 and biases generally smaller. This suggests that under typical yearly variability, including both calm summer periods and winter storms, the model can provide reliable estimates of wave climate. The peak period, however, tends to be underestimated, especially during high-energy events. It is still more reliable than the modeled peak period found in ERA5 with large variations.

Focusing on the short-term perspective, the analysis of the Sinterklaas storm highlights the model's challenges in high-energy conditions. While HurryWave captured the general storm evolution at Euro Platform, underestimations of peak period and wave height were evident at several offshore stations such as D15-A and North Cormorant. This suggests that the model struggles with extremes. ERA5 performed somewhat better in terms of storm peak representation, although it too displayed difficulties in peak period and wave direction.

Extending the perspective to the full reanalysis period (1950–2023) shows that HurryWave maintains consistent performance over long time series. Significant wave heights are reproduced with accuracy comparable to ERA5, including similar RMSE and scatter index values. Nevertheless, the underestimation of larger waves and the systematic underestimation of peak periods at high quantiles persists, echoing findings from the one-year and storm-scale analyses. Wave direction emerges as the weakest parameter in HurryWave: large errors and high variability compared to ERA5 are observed across

stations.

When compared to literature values reviewed in chapter 2, the model's biases fall within expected ranges. Previous studies have documented ERA5's underestimation of wave energy by up to 42% in certain regions (Afolabi et al., 2024), as well as its difficulty in capturing nearshore processes due to coarse resolution. HurryWave mirrors some of these limitations, but its relative advantage lies in reduced computational cost. At the same time, the persistent underestimation of peak period and the weaker performance in directional estimates align with earlier observations that simplified spectral models may trade off realism for efficiency.

The multi-scale reanalysis suggests that HurryWave is a robust model for long-term wave climate assessments in the North Sea, particularly for significant wave height. However, caution is needed when applying it to storm extremes or for applications where wave direction and peak period are critical. These findings are in line with the broader body of literature, underscoring the importance of model choice depending on the intended use.

4.4. Demonstrating model performance

4.4.1. Results of model performance analysis

To see if these statistics are significant, a hypothesis test was conducted using the methodology described in subsection 3.8.2. This test was conducted on the significant wave height H_{m0} and the peak period T_p.

Calculating number of effective points

Firstly, the number of effective points is calculated from the first lag in the autocorrelation of the error in the time series. The error is calculated as the difference between model value and measured value. The autocorrelation was then calculated on the errors and plotted to obtain the number of effective points which can be used ensuring the points are independent and identically distributed (iid). From the first lag, the number of effective points is calculated using equation 3.12. The plots of the error autocorrelation for significant wave height H_{m0} and peak period T_p, along with the associated first lag value (r_1) and the number of effective points (N_{eff}) are shown in 4.29 and 4.30.

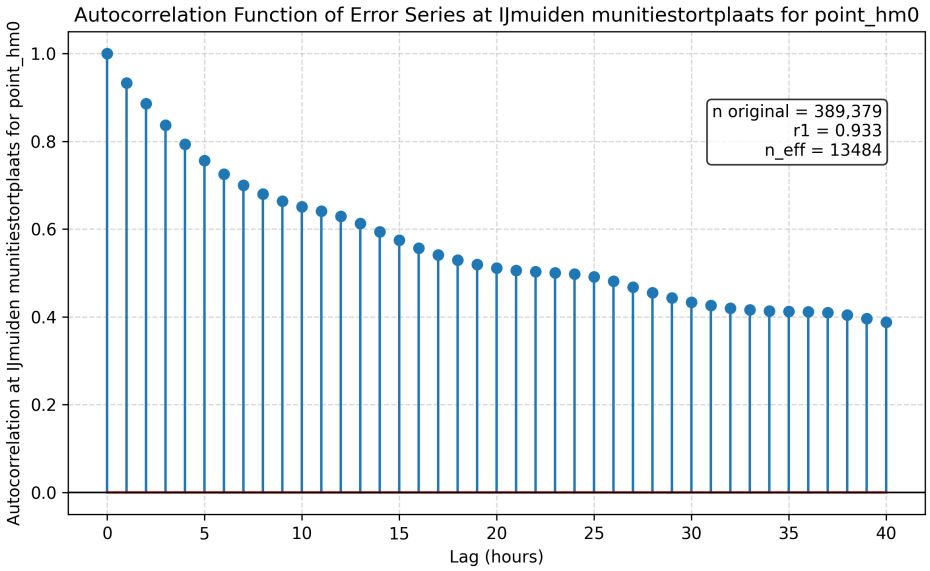

Figure 4.29: Autocorrelation function (ACF) of the model–observation error series for significant wave height (H_{m0}) at IJmuiden Munitiestortplaats, derived from the full 1950–2023 dataset. The strong persistence observed in the first lag ($r_1 = 0.86$) indicates that errors are temporally correlated, underscoring the need for autocorrelation-aware hypothesis testing methods such as block bootstrapping.

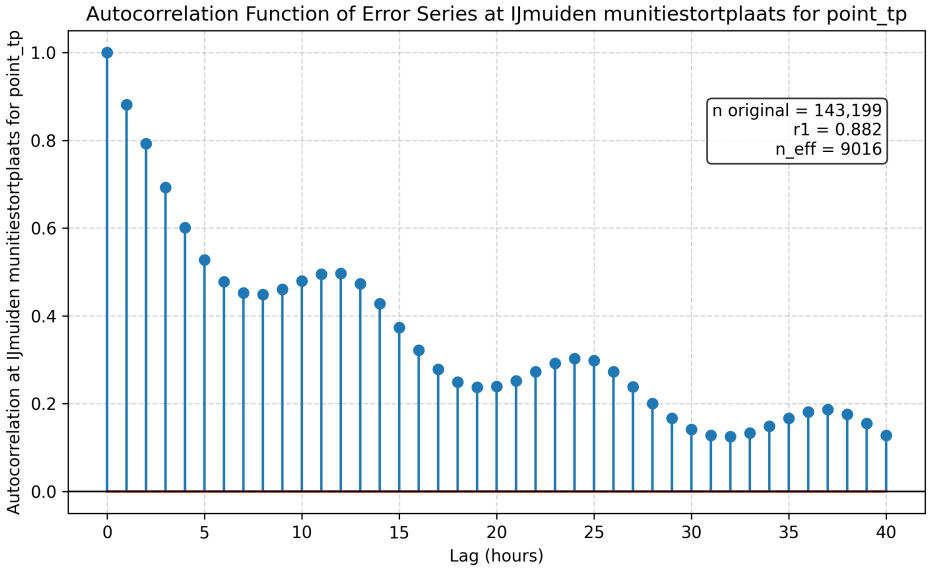

Figure 4.30: Autocorrelation function of the peak period (T_p) error series at IJmuiden Munitiestortplaats for the full reanalysis. The figure reveals significant autocorrelation at short lags and a repeating 12–24 hour oscillation pattern, suggesting tidal influence. This persistence in errors violates the assumption of independent samples, motivating the use of block bootstrap methods for subsequent statistical testing.

These figures show that the errors are heavily correlated in time, which is expected as larger waves tend to lead to larger errors as shown in figure 4.27. Given that large waves often occur in a storm, larger errors also would occur in a storm. The wave period error also seems to show seasonality where waves after 12 and 24 hours share a larger similarity with the original wave than those at 11,13,23, and 25 hours. A simple explanation for this could be the tides given the particular period to this seasonality of 12 and 24 hours.

The output of the number of effective points for each station is shown in table 4.14

Table 4.14: Autocorrelation results for significant wave height (H_{m0}) and peak wave period (T_p).

Platform	H_{m0}			T_p		
	Original size	Lag-1	N_{eff}	Original size	Lag-1	N_{eff}
Euro platform	358722	0.862	26624	143939	0.918	6117
Hoorn Q1-A	120845	0.680	23054	120845	0.750	17268
North Cormorant	159017	0.924	6306	–	–	–
K13 Alpha	392907	0.882	24569	140267	0.936	4644
F16-A	96864	0.921	3992	96864	0.936	3207
IJmuiden munitiestortplaats	389395	0.933	13486	143217	0.882	9019

Table 4.14 shows that the lag-1 autocorrelation is always above 0.8, meaning data is heavily auto correlated. Additionally, the effective sample size is always a large amount smaller than that of the original sample size. This quantitatively demonstrates the need for an autocorrelation aware method to test hypothesis on this wave data.

Visualization of the autocorrelation

The autocorrelation of both the significant wave height and peak period is plotted for all stations to observe the persistence and seasonality of the model output and the buoy data. These are plotted for IJmuiden munitiestortplaats in figures 4.31 and 4.32. This plot shows the very slow decay of correlation between both significant wave heights and peak periods. This means autocorrelation aware statistics need to be used in order to run hypothesis testing. Additionally, seasonality is observed on the wave period where correlation seems to increase every 12 hours, most likely due to tidal influence.

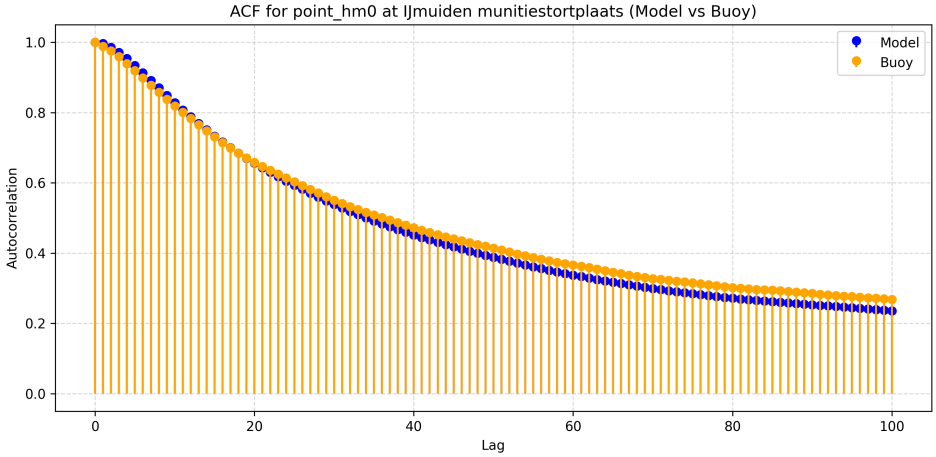

Figure 4.31: Autocorrelation function of the modeled and observed significant wave height (H_{m0}) at IJmuiden Munitiestortplaats. The slow decay of correlation over time demonstrates strong persistence in wave height, typical for marine environmental variables. This behavior confirms that conventional hypothesis testing would overestimate statistical significance without adjusting for autocorrelation.

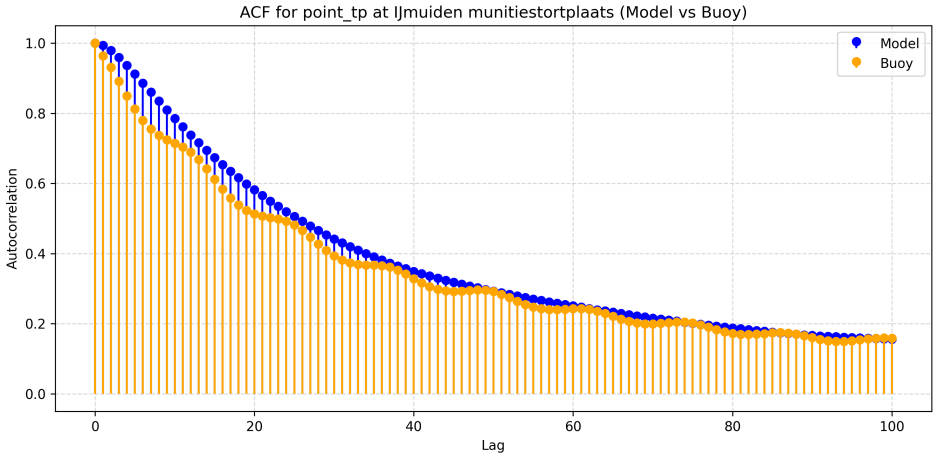

Figure 4.32: Autocorrelation function of modeled and observed peak period (T_p) at IJmuiden Munitiestortplaats over the full record. The cyclic peaks at 12-hour intervals suggest tidal modulation of wave period, further reinforcing the need for temporal dependence-aware statistical evaluation.

Results of the block bootstrap method

To remove this correlation, the block bootstrap method is used in hypothesis testing to avoid a violation of the independent identically distributed assumption. The bootstrap method was used as described in section 3.8.2. This methodology is used to create a synthetic dataset of the error metric from random points in the data while conserving the autocorrelation. This in turn, can be used to test certain hypothesis.

The parameters for the block bootstrap depend on the data itself. To avoid seasonality, a block size of 1 year is chosen with 2000 bootstrap samples as the need to predict the extremes around the tail was large. This differs from the 1000 samples used in literature (Becker & Roggenbuck, 2023), and is an assumption. The issue with this method choice is the lack of data from the peak period. Block bootstrap requires at least 20 blocks to be considered effective (Chandy et al., 2025), given the lack of data, peak period cannot be tested using this method.

A test to check if the blocks are indeed independent and identically distributed (iid) a visual test is performed using the autocorrelation of the means and standard deviations of each block in figure 4.33.

Another figure showing the evolution of the block mean and standard deviation of the HurryWave model and the measured buoy is shown in figure 4.34. These should not show a trend and should not differ too largely. These values do show some noise, however, clear trends or seasons are difficult to visualize qualitatively. Quantitatively, the Mann-Kendall and the Ljung-Box test are used to determine if a trend is present, or if the noise exhibited in the data is indeed white noise respectively. The iid assumption can be taken based on evidence from these tests. Figure 4.34 shows that IJmuiden munitiestortplaats passes the tests, however most stations do not. As an example, Euro platform fails the Mann-Kendall test showing that there is a long term trend in the data. This data is difficult to de-trend as this could affect the comparisons between the HurryWave model and the buoy data.

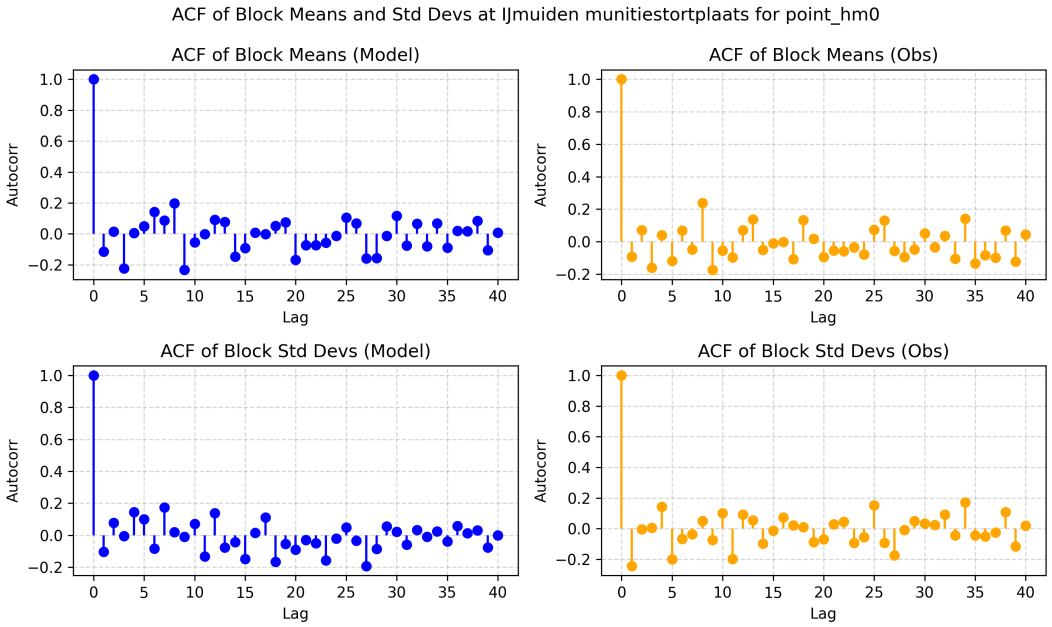

Figure 4.33: Autocorrelation of the block means and standard deviations of H_{m0} at IJmuiden Munitiestortplaats, used to test the independence of bootstrap blocks. The minimal correlation between consecutive blocks indicates that the chosen block length (1 year) sufficiently reduces temporal dependence for use in block bootstrap hypothesis testing.

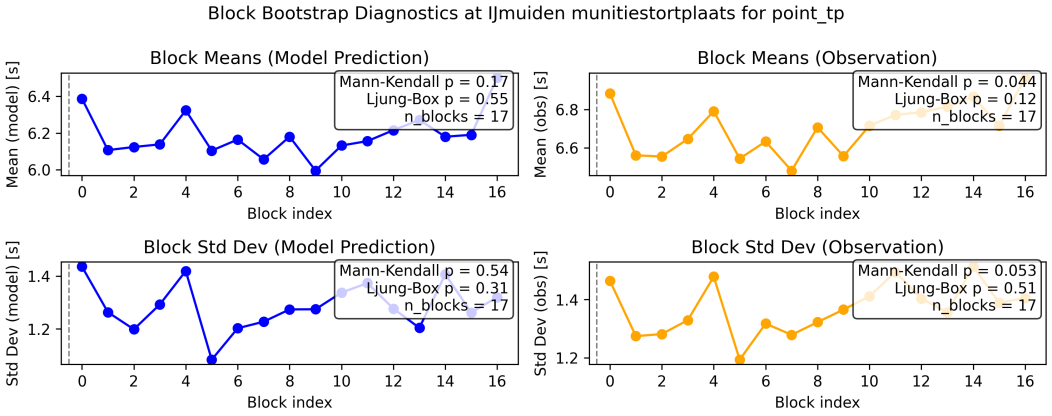

Figure 4.34: Time series of block mean and standard deviation for peak period (T_p) at IJmuiden Munitiestortplaats. Each point represents the average and spread of one yearly block. The figure shows relatively stable variability with no clear long-term trend, supporting the assumption of stationarity necessary for bootstrap resampling, despite some interannual fluctuations.

A table summarizing the number of blocks and the numerical values of these tests is shown in tables 4.15 and 4.16

Table 4.15: Bootstrap block statistics for H_{m0} at selected stations

Station	# Blocks	Block means MK test pass (p>0.05)	Block means LB test pass (p>0.05	Block std MK test pass (p>0.05	Block std LB test pass (p>0.05
Euro platform	41	No	Yes	No	Yes
K13 Alpha	45	Yes	Yes	No	Yes
IJmuiden Munitistortplaats	45	Yes	Yes	Yes	Yes
North Cormorant	19	Yes	Yes	Yes	Yes
F16-A	12	Yes	Yes	Yes	Yes
Hoorn Q1-A	14	Yes	No	Yes	No

Table 4.16: Bootstrap block statistics for Tp at selected stations

Station	# Blocks	Block means MK test pass (p>0.05)	Block means LB test pass (p>0.05	Block std MK test pass (p>0.05	Block std LB test pass (p>0.05
Euro platform	17	Yes	Yes	Yes	Yes
K13 Alpha	17	Yes	Yes	No	Yes
IJmuiden Munitistortplaats	17	No	Yes	Yes	Yes
F16-A	12	Yes	Yes	Yes	Yes
Hoorn Q1-A	14	Yes	Yes	Yes	Yes

Tables 4.15 and 4.16 show that very few stations are fully compliant with all necessary requirements for hypothesis testing. These requirements include the blocks being iid and the minimum number of blocks being at least 20. Nevertheless, the final hypothesis test can only be valid for one station and one wave parameter, IJmuiden Munitiestortplaats for significant wave height. This station and parameter are therefore used in further analysis.

The block bootstrap is then performed using the methodology described in 3.8.2. The output of this bootstrap is shown in figure 4.35 at Euro platform for RMSE, Scatter index, and bias for significant wave height (H_{m0}).

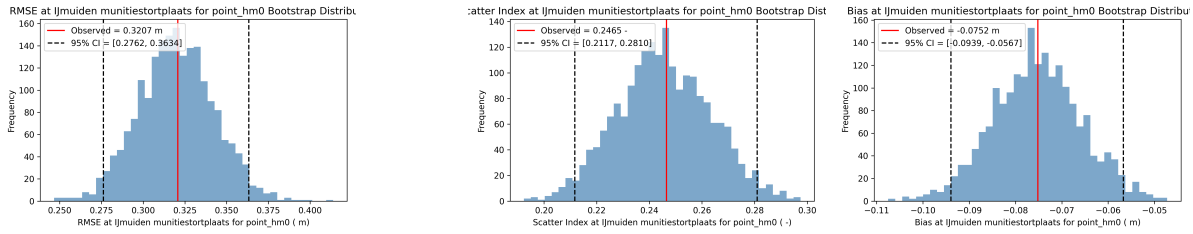

Figure 4.35: Distribution of bootstrap samples for RMSE (left), scatter index (center), and bias (right) of significant wave height (H_{m0}) at IJmuiden Munitiestortplaats. These histograms represent 2000 resampled metrics computed via block bootstrap using 1-year blocks. The narrow spread and consistent central tendency confirm the stability of the model's performance and support the statistical significance of its skill estimates.

The statistical results obtained in table 4.11 in subsection 4.4.1 are then verified through a statistical test, the output of which is shown in table 4.17.

Table 4.17: Statistical test results for H_{m0}.

Station	Metric	Value	95% CI	H_0 tested	p-value
IJmuiden munitiestortplaats	Bias	-0.0752 m	[-0.0931, -0.0581] m	$Bias \geq 0m$	$< 5.0 \times 10^{-4}$
	RMSE	0.3207 m	[0.2770, 0.3667] m	$RMSE \geq 0.5m$	$< 5.0 \times 10^{-4}$
	SI	0.2465	[0.2125, 0.2798]	$SI \geq 0.4$	$< 5.0 \times 10^{-4}$

This method shows that significant wave height is modeled accurately by the HurryWave model at IJmuiden munitiestortplaats given an RMSE value below 0.5m, a scatter index below 0.4 and a negative bias showing a slight underestimation. This is shown in the scatter plot of the modeled significant wave height against the measured one in figure 4.36.

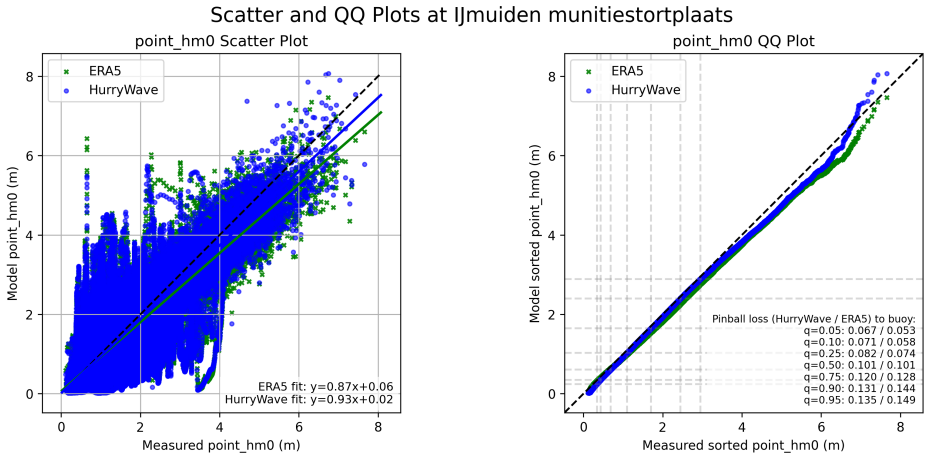

Figure 4.36: Scatter (left) and QQ (right) plot comparing modeled versus observed significant wave height (H_{m0}) at IJmuiden Munitiestortplaats for the entire 1950–2023 period. The dense clustering around the 1:1 line demonstrates a strong correlation and small bias between HurryWave and observations, validating the model's ability to reproduce wave height variability over multi-decadal timescales.

4.4.2. Discussion of model performance analysis

The statistical tests applied to the HurryWave reanalysis provide a more rigorous picture of model skill than the raw performance metrics alone. By explicitly accounting for autocorrelation in the time series, the analysis avoided the risk of inflated significance levels that would arise from treating highly correlated environmental data as independent. The block bootstrap methodology allowed effective sample sizes to be estimated and ensured that confidence intervals for the bias, RMSE, and scatter index reflected the persistence inherent in wave data.

The results show that at least one station, IJmuiden Munitiestortplaats. passed the requirements for independent and identically distributed (iid) sampling blocks, enabling a valid hypothesis test. At this location, the HurryWave model was demonstrated to simulate significant wave height accurately, with RMSE consistently below 0.5 m, scatter index below 0.4, and a negative bias confirming a slight underestimation.

The method also highlighted important limitations. Only a subset of stations could be included in the formal hypothesis tests, as many failed to meet iid assumptions or had insufficient block counts. This restriction means the statistical "proof" of performance cannot be generalized to all stations or to all wave parameters. Peak period, in particular, could not be validated with the same rigor due to shorter data records and strong seasonal autocorrelation. Furthermore, the reliance on block bootstrapping assumes that persistence in errors is uniform within blocks, which may not always hold during rapidly evolving storm conditions.

While the methodology successfully demonstrated that HurryWave achieves statistically significant skill in simulating significant wave heights at IJmuiden munitiestortplaats, the conclusions are spatially and temporally constrained. The limitations underline the need for longer and more consistent buoy records, as well as further methodological refinements to extend statistical validation to peak period and directional parameters. This reflects a broader challenge in wave model evaluation: robust statistical proof of model performance is often only feasible where long, high-quality observational datasets exist.

4.5. Analysis of the wave spectra

An analysis of the wave spectra was conducted using the output of the HurryWave model, and data collected from spectral buoys. These buoys include stations A12, Hoorn Q1-A, and Euro platform shown

in figure 3.11.. The spectral data were compared during the Sinterklaas storm from 5 to 6 December 2013. The data were compared each hour at each platform. The data were then plotted in polar plots as shown in figure 4.37.

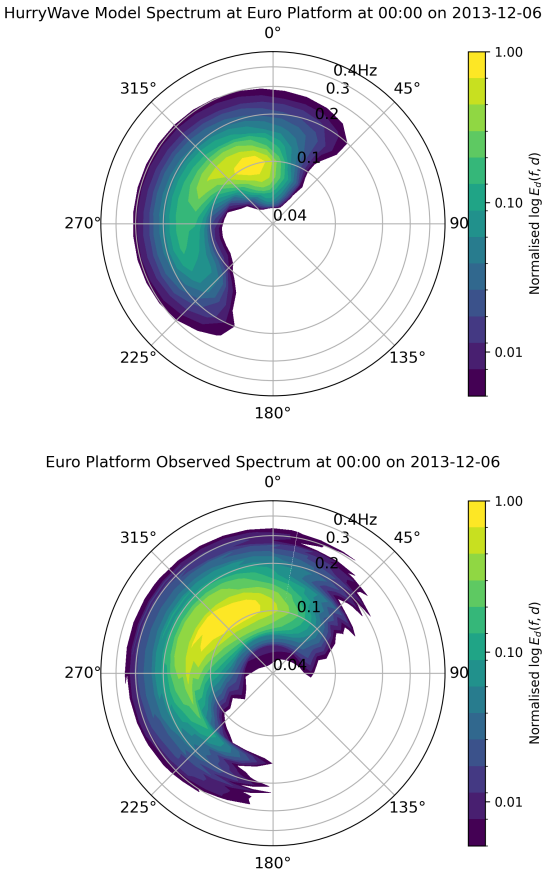

Figure 4.37: Comparison of wave energy spectra from HurryWave (top) and spectral buoy data (bottom) at Euro Platform on 6 December 2013 at 00:00. The plots show wave energy as a function of frequency and direction during the Sinterklaas storm. The HurryWave model spectrum is narrowly concentrated around 0.1 Hz and 330°, whereas the buoy data shows broader energy between 300–350°.

Figure 4.37 shows the energy distribution as a function of frequency and direction for both the output from the HurryWave model and the spectral buoy data. This data has been collected and treated as shown in section 3.9. From this figure a few striking differences can be observed. Firstly, the data shows a much larger spread than the model output. It seems the model output is concentrated near 0.1Hz (or 10s period) and around 330 degrees. This is different to the spectral buoy data, which shows the largest energy around 0.12 to 0.14 Hertz (or 7.1 to 8.3 seconds in period). Additionally, the energy is spread out between 300 and 350 degrees. This could cause HurryWave to overestimate the wave periods and misrepresent the direction. Figure 4.19 shows that indeed on June 6th 2013 at 00:00, the wave period is underestimated. The analysis on the wave direction also lines up with figure 4.24 where (in IJmuiden) the HurryWave model tended to "overestimate" the change in direction at the time and showed "higher" angles than the buoy. This means HurryWave tended to perceive the waves as more Northern than the buoy that perceived it as more Western.

This directional bias changes at station Hoorn Q1-A, where at midnight , the direction is found to be more Northern in the data. As shown in figure 4.24 the wave direction is shifting from NorthWest to North. The speed at which the shift occurs differs between stations. In this case, the change in direction occured too slowly. In figure 4.38, the spectra at Hoorn Q1-A is plotted on 5 December 2013 at 21:00 for both HurryWave model output and spectral buoy data. In figure 4.39, the same is plotted at midnight on 6 December 2013.

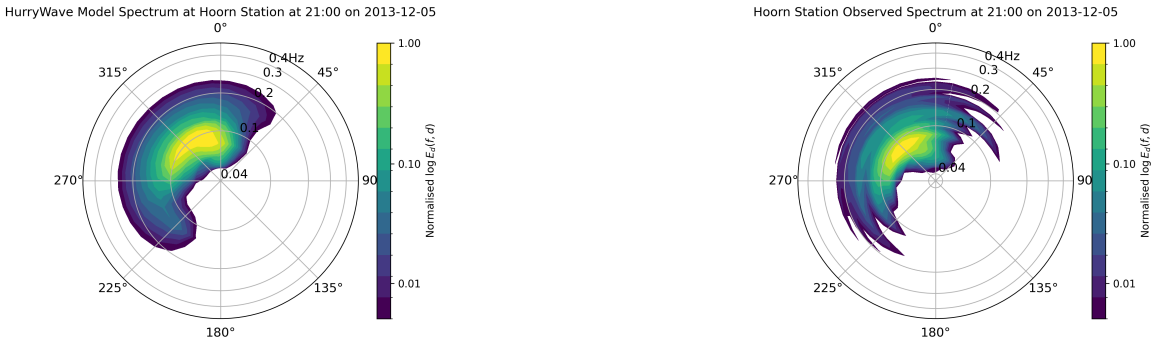

Figure 4.38: Comparison of directional-frequency wave spectra from HurryWave (left) and spectral buoy data (right) at Hoorn Q1-A on 5 December 2013 at 21:00. Both plots represent the late stages of the Sinterklaas storm's directional transition from North West to north. The model's slower response to changing wave direction suggests limited coupling between evolving wind conditions and wave propagation.

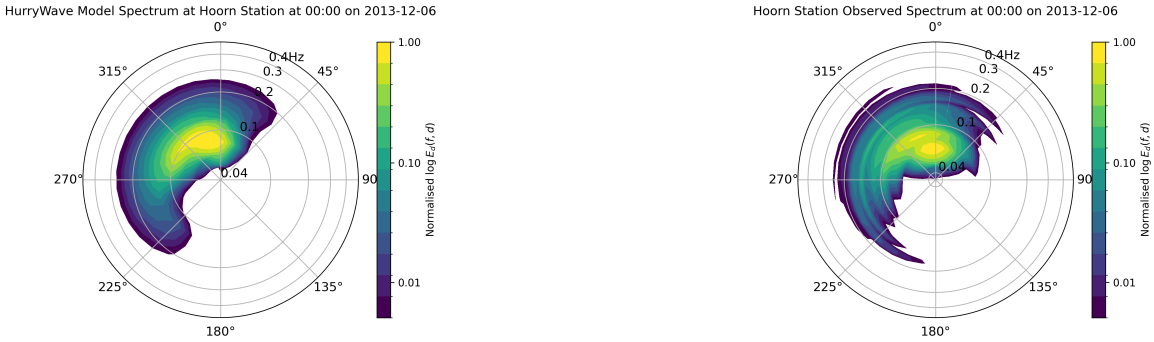

Figure 4.39: Wave energy spectra from HurryWave (left) and spectral buoy data (right) at Hoorn Q1-A on 6 December 2013 at 00:00. During this period, observed waves had rotated fully to a northerly direction, with buoy data showing a concentrated peak near 0°. HurryWave, however displays limited directional shift. This lag in wave turning confirms that the model under represents directional evolution, likely due to error in wind direction or the absence of nonlinear triad interactions

Figure 4.38 and 4.39 clearly show the delay in direction change, where the data moves from 315 to 0 degrees between 21:00 and 00:00 the next day, but the model shows very little directional change. The spread of the large energy portion across the direction is also larger in the model output, whereas it is much more concentrated in one direction in the spectra buoy data. This is also true in deeper waters at platform A12. This station is located further offshore, and is exposed to a larger amount of energy. This is shown in figure 4.40 where the energy distribution is plotted at 00:00 on 6 December 2013.

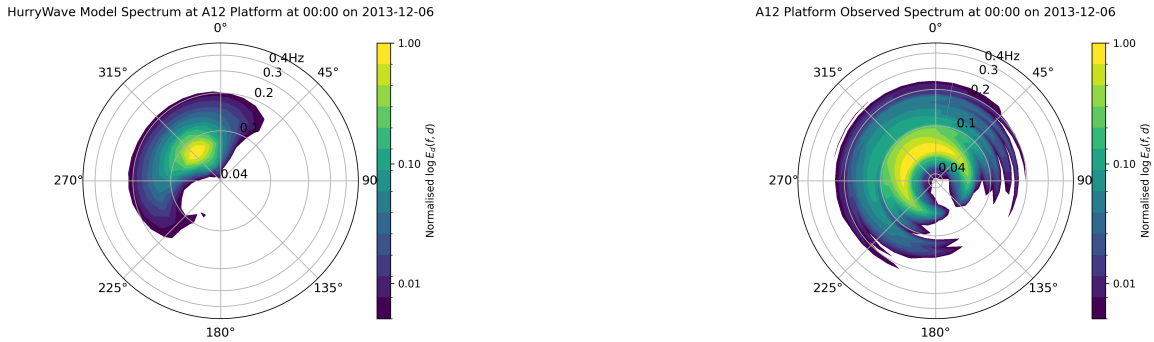

Figure 4.40: Comparison of modeled and measured wave spectra at platform A12 on 6 December 2013 at 00:00. The HurryWave model (left) shows narrowly distributed energy centered near 270° and 0.1 Hz (10 s period), while the spectral buoy data (right) exhibits a much broader directional spread (270–45°) and longer dominant periods (15 s). This highlights the model's tendency to underestimate long-period swell energy and angular dispersion in deeper waters, consistent with the observed underestimation of peak period and over-concentration of directional energy in the North Sea hindcast.

Figure 4.40 shows an immense directional spread in the data at platform A12 on 6 December 2013 at midnight. The largest proportion of energy ranges from 270 to 45 degrees. The waves recorded are also very long, around a period 15 seconds. This is very different from the model which shows a very low spread around 270 degrees and a much lower period at around 10s. This also confirms the large underestimation of the wave period in deeper waters by HurryWave.

Next to the qualitative assessment, a quantitative assessment was also performed by analyzing different formulations for wave period as described in section 3.9.

Different wave periods were compared over this storm period and the statistics for each period are given in table 4.18

Table 4.18: Model performance statistics for different periods

Platform	Parameter	RMSE [s]	Scatter Index [-]	Bias [s]
A12 platform	T_p	2.91	0.27	-2.13
	T_{m01}	0.99	0.12	-0.71
	T_{m-10}	0.24	0.08	-0.19
	T_{m02}	0.76	0.10	-0.51
Euro platform	T_p	1.58	0.22	0.55
	T_{m01}	0.95	0.17	0.74
	T_{m-10}	0.15	0.06	0.10
	T_{m02}	0.98	0.19	0.76
Platform Hoorn Q1-A	T_p	2.50	0.26	-1.15
	T_{m01}	0.61	0.09	-0.08
	T_{m-10}	0.15	0.05	-0.09
	T_{m02}	0.55	0.09	0.10

Table 4.18 compares the model performance for different definitions of the wave period across three observation platforms: A12 platform, Euro platform, and platform Hoorn Q1-A. The metrics root-mean-square error (RMSE), scatter index (SI), and bias are used to assess model accuracy and systematic deviation from observations.

Overall, the results show that the model performs best for the mean period based on the negative first spectral moment (T_{m-10}) across all locations. This period definition yields the lowest RMSE (0.15–0.24 s), the smallest scatter index (0.05–0.08), and minimal bias, indicating very close agreement between modeled and observed long-period swell components. The reason for this strong performance is that

T_{m-10} emphasizes low-frequency energy, which is often well captured by spectral wave models, as swell propagates more predictably over long distances.

By contrast, the peak period (T_p) exhibits the largest discrepancies, with RMSE values up to 2.9 s and high scatter indices (0.22–0.27). The large negative biases at A12 and Hoorn suggest that the model tends to underestimate the dominant spectral peak period, which could result from insufficient resolution of multi-peaked spectra during mixed sea states (coexisting wind sea and swell).

The mean spectral periods T_{m01} and T_{m02} show intermediate performance. Both are more stable than T_p, but less accurate than t_{m-10}. The RMSE values range between roughly 0.55 and 0.99 s, with small negative biases at A12 and Hoorn and positive biases at EPL. The difference between T_{m01} and T_{m02} is subtle but consistent: T_{m02} slightly outperforms T_{m01} at Hoorn (0.55 s vs. 0.61 s RMSE), likely because T_{m02} gives more weight to higher frequencies and thus better captures locally generated, short-period waves.

The comparison demonstrates that period definitions based on spectral moments (T_{m01}, T_{m02}, and especially T_{m-10}) provide more robust and consistent estimates of modeled wave periods than the traditional peak period. The superior performance of T_{m-10} suggests that this metric may serve as a more reliable indicator of overall wave dynamics in both offshore and nearshore conditions, particularly when assessing mixed sea states where the spectral peak is ambiguous.

5
Discussion

5.1. Key findings and discussion

This study provided a first comprehensive evaluation of the HurryWave model in the North Sea, testing it across multiple temporal scales and comparing its performance to buoy observations and the ERA5 dataset.

A key finding is that HurryWave reproduces significant wave height (H_{m0}) with accuracy comparable to ERA5, often with smaller biases and similar scatter index values. For example, in the 2023 validation at Euro Platform, HurryWave achieved an RMSE of 0.38 m and a scatter index of 0.18, compared to ERA5 with 0.46 m and 0.21, respectively. Over the full 1950–2023 period at IJmuiden, HurryWave bias remained close to zero (–0.02 m), with SI = 0.20 and RMSE = 0.44 m, confirming stable long-term skill. The extreme wave heights are however underestimated, this is also the case for ERA5 and could be due to the wind input underestimating large wind speeds. This is investigated further in section 5.2. The block bootstrap test at IJmuiden Munitiestortplaats further demonstrated model performance at over 95% statistical significance, making this station the most robust reference for HurryWave's reliability in the North Sea.

However, performance varied between parameters. Peak period (T_p) was systematically underestimated, especially during storm events. In 2023, RMSE values for T_p were 1.8 s for HurryWave, compared to 1.4 s for ERA5, and during the Sinterklaas storm underestimations reached up to 2–3 s at offshore platforms. This trend is not unique to HurryWave: ERA5 itself is known to underestimate wave energy and misrepresent high periods (Afolabi et al., 2024). Still, the persistence of this underestimation across time scales suggests that HurryWave's simplified treatment of wind input and nonlinear interactions trades accuracy for computational speed. This underestimation can result from several factors. Firstly, the wind data may not be accurate, in this case the WAM model used in ERA5 would also underestimate the wave period as detailed further in 5.2. Another reason for the underestimation could come from the way the energy is transferred from wind to waves. Currently this is done using the wind drag formulation by Zijlema et al. (2012), which is a simple way of transferring energy. Other methods could be investigated to see if the underestimation persists. Lastly, ignored nonlinear effects can lead to underestimation of wave periods as shown in section 5.4.

Mean wave direction emerged as the weakest-performing parameter, with mean absolute errors of 22–28° at several stations and scatter indices exceeding 0.25. This aligns with previous findings that wave direction is among the hardest parameters to model in the North Sea due to its sensitivity to local wind variability and spectral resolution (van Nieuwkoop & Gautier, 2015). The wave direction issues lies mostly in the moments the waves turn as shown in figure 4.24. These turns occur either too quickly or too slowly. In many cases the ERA5 wave simulations also show this behavior as the wind data is inaccurate. However, larger inaccuracies occur in the HurryWave model alone as shown further in section 5.2. The differences in wave direction could also stem from the inaccuracies in modeling the wave period as described above which would change the speed at which waves arrive. This could lead to delays in changing direction leading to large inaccuracies in the wave direction. Lastly, ignored nonlinear effects can lead to delay in transferring wave energy accurately leading to later direction changes as shown in section 5.4.

When tested against extremes, such as the Sinterklaas storm, HurryWave captured the overall storm development and timing but underestimated both wave height (by up to 0.7 m) and peak period (by up to 2 s) at offshore stations. This under performance is due to the absence of boundary conditions limiting incoming wave energy from the Atlantic and Arctic oceans as shown further in section 5.3.

Finally, analysis of the wave spectra revealed concentrated energy distributions compared to buoy observations, which showed broader spreads in both frequency and direction. This indicates that HurryWave may misrepresent multimodal seas, an issue particularly relevant for the North Sea, where wind-sea and swell often coexist (Portilla-Yandún et al., 2015). The use of mean period as a comparison parameter instead of the peak period shows HurryWave is better at estimating the general sea state rather than the peak frequency. The mean periods of the first and second order spectral moments (T_{m01} and T_{m02}) show a better agreement with the spectral buoy data over the Sinterklaas storm in December 2013. The negative first order spectral moment (T_{m-10}) was shown to have the strongest agreement between model and data. This parameter emphasizes long long period swells which shows HurryWave in this case performs better in modeling longer waves during this storm. The analysis of these parameters can be extended to longer periods to better understand the performance of HurryWave in modeling different wave parameters.

Taken together, these results show that HurryWave is robust for long-term wave height climatology, however caution is needed for applications requiring accurate storm extremes, spectral details, or wave direction.

5.2. Validity of the ERA5 wind input data

An explanation to the differences observed between the wave buoy measurements and the HurryWave model could be explained through the inaccurate climate wind data from ERA5. This is true for both wind speed and direction, as those are the primary drivers for wave height, period and direction. A study was conducted on the wind data comparing measured wind speed and direction obtained from the Royal Dutch meteorological institute (KNMI) at the stations shown in table 3.2 (KNMI, n.d.). A scatter plot and a QQ plot are shown in figure 5.1 where the modeled ERA5 wind speed is compared to the measured wind by KNMI at Euro platform during the Sinterklaas storm week (1-8 December 2013).

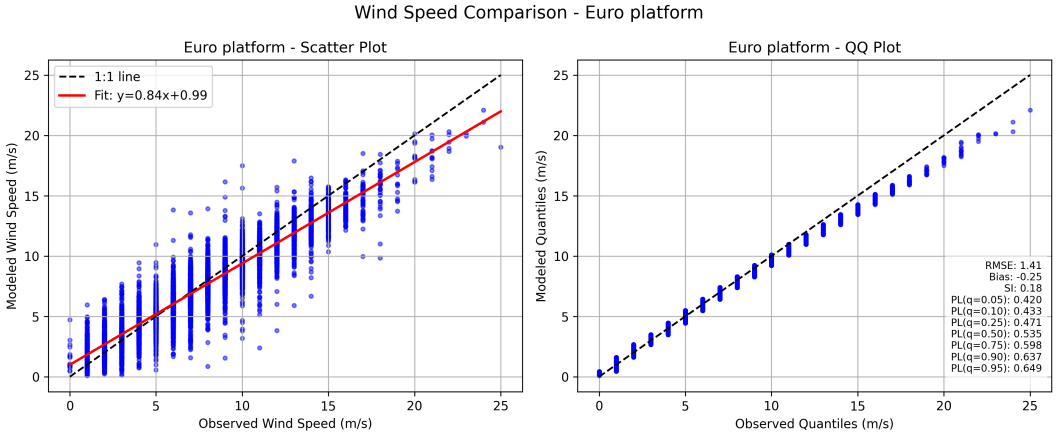

Figure 5.1: Scatter and quantile–quantile comparison of ERA5-modeled and KNMI-measured wind speeds at Euro Platform during the Sinterklaas storm (1–8 December 2013). The results show that ERA5 systematically underestimates high wind speeds, especially during storm peaks, while maintaining good agreement at lower magnitudes. This underestimation directly influences wave model performance, contributing to the underprediction of significant wave height and peak period observed in both HurryWave and ERA5-derived WAM simulations.

This figure clearly shows the underestimation of large wind speeds, especially during the storm. The KNMI measurements have a resolution of 1 m/s, which explains the vertical lines, however this is enough to observe the underestimation at high quantiles. These results align with the results found in Gandoin and Garza (2024) which highlights underestimation of large wind speeds in ERA5. This is one of the root causes in the underestimation of the wave height. Larger wind speeds generate larger waves which in this case are underestimated due to incorrect input. Another flaw in the HurryWave model in the current setup is the inaccurate representation of wave direction, especially in moments where waves change direction. Figure 4.38 and 4.39 clearly show this slow turning in the spectrum. This is also reflected when testing on a single storm in figure 4.24 where HurryWave has its waves change direction at an earlier time than shown in the buoy measurements. The large angular RMSE is a result of these

differences, being larger than other previous hindcast studies shown in section 2.5. The error in the wind direction modeling by ERA5 could lead to these inaccuracies. The error between the modeled and measured wind at Euro Platform for 2023 is shown in figure 5.2, with a histogram shown in figure 5.3.

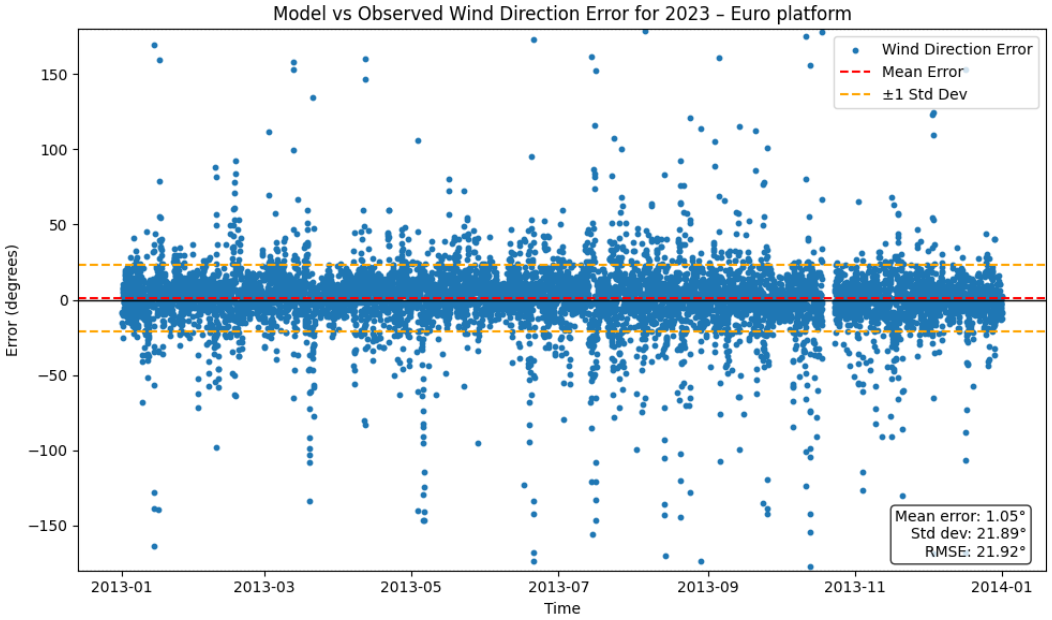

Figure 5.2: Time series comparison of ERA5-modeled and KNMI-measured wind direction at Euro Platform for 2023. The large deviations during wind shifts demonstrate that ERA5 fails to reproduce rapid directional changes, particularly during storm events. Such directional inaccuracies propagate into the wave simulations, causing phase lags and misaligned turning angles in modeled wave direction.

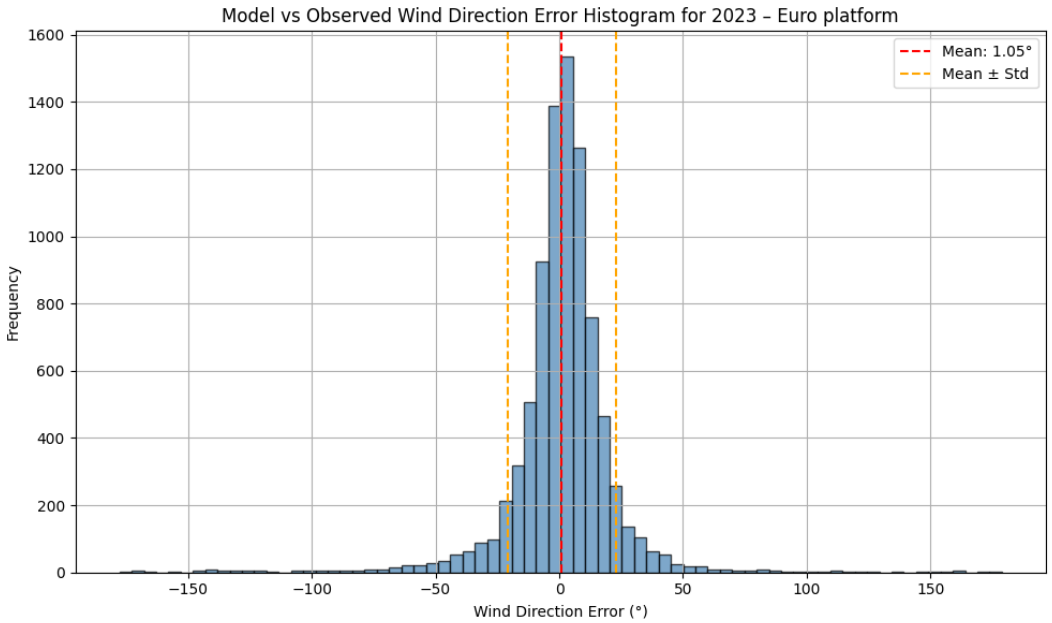

Figure 5.3: Histogram of wind direction error between ERA5-modeled and KNMI-measured winds at Euro Platform for 2023. The error distribution shows a standard deviation exceeding 20°, consistent with the angular RMSE of 21.9°. This confirms that ERA5 wind direction contains significant uncertainty, reducing the reliability of derived wave direction simulations and contributing to the directional discrepancies observed in HurryWave outputs.

These show large differences, especially given the angular RMSE of 21.92 degrees. This means the wind data lacks quality compared to the measurements and is not suitable at this moment for precise studies. This reflects the importance of accurate input data when testing a model against observations.

The values used as input of u_{10} and v_{10} being the x and y wind speed at 10m elevation respectively, could also be a source of uncertainty. ERA5 calculates the wind stress on the boundary between the surface using a coupled model between air and water (ECMWF, n.d.-a). From this boundary layer stress, the wind speeds at 10m elevation are calculated. This 10m elevation wind speed is then used as input to HurryWave, where it is then used again to calculate the stress on the boundary layer using a different formula, in this case, a simple drag formulation. This simplification could lead to differences in wind stress, not observable when comparing the 10m elevation wind speeds.

5.3. Validity of absence of boundary conditions

In this study, boundary conditions are not included as explained in section 1.4. This is done to simplify the model, avoiding errors caused by the inaccuracies in the data prescribed at the boundary. The validity of this assumption was however not tested. A test was therefore conducted to explain the inaccuracies observed in the deeper water stations such as North Cormorant. This station largely underestimates wave height which could be caused by the wind input, however this specific stations underestimates the wave height systematically, not only at large quantiles.

An investigation into enlarging the area to find whether the waves coming from outside the North Sea could have an effect is therefore performed, enlarging the area from the one described in section 1.4 to a larger one defined by these values. The area is increased following the table shown in table 5.1 and in figure 5.4. The model is then run on the input of 2023 simulating a full year in an enlarged domain to compare to the yearly simulations made in 2023.

Coordinate	Original domain	Extended domain
North	65	80
West	-12	-25
South	48	45
East	10	10

Table 5.1: Extended model domain coordinates

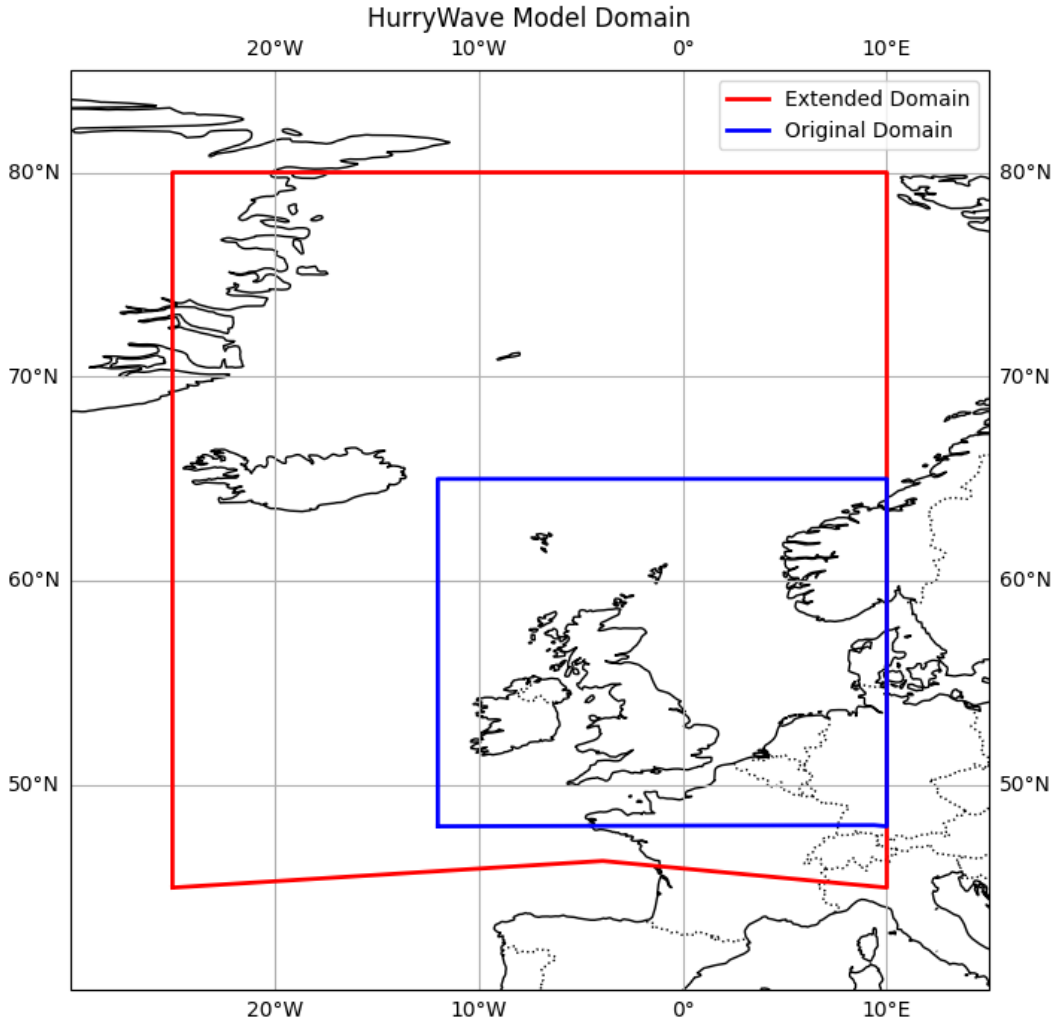

Figure 5.4: Geographic extent of the original (blue) and extended (red) HurryWave domains. The extended domain includes portions of the North Atlantic and Arctic inflow regions, allowing for the entry of swell energy otherwise excluded in the base setup. This configuration was used to test the sensitivity of offshore wave height underestimations to missing boundary input.

The modeled wave parameters match the buoy measurements more accurately, specifically at North Cormorant as shown in figure 5.5.

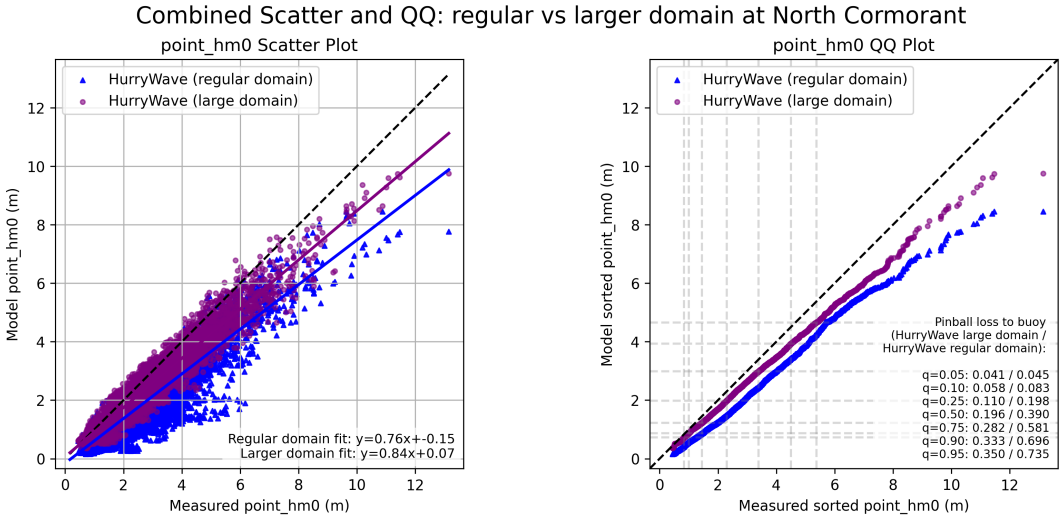

Figure 5.5: Comparison of HurryWave model output and buoy data at North Cormorant during 2023 using the extended model domain. The left panel shows scatter plots of modeled versus observed H_{m0}, while the right panel shows sorted quantile plots. The inclusion of boundary inflow from the Arctic and Atlantic Oceans significantly improves model performance, reducing underestimation of large waves and aligning quantile distributions with buoy observations.

Figure 5.5 shows a significant improvement in modeling wave height, specifically at North Cormorant, compared to 4.15. This proves that penetrating waves from the Atlantic and Arctic, as shown in figure 5.6, have a large effect on the modeled wave height, especially in deeper areas such as North Cormorant.

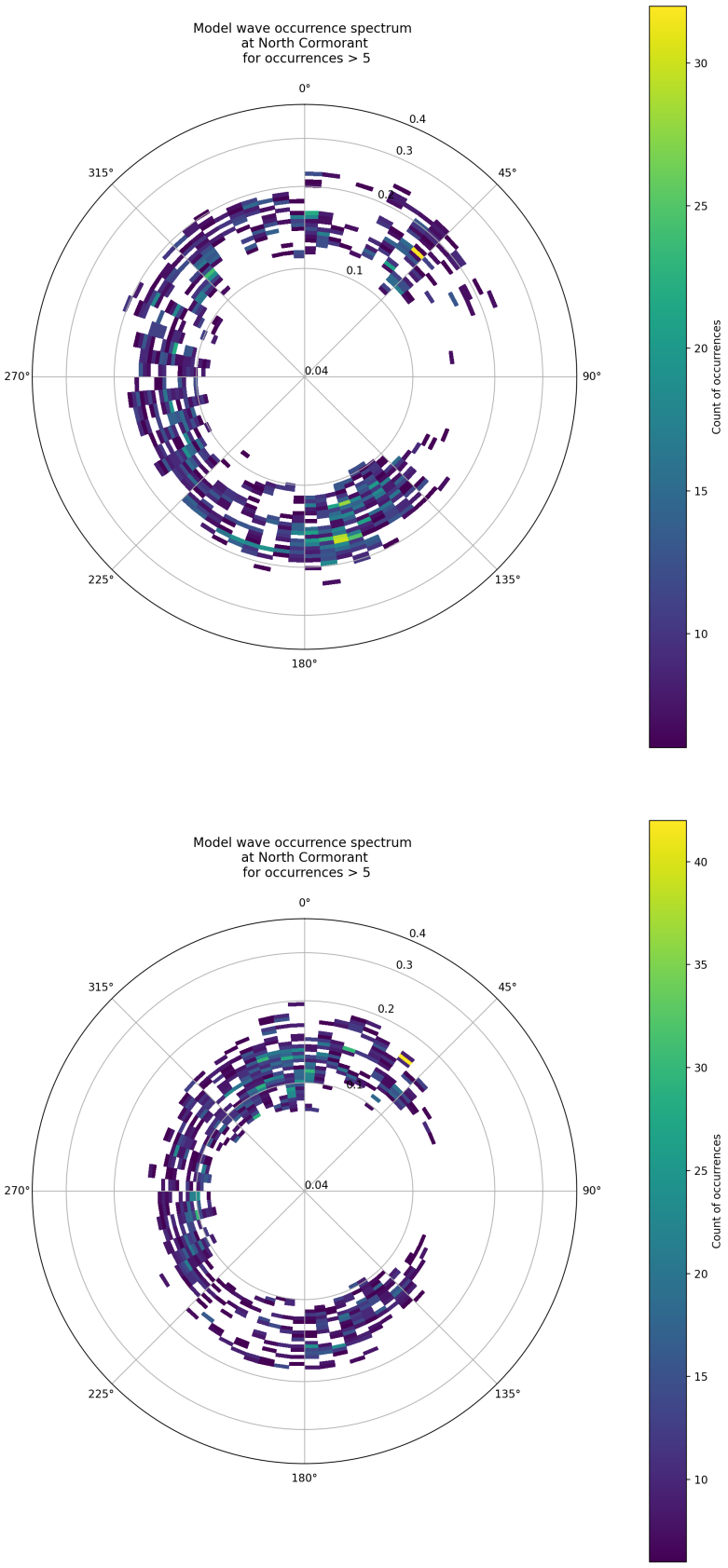

Figure 5.6: Polar plot of the occurrence density of modeled wave events at North Cormorant, as a function of direction and frequency for the regular domain (top) and the extended domain (bottom). A dominant direction emerges from the South East in the regular domain, whereas the extended domain shows an increase of waves coming from the North and the West.

Other stations such as Euro platform, further from the boundaries, do not exhibit much improvement in modeling significant wave height as shown in figure 5.7. Additional comparisons for other stations are shown in Appendix G.

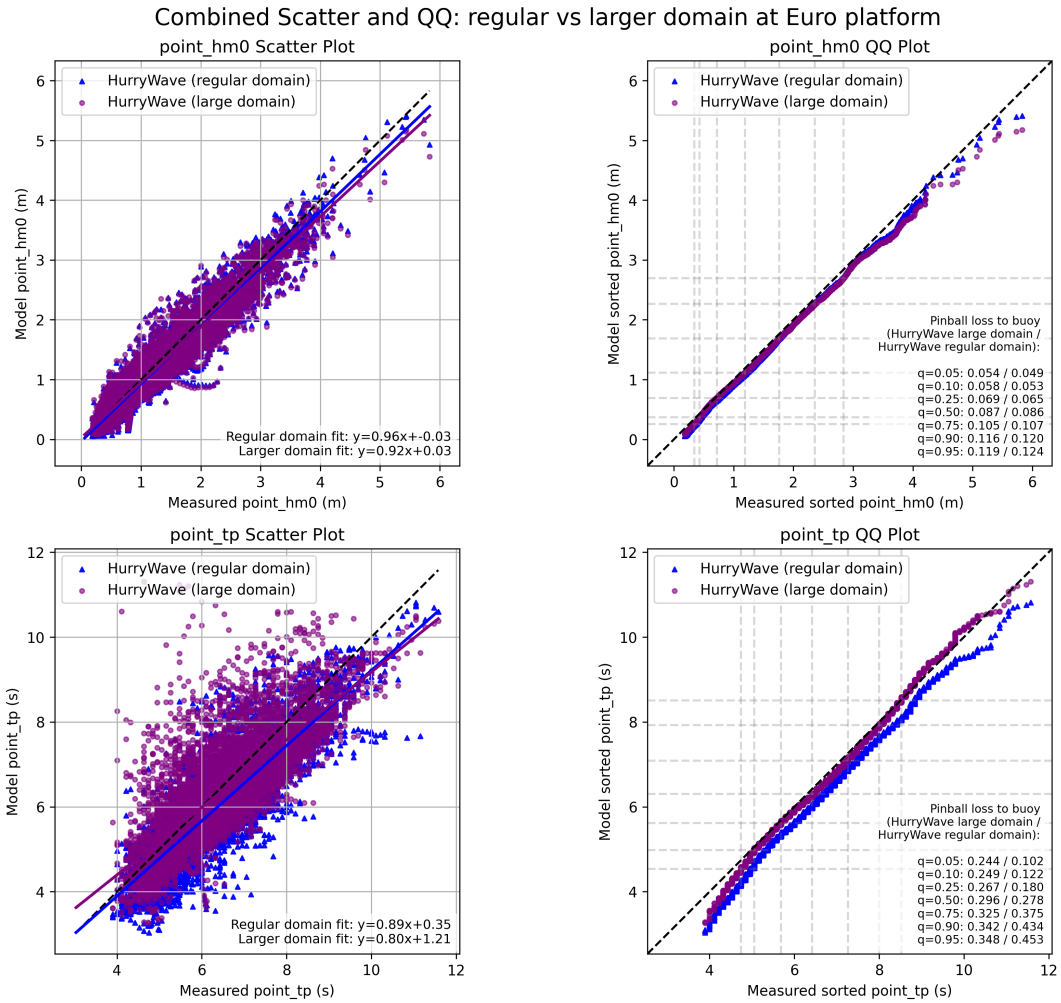

Figure 5.7: Comparison of HurryWave model output and buoy data at Euro platform during the 2023 using the extended model domain. The left panel shows scatter plots of modeled versus observed H_{m0}, while the right panel shows sorted quantile plots. The inclusion of boundary inflow from the Arctic and Atlantic Oceans does not improve the model performance, even reducing the accuracy of points at higher quantiles

Figure 5.7 shows that the assumption made of excluding boundary conditions does not improve results at stations near the Dutch shore such as Euro platform. In stations in deeper water, such as North Cormorant shown in figure 5.5, these boundary conditions are relevant and need to be included to accurately capture waves coming from the North and West.

5.4. Importance of wave interactions

Choosing the appropriate model size and boundaries for the application of the model is therefore crucial, as a too small model size could lead to inaccurate results.

The absence of triad interactions in the HurryWave model can lead to an underestimation of the wave period and a misrepresentation of the wave direction, particularly when the directional distribution is changing. Triad interactions are responsible for transferring energy rapidly toward lower wavenumbers (longer periods) and for coupling energy between different propagation directions. When these interactions are omitted, the simulated spectrum remains too narrow and evolves too slowly toward longer waves, causing the mean and peak wave periods to be underestimated. Moreover, because triads

also generate both bound and free wave components with distinct directional signatures, their absence reduces angular energy spreading and can yield unrealistically coherent wave fields (Zhang & Pal, 2025). This is especially relevant in intermediate water depths, where triad effects strengthen as nonlinearity increases. This is seen in the difference in dispersion between the modeled and measured spectrum shown in figure 4.40, where the spectrum modeled by HurryWave is consistent and shows a much smaller amount of dispersion compared to the measured spectrum.

The importance of triad interactions can be quantified using the Ursell number, defined as

$$\mathrm{Ur} = \frac{H}{k^2 d^3} = \frac{HL^2}{4\pi^2 d^3}, \tag{5.1}$$

where H is the wave height (here taken as the significant wave height), L the wavelength, and d the local water depth.

The Ursell number expresses the relative importance of nonlinear (triad) effects versus dispersion. For $\mathrm{Ur} \ll 1$, waves behave linearly and triad interactions are negligible, whereas for increasing Ur, nonlinear energy transfers through triads become progressively more significant. While the SWAN model treats triad interactions as relevant for $\mathrm{Ur} > 0.1$ (The SWAN team, 2006), recent findings suggest that noticeable nonlinear effects already emerge for $\mathrm{Ur} \gtrsim 0.2$–0.3, and that the system becomes fully nonlinear when $\mathrm{Ur} \gtrsim 1$ (Zhang & Pal, 2025). Consequently, in intermediate-depth conditions, the neglect of triad interactions in HurryWave can lead to an underestimation of long-period energy and an inaccurate representation of evolving wave directionality. The Ursell number is derived using equation 5.1 by using the method to calculate the wave length described in equations 3.3. This is plotted for the buoy data at IJmuiden munitiestortplaats in figure 5.8.

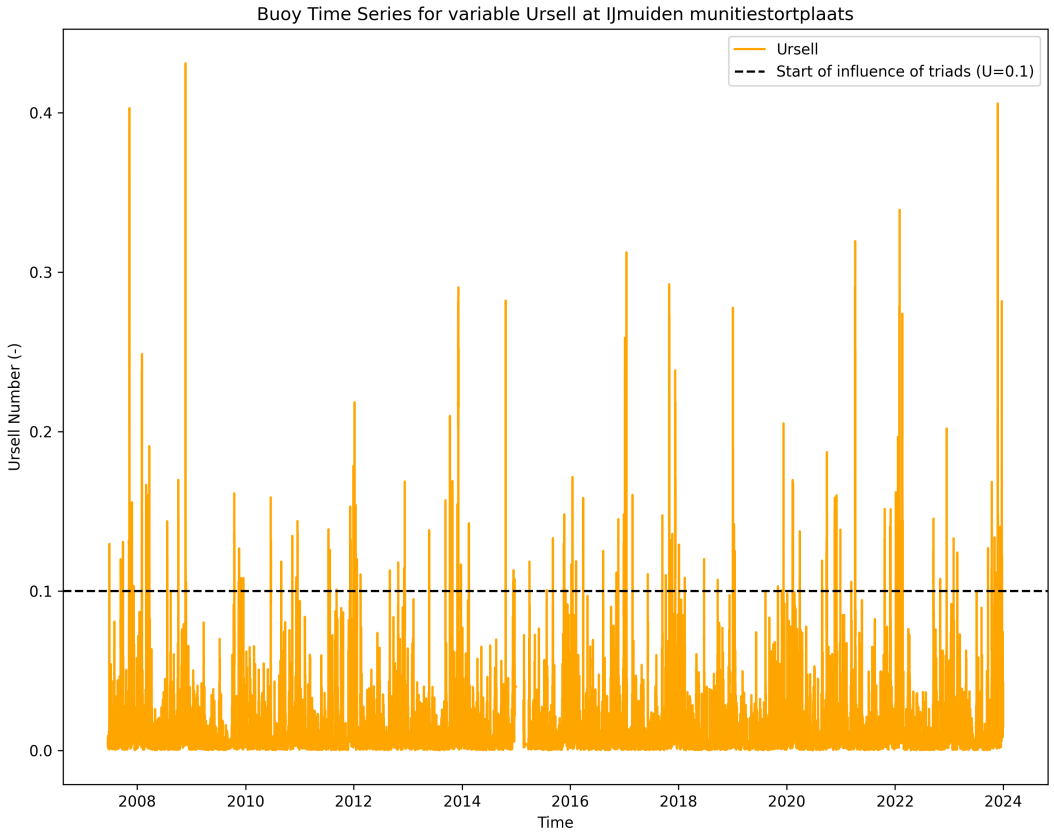

Figure 5.8: Time series of the Ursell number (Ur) computed from buoy data at Euro Platform using equation 5.1. Values above $\mathrm{Ur} \approx 0.2$ indicate a transition toward nonlinear wave behavior, where triad interactions become important. The frequent occurrence of $\mathrm{Ur} > 0.3$ confirms that the local sea state is often nonlinear, implying that neglecting triad interactions in HurryWave leads to underestimation of long-period energy and misrepresentation of wave directionality.

This figure shows that the sea at this location tends to sometimes be in a nonlinear state with high Ursell numbers. This means that triad interactions are occurring and could explain the misrepresentation of the period and some of the differences in wave direction which are not attributed to the wind input. Other measuring stations tend to have lower Ursell numbers as shown in appendix Appendix I, which means triad interaction would have less of an effect in those locations.

Another explanation for the missing performance in modeling wave period would be the assumption made when calculating quadruplet wave interactions. These typically occur in deeper waters, which applies to most of the North Sea. These quadruplet interactions tend to be misrepresented using the DIA model (S. Hasselmann & Hasselmann, 1985) as shown in Benoit (2006). This study shows that the DIA approximation is not suitable in fast changing wind direction. This, along with the error in input wind direction shown in section 5.2 could explain the difficulty of HurryWave in modeling wave direction when this direction changes.

5.5. Limitations of the study

While this study provides valuable first insights, several limitations constrain the interpretation of the results.

Data availability was a primary limitation. The number of buoy stations used was limited, with key gaps in parameters such as peak period at North Cormorant. Furthermore, the availability of $T_{1/3}$ data only in later years restricted the temporal coverage of this analysis. The limited amount of data also hindered the statistical demonstration of model performance. At least 20 years of data are needed for this demonstration which could not be achieved using the period data from buoy measurements. Satellite products, such as radar altimetry or SAR, were not included, limiting the ability to assess spatial patterns of model bias. Previous studies have demonstrated the value of combining buoy and satellite data for wave model validation (van Nieuwkoop et al., 2024), suggesting that future assessments of HurryWave could benefit from such integration.

Data quality was also a limitation with many outliers found and removed as shown in tables 3.4 and 3.5. Some outliers were not removed in this framework which could lead to errors when comparing the model output to the incorrect measured values as shown in section 3.2.

The conversion of the observed parameters, significant wave height (H_s), and period of the 1/3 highest waves ($T_{1/3}$) could also cause some error when comparing to the parameters in the frequency domain H_{m0} and T_p. The conversions assume that $H_{m0} = H_s$ and $T_p = T_{1/3}/0.95$, which is not always the case. This is especially true when looking at significant wave height which can differ by up to 5%, especially in sea states with a broad spectrum and low energy (Teutsch et al., 2020).

This study only treats the model run with the input settings shown in table 3.7 and the setup given in section 3.4. This could lead to differences in results when using other parameters or other setups. As an example, including longer waves by extending the frequency range can lead to vastly different peak period results as shown in subsection 4.2.2. Other large scale assumptions included in section 1.4, such as stationary water levels, were made which could account for the differences between the HurryWave model output and the measurements.

Statistical testing also had constraints. The block bootstrap method used for hypothesis testing depended on sufficiently long and independent data records. Only one station (IJmuiden Munitiestortplaats) passed these requirements, meaning statistical proof of model performance could not be generalized to the entire North Sea. This is a common challenge in wave model validation, where long, continuous, and high-quality datasets are rare (Lavidas & Polinder, 2019).

Finally, spectral validation was restricted to a short storm period. While revealing important differences, the analysis was too limited to draw conclusions about long-term performance in multimodal seas.

5.6. Future work

Building on these findings, several avenues of future research can strengthen and validate HurryWave further.

Wave Partitioning

Partitioning wave spectra into wind-sea and swell components would allow for a more detailed assessment of how HurryWave performs under multimodal conditions. This would help identify whether biases stem primarily from wind-sea misrepresentation or swell propagation. A wave splitting methodology such as suggested in Portilla-Yandún et al. (2015) could help gain more understanding into model performance for specific wave families.

Extended Spectral Analysis

Rather than analyzing only short storm periods, spectral validation should be performed over multiple seasons or decades. This would allow to evaluate the HurryWave capability to capture the variability of spectral parameters over climatological timescales, and to evaluate whether underestimations of peak period persist consistently.

Extended wave parameter analysis

Instead of looking at only 3 integrated parameters, future efforts could focus on other parameters such as directional spreading. The outcome of this analysis could help explain some of the results observed in this study.

Inclusion of Boundary Conditions

Running HurryWave with Atlantic and Arctic boundary forcing would test whether offshore underestimations are due to missing swell input. This is particularly relevant for the North Sea, where swell penetration from the North Atlantic is well documented (Lavidas & Polinder, 2019).

Expanded Observational Datasets

Incorporating satellite altimetry, SAR, and newer buoy datasets would provide better spatial coverage. Assimilation experiments could further test how observational integration affects performance.

Application in Climate Scenarios

Given its computational efficiency, HurryWave could be applied to long-term climate simulations such as SEAS5. Future work should assess whether biases scale with changing wind climates and whether it can provide robust input for coastal impact studies.

6
Conclusion

This thesis set out to evaluate the performance of the newly developed HurryWave model in the North Sea and to determine whether its computational efficiency comes at the expense of accuracy. Against this backdrop, the study also considered the broader policy shift in the Netherlands toward reliance on SEAS5 datasets for flood safety assessments, which emphasizes the need for models that can deliver reliable results over long time horizons while remaining computationally feasible.

The research was guided by the following main question:

> **How does HurryWave compare to historical modeled and measured data, in terms of accuracy in modeling wave parameters in the North Sea?**

This overarching question was further broken down into four sub-questions:

1. What are the key differences in methodology and assumptions between the HurryWave model and other wave models?
2. How does HurryWave perform against measured data and the ERA5 dataset in modeling wave height, period, and direction in the North Sea?
3. How does HurryWave perform against measured data in modeling the wave spectra in the North Sea?
4. What are the advantages and limitations to the use of the HurryWave model in the North Sea?

Returning to the research question, the study confirms that HurryWave can provide a viable alternative to established models like SWAN and WAM in the North Sea, but only under certain conditions.

In relation to the first sub-question, this study showed that the HurryWave model distinguishes itself from established models such as SWAN and WAM through its prioritization of computational efficiency over physical detail. Its explicit first-order upwind scheme enables rapid simulations, but the lack of processes such as triad interactions and the simplification of quadruplets constrains its realism. These simplifications lead to an under representation of nonlinear energy transfer, which are essential for accurately reproducing wave conditions in the North Sea.

Addressing the second sub-question, HurryWave demonstrated strong skill in reproducing significant wave height across the North Sea, with biases and scatter index values often outperforming ERA5. This indicates that the model's representation of wind-wave growth and energy dissipation is robust for average conditions. However, large wave heights are underestimated, which would have a consequence when analyzing extremes. This is mainly due to the general underestimation of high wind speeds in the ERA5 input compared to KNMI observations. Peak period was also systematically underestimated, and mean wave direction performance showed variability when the mean wave direction was changing. The underestimation of the peak period is likely driven by the reduced transfer of energy to lower frequencies due to the absence of triads and the simplification of quadruplets. Directional discrepancies can also be attributed to the inaccuracies in wind direction, and the slow transfer of energy due to the absence of triads. Next to these observations, the block bootstrap methodology successfully demonstrated that HurryWave achieves statistically significant skill in simulating significant wave heights at IJmuiden munitiestortplaats, however, the conclusions are spatially and temporally constrained with lack of data, together with remaining trends and seasonality denying the ability to demonstrate performance at other stations or for other parameters.

For the third sub-question, the spectral analysis revealed that HurryWave tends to concentrate energy too narrowly and under represent multimodal seas. This spectral narrowing reflects the lack of nonlinear triad interactions, which redistribute energy between frequency bands. As a result, multimodal wave conditions common in the North Sea, especially during storms, are not adequately captured, limiting the model's spectral realism. The mean period extracted from the spectrum showed that HurryWave performed better at modeling all mean periods compared to peak periods, showcasing its ability to represent overall sea states over the peak frequency. The negative first order mean period performed best with very low root mean square error values, which shows HurryWave tends to better represent lower frequency waves over fast changing wind waves during a storm.

With respect to the fourth sub-question, the main advantage of HurryWave is its very fast computational speed, which makes it highly attractive for long-term climatological applications and large-scale reanalysis. Its limitations lie in the underestimation of extremes, directional accuracy, and spectral realism, meaning it should complement rather than replace existing models in operational or safety-critical contexts.

Taken together, these findings indicate that while HurryWave is a strong candidate for applications that require efficient and large-scale wave height simulations, caution is warranted for operational forecasting or safety assessments that demand accurate storm extremes, wave direction, or detailed spectral properties. The objectives of the research have been addressed: the model's theoretical underpinnings were assessed, its performance compared against historical data, and its ability to represent wave spectra tested through spectral analysis. Each objective contributed to a clearer understanding of the model's strengths and trade-offs when compared to other wave models.

The findings must also be seen in light of the shifting policy context. The move toward SEAS5-based assessments reflects a demand for consistency, long time series, and computational efficiency in coastal risk management. HurryWave, with its speed advantage, is well positioned to complement this policy transition, provided its limitations are carefully accounted for. Its role may therefore be most impactful in long-term climate applications aligned with BOI-2023, while for storm-scale safety assessments, additional validation and model refinement remain necessary.

HurryWave advances the possibilities for efficient wave modeling in the North Sea, but it should be deployed selectively, balancing its strengths in modeling significant wave height with recognition of its current limitations in storm extremes. Continued development should focus on incorporating boundary conditions, improving spectral energy redistribution, and refining the wind input to enhance model realism.

Bibliography

Afolabi, L. A., Russo, S., Lo Re, C., Ludeno, G., Nardone, G., Vicinanza, D., & Contestabile, P. (2024). Underestimation of wave energy from era5 datasets: Back analysis and calibration in the central tyrrhenian sea. *Energies*, *18*(1). https://doi.org/10.3390/en18010003

Barnard, P. L., Erikson, L. H., Foxgrover, A. C., Hart, J. A. F., Limber, P., O'Neill, A. C., van Ormondt, M., Vitousek, S., Wood, N., Hayden, M. K., & Jones, J. M. (2019). Dynamic flood modeling essential to assess the coastal impacts of climate change. *Scientific Reports*, *9*, 4309. https://doi.org/10.1038/s41598-019-40742-z

Battjes, J., & Janssen, J. (1978). Energy loss and set-up due to breaking random waves. https://doi.org/10.9753/icce.v16

Becker, J. M., & Roggenbuck, O. (2023). Prediction of Significant Wave Heights with Engineered Features from GNSS Reflectometry [Publisher: Multidisciplinary Digital Publishing Institute]. *Remote Sensing*, *15*(3), 822. https://doi.org/10.3390/rs15030822

Bell, B., Hersbach, H., Simmons, A., Berrisford, P., Dahlgren, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Radu, R., Schepers, D., Soci, C., Villaume, S., Bidlot, J.-R., Haimberger, L., Woollen, J., Buontempo, C., & Thépaut, J.-N. (2021). The era5 global reanalysis: Preliminary extension to 1950. *Quarterly Journal of the Royal Meteorological Society*, *147*(741), 4186–4227. https://doi.org/10.1002/qj.4174

Benoit, M. (2006). Implementation and test of improved methods for evaluation of nonlinear quadruplet interactions in a third generation wave model, 526–538. https://doi.org/10.1142/9789812709554_0046

Bidlot, J. (2020, February). Ocean wave model output parameters. https://confluence.ecmwf.int/download/attachments/59774192/wave_parameters.pdf

Bosboom, J., & Stive, M. J. (2023). *Coastal dynamics* [DOI 10.5074/T.2021.001]. TU Delf Open.

Bretherton, C. S., Widmann, M., Dymnikov, V. P., Wallace, J. M., & Bladé, I. (1999). The Effective Number of Spatial Degrees of Freedom of a Time-Varying Field [Section: Journal of Climate]. Retrieved August 18, 2025, from https://journals.ametsoc.org/view/journals/clim/12/7/1520-0442_1999_012_1990_tenosd_2.0.co_2.xml

Cavaleri, L., & Rizzoli, P. M. (1981). Wind wave prediction in shallow water: Theory and applications. https://doi.org/10.1029/JC086iC11p10961

Chandy, M., Schifano, E., & Yan, J. (2025). On Sample Size Needed for Block Bootstrap Confidence Intervals to Have Desired Coverage Rates. *American Journal of Undergraduate Research*, *20*(4), 3–16. https://doi.org/10.33697/ajur.2024.101

Curto, V. (2014). Reshaping of beach nourishments under high angle of wave incidence. Retrieved March 25, 2025, from https://repository.tudelft.nl/record/uuid:5c26b066-8fe7-4440-9641-d5c8946cac01

de Valk, C., & van den Brink, H. (2023). Update van de statistiek van extreme zeewaterstand en wind op basis van meetgegevens en modelsimulaties [KNMI number: TR-406]. https://www.knmi.nl/kennis-en-datacentrum/publicatie/update-van-de-statistiek-van-extreme-zeewaterstand-en-wind-op-basis-van-meetgegevens-en-modelsimulaties

Deltares. (n.d.). Deltares guiding principles. https://cms.deltares.nl/assets/common/downloads/Deltares-Guiding-Principles-1.pdf

Destination Earth. (n.d.). Disaster Risk Mitigation & Climate Adaptation. Retrieved February 16, 2025, from https://destination-earth.eu/use-cases/disaster-risk-mitigation-climate-adaptation/

ECMWF. (n.d.-a). Era5: Data documentation - copernicus knowledge base - ecmwf confluence wiki. Retrieved March 16, 2025, from https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation

ECMWF. (n.d.-b). The family of era5 datasets - copernicus knowledge base - ecmwf confluence wiki. Retrieved March 16, 2025, from https://confluence.ecmwf.int/display/CKB/The+family+of+ERA5+datasets

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife [Publisher: Institute of Mathematical Statistics]. *The Annals of Statistics*, *7*(1), 1–26. https://doi.org/10.1214/aos/1176344552

Eldeberky, Y. (1996, October). *Nonlinear Transformation of Wave Spectra in the Nearshore Zone* [Doctoral dissertation, Delft University of Technology]. https://resolver.tudelft.nl/uuid:707ca57d-81c3-4103-bc6e-aae1c90fce63

Elshinnawy, A. I., & Antolínez, J. A. Á. (2023). A changing wave climate in the Mediterranean Sea during 58-years using UERRA-MESCAN-SURFEX high-resolution wind fields. *Ocean Engineering*, *271*, 113689. https://doi.org/10.1016/j.oceaneng.2023.113689

Gandoin, R., & Garza, J. (2024). Underestimation of strong wind speeds offshore in ERA5: Evidence, discussion and correction [Publisher: Copernicus GmbH]. *Wind Energy Science*, *9*(8), 1727–1745. https://doi.org/10.5194/wes-9-1727-2024

Goda, Y. (2010). REANALYSIS OF REGULAR AND RANDOM BREAKING WAVE STATISTICS. *Coastal Engineering Journal*, *52*(01), 71–106. https://doi.org/10.1142/S0578563410002129

Groll, N., & Weisse, R. (2017). A multi-decadal wind-wave hindcast for the North Sea 1949–2014: coastDat2 [Publisher: Copernicus GmbH]. *Earth System Science Data*, *9*(2), 955–968. https://doi.org/10.5194/essd-9-955-2017

Hasselmann, K., Barnett, T., Bouws, E., Carlson, H., Cartwright, D., Enke, K., Ewing, J., Gienapp, H., Hasselmann, D., Kruseman, P., Meerbrug, A., Müller, P., Olbers, D., Richter, K., Sell, W., & Walden, H. (1973). Measurements of wind-wave Growth and swell decay during the joint north sea wave project (jonswap). *Erganzungsheft zur Deutschen Hydrographischen Zeitschrift Reihe A*, *12*.

Hasselmann, S., & Hasselmann, K. (1985). Computations and parameterizations of the nonlinear energy transfer in a gravity-wave spectrum. part i: A new method for efficient computations of the exact nonlinear transfer integral. *Journal of Physical Oceanography*. Retrieved March 15, 2025, from https://journals.ametsoc.org/view/journals/phoc/15/11/1520-0485_1985_015_1369_capotn_2_0_co_2.xml

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., . . . Thépaut, J.-N. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Jesper H. Andersen, Andy Stock, Stefan Heinänen, Miia Mannerla, & Morten Vinther. (2013, March). *Human uses, pressures and impacts in the eastern North Sea* (Technical report No. 18). Aarhus University, Department of Bioscience, DHI, Denmark, HELCOM Secretariat, National Institute of Aquatic Resources, Technical University of Denmark. http://dce.au.dk/en

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D., Weisheimer, A., Balsamo, G., Keeley, S. P. E., Mogensen, K., Zuo, H., & Monge-Sanz, B. M. (2019). SEAS5: The new ECMWF seasonal forecast system [Publisher: Copernicus GmbH]. *Geoscientific Model Development*, *12*(3), 1087–1117. https://doi.org/10.5194/gmd-12-1087-2019

Kim, S., Mori, N., Mase, H., & Yasuda, T. (2015). The role of sea surface drag in a coupled surge and wave model for Typhoon Haiyan 2013. *Ocean Modelling*, *96*, 65–84. https://doi.org/10.1016/j.ocemod.2015.06.004

KNMI. (n.d.). KNMI - Uurgegevens van Noordzee stations. Retrieved October 19, 2025, from https://www.knmi.nl/nederland-nu/klimatologie/uurgegevens_Noordzee

KNMI - About KNMI. (n.d.). Retrieved March 21, 2025, from https://www.knmi.nl/over-het-knmi/about

Komen, G. J., Hasselmann, S., & Hasselmann, K. (1984). On the existence of a fully developed wind-sea spectrum. *Journal of Physical Oceanography*, 1271–1285. Retrieved March 15, 2025, from https://journals.ametsoc.org/view/journals/phoc/14/8/1520-0485_1984_014_1271_oteoaf_2_0_co_2.xml

Kunsch, H. R. (1989). The Jackknife and the Bootstrap for General Stationary Observations [Publisher: Institute of Mathematical Statistics]. *The Annals of Statistics*, *17*(3), 1217–1241. https://doi.org/10.1214/aos/1176347265

Lavidas, G., & Alday, M. (2022). *Wave energy assessment | north sea wave database* (tech. rep.) (ISBN 978-94-6366-628-2). TU Delft. Retrieved March 16, 2025, from https://repository.tudelft.nl/record/uuid:c9e92528-d0dd-4b6d-a354-e0f117560e8e

Lavidas, G., & Polinder, H. (2019). North sea wave database (nswd) and the need for reliable resource data: A 38 year database for metocean and wave energy assessments [Number: 9 Publisher:

Multidisciplinary Digital Publishing Institute]. *Atmosphere*, *10*(9), 551. https://doi.org/10.3390/atmos10090551

Leo. H. Holthuijsen. (2007). *WAVES IN OCEANIC AND COASTAL WATERS*. Cambridge University Press.

Ministerie van Infrastructuur en Waterstaat. (n.d.). About rijkswaterstaat. Retrieved March 21, 2025, from https://www.rijkswaterstaat.nl/en/about-us

Pegel, H. (2014, April). Samenwerken en gedragscode publiek opdrachtgeverschap [Documentnummer: 758664]. https://open.rijkswaterstaat.nl/open-overheid/onderzoeksrapporten/@48950/samenwerken-gedragscode-publiek/

Ponce de Leon, S., Bettencourt, J., van Vledder, G., & Doohan, P. (2018). Performance of WAVEWATCH-III and SWAN Models in the North Sea. https://doi.org/10.1115/OMAE2018-77291

Portilla-Yandún, J., Cavaleri, L., & Van Vledder, G. P. (2015). Wave spectra partitioning and long term statistical distribution. *Ocean Modelling*, *96*, 148–160. https://doi.org/10.1016/j.ocemod.2015.06.008

Reistad, M., Breivik, Ø., Haakenstad, H., Aarnes, O. J., Furevik, B. R., & Bidlot, J.-R. (2011). A high-resolution hindcast of wind and waves for the North Sea, the Norwegian Sea, and the Barents Sea - Reistad - 2011 - Journal of Geophysical Research: Oceans - Wiley Online Library. *Journal of Geophysical Research: Oceans*, *116*(C5). https://doi.org/10.1029/2010JC006402

Rijnsdorp, D. P., Reniers, A. J. H. M., & Zijlema, M. (2021). Free infragravity waves in the north sea [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2021JC017368]. *Journal of Geophysical Research: Oceans*, *126*(8), e2021JC017368. https://doi.org/10.1029/2021JC017368 e2021JC017368 2021JC017368.

Ris, R. C., Holthuijsen, L. H., & Booij, N. (1999). A third-generation wave model for coastal regions: 2: Verification. *Journal of Geophysical Research*, *104*(C4), 7667–768. https://doi.org/10.1029/1998JC900123

Ritter, A., & Muñoz-Carpena, R. (2013). Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments. *Journal of Hydrology*, *480*, 33–45. https://doi.org/10.1016/j.jhydrol.2012.12.004

Roseser, S., & Copeland, S. (2020). Tu delft code of conduct. https://doi.org/10.4233/uuid:704e72b8-6b14-4cf1-a931-9c0f93c50152

Scholten, M. (2024, November). Evaluation Machine Learning Model of SWAN-Kuststrook. https://www.deltares.nl/expertise/publicaties/evaluation-machine-learning-model-of-swan-kuststrook-memo

Settelmaier, J. B., Gibbs, A., Santos, P., Freeman, T., & Gaer, D. (2011). SIMULATING WAVES NEARSHORE (SWAN) MODELING EFFORTS AT THE NATIONAL WEATHER SERVICE (NWS) SOUTHERN REGION (SR) COASTAL WEATHER FORECAST OFFICES (WFOs). https://ams.confex.com/ams/91Annual/webprogram/Manuscript/Paper183540/Settelmaier_13A.4.pdf

System details. (n.d.). Retrieved October 5, 2025, from https://servicedesk.surf.nl/wiki/spaces/WIKI/pages/92668136/System+details

Teutsch, I., Weisse, R., Moeller, J., & Krueger, O. (2020). A statistical analysis of rogue waves in the southern North Sea [Publisher: Copernicus GmbH]. *Natural Hazards and Earth System Sciences*, *20*(10), 2665–2680. https://doi.org/10.5194/nhess-20-2665-2020

The SWAN team. (2006). SWAN USER MANUAL. https://falk.ucsd.edu/modeling/swanuse.pdf

The WAVEWATCH III R Development Group (WW3DG). (2016). User manual and system documentation of WAVEWATCH III version 5.16 [Tech. Note 329]. https://polar.ncep.noaa.gov/waves/wavewatch/manual.v5.16.pdf

van Asselt, K. (2025, February). HW Documentation.

van Nieuwkoop, J., & Gautier, C. (2015, November). Validation of SWAN wave directions. https://open.rijkswaterstaat.nl/open-overheid/onderzoeksrapporten/@154823/validation-swan-wave-directions/

van Nieuwkoop, J., Madelief, D., Eleveld, M., de Korte, E., & Gawehn, M. (2024, December). *Technology scan wave information from satellites* (tech. rep. No. 11210320-021-BGS-0001). Deltares. https://publications.deltares.nl/11210320_021_0001.pdf

van Vledder, G. P. (2001). *Improved algorithms for computing the non-linear quadruplet wave-wave interactions in deep and shallow water* (tech. rep.). Alkyon Hydraulic Consultancy & Research. https://www.

ecmwf.int/sites/default/files/elibrary/2001/13251-improved-algorithms-computing-non-linear-quadruplet-wave-wave-interactions-deep-and-shallow.pdf

van der Lugt, M., van Asselt, K., van Ormondt, M., & van Dongeren, A. (2024, August). *Hurrywave - an instationary wave model on ocean scales* (Validation report No. 11208970-002-ZKS-0001). Deltares.

Wageningen University & Research. (2021, February). North Sea. Retrieved February 16, 2025, from https://www.wur.nl/en/research-results/dossiers/file/north-sea-1.htm

Wave measuring buoy | Digital North Sea. (n.d.). Retrieved March 16, 2025, from https://www.digitalnorthsea.nl/data/hydrological-data/wave-measuring-buoy

Wavespectra — wavespectra 4.4.1 documentation. (n.d.). Retrieved August 18, 2025, from https://wavespectra.readthedocs.io/en/latest/

Weisse, R., von Storch, H., Niemeyer, H. D., & Knaack, H. (2012). Changing north sea storm surge climate: An increasing hazard? *Ocean & Coastal Management*, *68*, 58–68. https://doi.org/10.1016/j.ocecoaman.2011.09.005

Wolf, J., Hargreaves, J. C., & Flather, R. A. (2000, December). *Application of the SWAN shallow water wave model to some UK coastal sites* (Publication - Report) (Num Pages: 51). Proudman Oceanographic Laboratory. Birkenhead. Retrieved September 27, 2025, from https://nora.nerc.ac.uk/id/eprint/3917/

Wróbel-Niedźwiecka, I., Drozdowska, V., & Piskozub, J. (2019). Effect of drag coefficient formula choice on wind stress climatology in the North Atlantic and the European Arctic. *Oceanologia*, *61*(3), 291–299. https://doi.org/10.1016/j.oceano.2019.02.002

Wu, J. (1982). Wind-stress coefficients over sea surface from breeze to hurricane. *Journal of Geophysical Research: Oceans*, *87*(C12), 9704–9706. https://doi.org/10.1029/JC087iC12p09704

Zhang, Z., & Pal, Y. (2025). Role of triad interactions in spectral evolution of surface gravity waves in deep water. *Journal of Fluid Mechanics*, *1008*, A41. https://doi.org/10.1017/jfm.2025.234

Ziegler, P. (1975). The Geological Evolution of the North Sea Area in the Tectonic Framework of North Western Europe. *Norges geol. Unders.*, *316*, 1–27. Retrieved February 16, 2025, from https://www.ngu.no/filearchive/NGUPublikasjoner/NGUnr_316_Bulletin_29_Ziegler_1_27.pdf

Zijl, F., & Groenenboom, J. (2019). Development of a sixth generation model for the NW European Shelf (DCSM-FM 0.5nm). https://publications.deltares.nl/11203715_004.pdf

Zijlema, M., van Vledder, G. P., & Holthuijsen, L. H. (2012). Bottom friction and wind drag for wave models. *Coastal Engineering*, *65*, 19–26. https://doi.org/10.1016/j.coastaleng.2012.03.002

Appendices

Appendix A
Github repository and contact

The code used to extract data and process HurryWave is stored on the author's Github repository at: https://github.com/DanielvanderHoorn/cht_hurrywave

Appendix B
Description of input parameters

- `mmax`, `nmax` (Number of grid cells in x- and y-direction)
  Defines the total number of grid cells in the domain. More cells result in finer spatial resolution and better representation of bathymetry and wind fields, but increase computation time approximately proportional to $mmax \times nmax$ per timestep (The SWAN team, 2006).
- `dx`, `dy` (Grid spacing)
  Define the physical grid resolution. Finer grids resolve coastal and bathymetric gradients better but may require smaller timesteps to maintain numerical stability (CFL condition), leading to higher runtime (The SWAN team, 2006).
- `dt` (Time step)
  Controls the model integration time step. Larger dt values speed up computations but risk numerical instability if CFL > 1. Choosing dt close to the maximum stable value optimizes performance (The SWAN team, 2006).
- `tspinup` (Spin-up time)
  Allows wave spectrum to reach equilibrium before the analysis period. Longer spin-up yields more realistic initial conditions but extends total runtime (Settelmaier et al., 2011).
- `dtmapout`, `dthisout`, `dtsp2out` (Output intervals)
  Determine the frequency of map and point outputs. Shorter intervals provide better temporal resolution but increase file size.
- `dtwnd` (Wind update interval)
  Sets the frequency of wind field updates. Smaller intervals improve accuracy during fast-evolving storms but slightly increase CPU usage due to more frequent interpolation.
- `freqmin`, `freqmax` (Spectral frequency range)
  Specify the resolved wave frequency band. Too narrow a range can miss swell or short waves; too wide increases computation without significant benefit. Default values (0.04–0.5 Hz) match typical deep-water applications as shown in The SWAN team (2006), and in Wolf et al. (2000).
- `nsigma` (Number of frequency bins)
  Controls spectral resolution in frequency space. Increasing $nsigma$ yields a more detailed spectrum but increases runtime linearly (The SWAN team, 2006). It is here chosen to be 12 as shown in (The SWAN team, 2006).
- `ntheta` (Number of directional bins)
  Sets angular resolution of the spectrum. More bins better capture directional spreading but proportionally increase computational effort. $ntheta$ = 36 (10° bins) is a standard compromise (The SWAN team, 2006).
- `gammajsp` (JONSWAP peak enhancement factor)
  Defines sharpness of the spectral peak. Default value 3.3 reproduces North Sea conditions (K. Hasselmann et al., 1973). This value does not have a specific effect on computation time.
- `quadruplets` (Nonlinear four-wave interactions)
  Activates the DIA approximation of nonlinear quadruplet interactions (S. Hasselmann & Hasselmann, 1985). Improves spectral energy transfer between frequencies but can account for a significant proportion of runtime.
- `winddrag`, `cdcap` (Wind input formulation)
  Determine wind-to-wave energy transfer. The Zijlema et al. (2012) formulation is preferred for

extreme winds, avoiding unrealistically high drag coefficients as shown in section 2.2, whereas the formulation by Wu (1982) is simpler but may overestimate drag under storm conditions as shown in (Kim et al., 2015).

- `fbed` (Bed friction coefficient)
  Governs dissipation of wave energy by bottom friction using the JONSWAP formulation (K. Hasselmann et al., 1973). This has minimal effect on runtime, but is critical for shallow water wave height accuracy.
- `dmx1`, `dmx2` (Dispersion parameters)
  Controls numerical diffusion in the wave advection scheme. They influence how sharply wave fronts are represented during spatial propagation of the wave spectrum. Lower values reduce numerical diffusion, resulting in sharper wave fronts and more accurate spatial gradients. However, these lower values increase the risk of producing unphysical oscillations or numerical instability in the solution (The SWAN team, 2006). Importantly, variations in `dmx1` and `dmx2` have negligible impact on computational cost because they do not fundamentally change the time step or grid resolution, but only the numerical scheme's diffusion characteristics (The SWAN team, 2006).

Appendix C
Additional spin up results

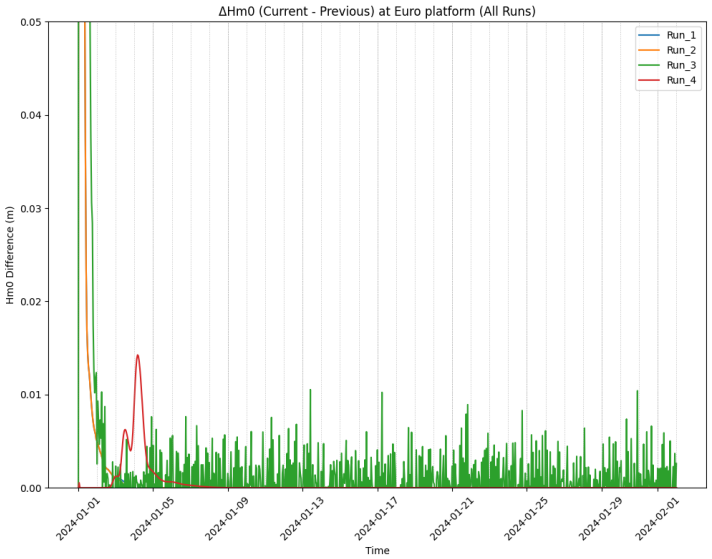

Figure Appendix C.1: Spin-up assessment for significant wave height at Euro platform.

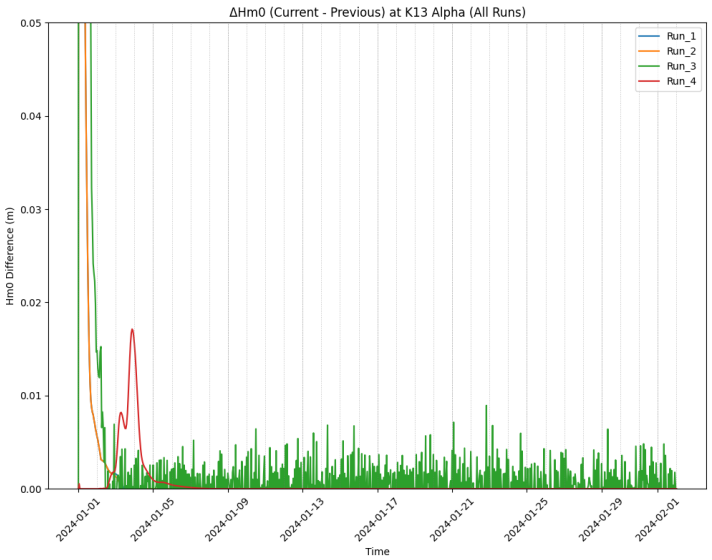

Figure Appendix C.2: Spin-up assessment for significant wave height at K13 Alpha.

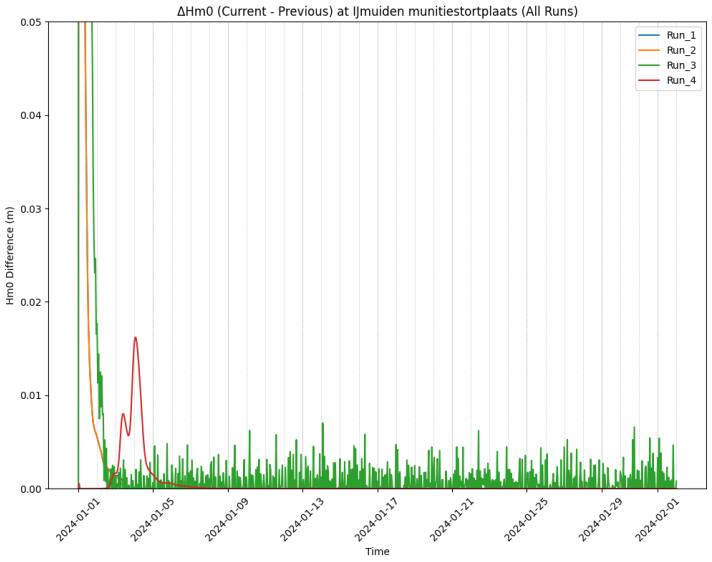

Figure Appendix C.3: Spin-up assessment for significant wave height at Ijmuiden munitiestortplaats.

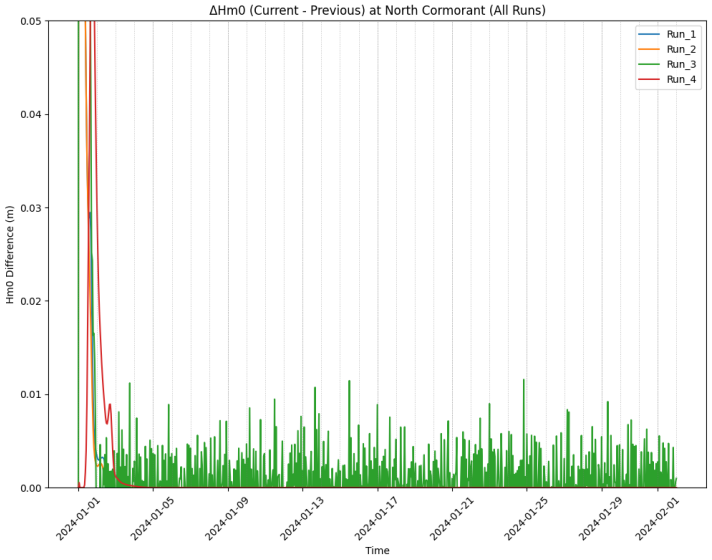

Figure Appendix C.4: Spin-up assessment for significant wave height at North Cormorant.

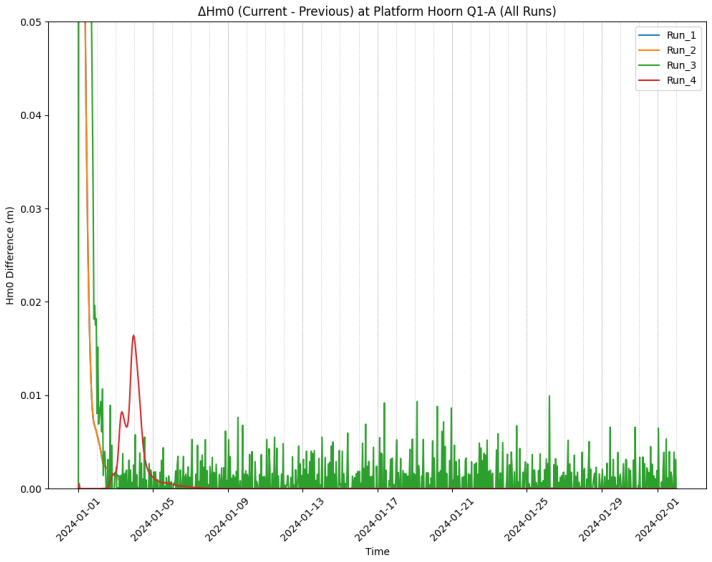

Figure Appendix C.5: Spin-up assessment for significant wave height at Platform Hoorn Q1-A.

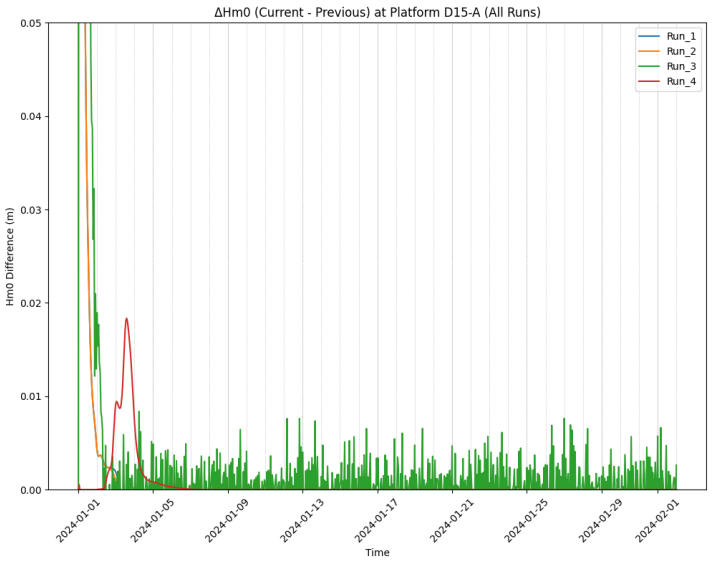

Figure Appendix C.6: Spin-up assessment for significant wave height at Platform D15-A.

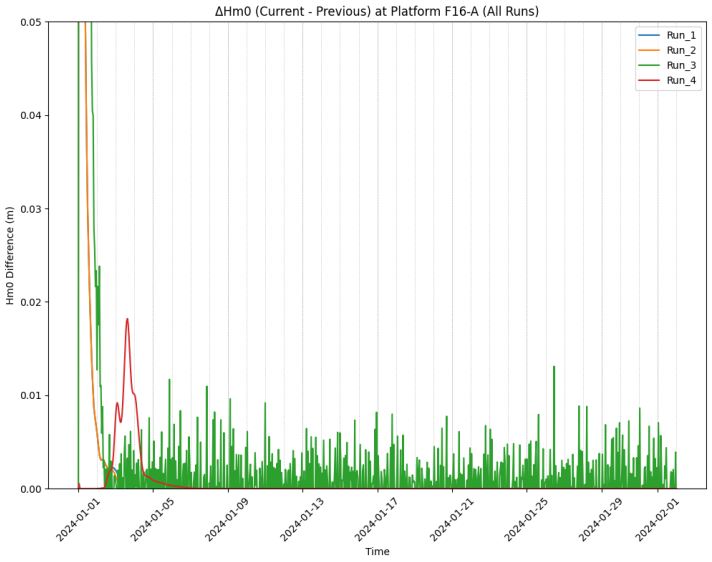

Figure Appendix C.7: Spin-up assessment for significant wave height at Platform F16-A.

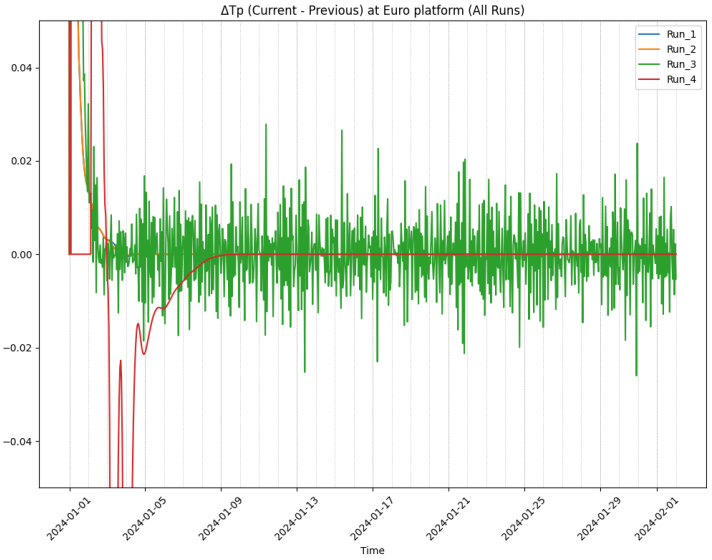

Figure Appendix C.8: Spin-up assessment for peak period at Euro platform.

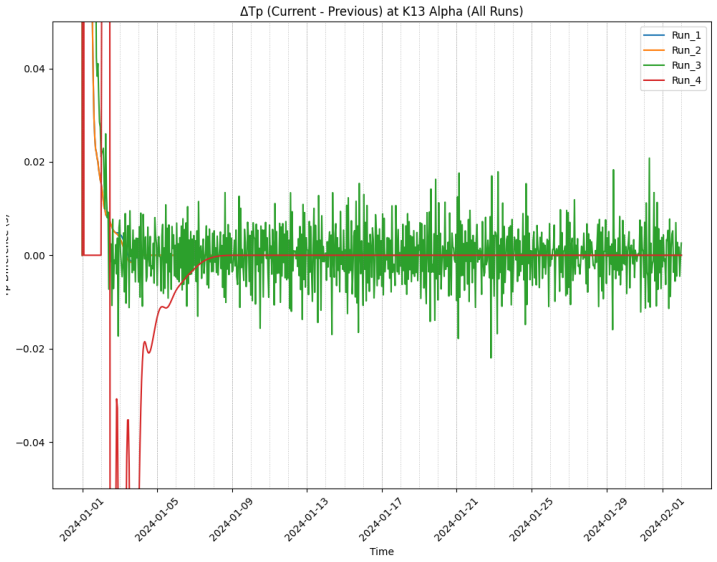

Figure Appendix C.9: Spin-up assessment for peak period at K13 Alpha.

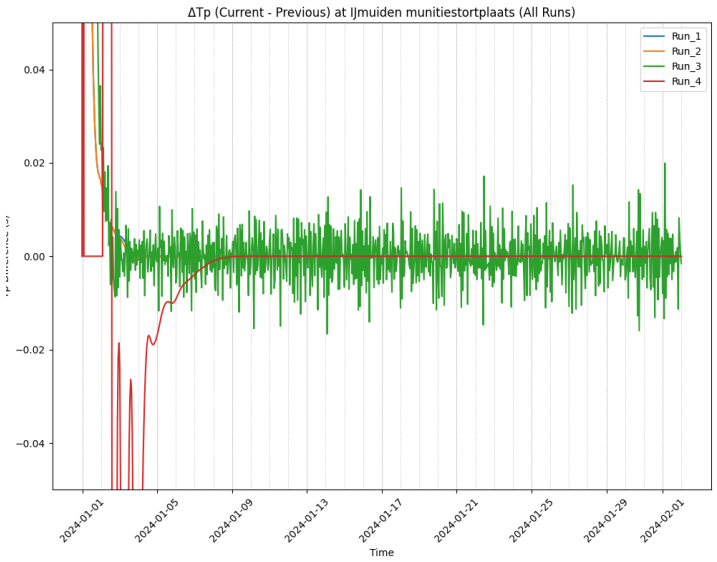

Figure Appendix C.10: Spin-up assessment for peak period at Ijmuiden munitiestortplaats.

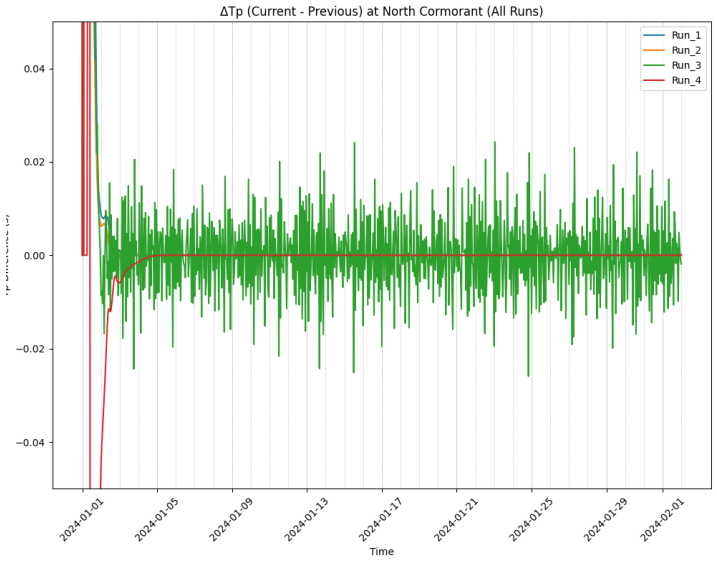

Figure Appendix C.11: Spin-up assessment for peak period at North Cormorant.

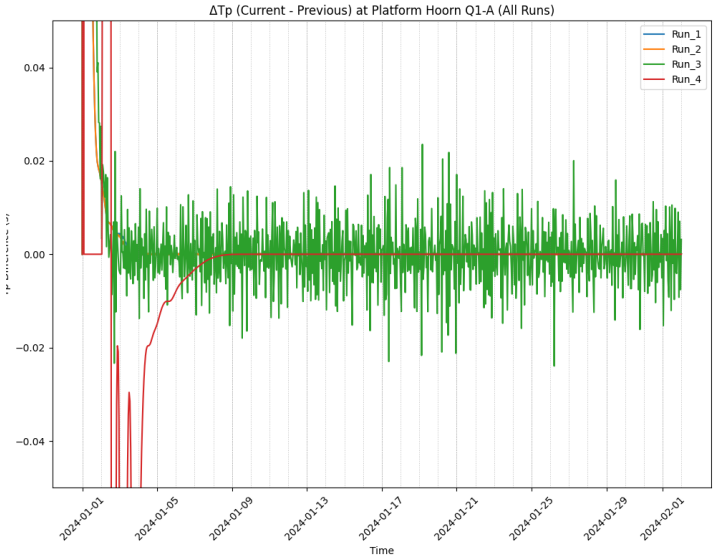

Figure Appendix C.12: Spin-up assessment for peak period at Platform Hoorn Q1-A.

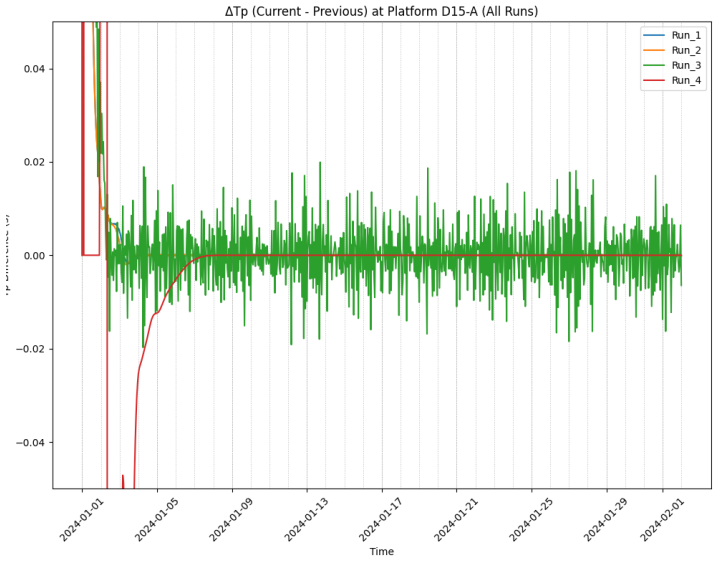

Figure Appendix C.13: Spin-up assessment for peak period at Platform D15-A.

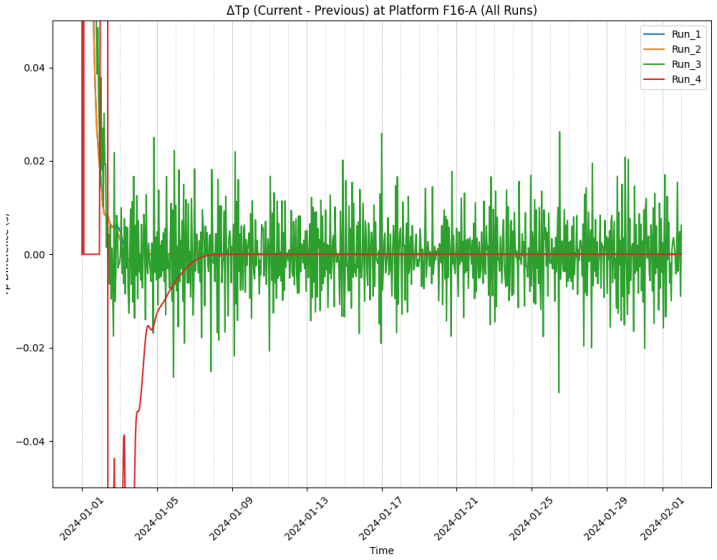

Figure Appendix C.14: Spin-up assessment for peak period at Platform F16-A.

Appendix D
Additional 2023 results

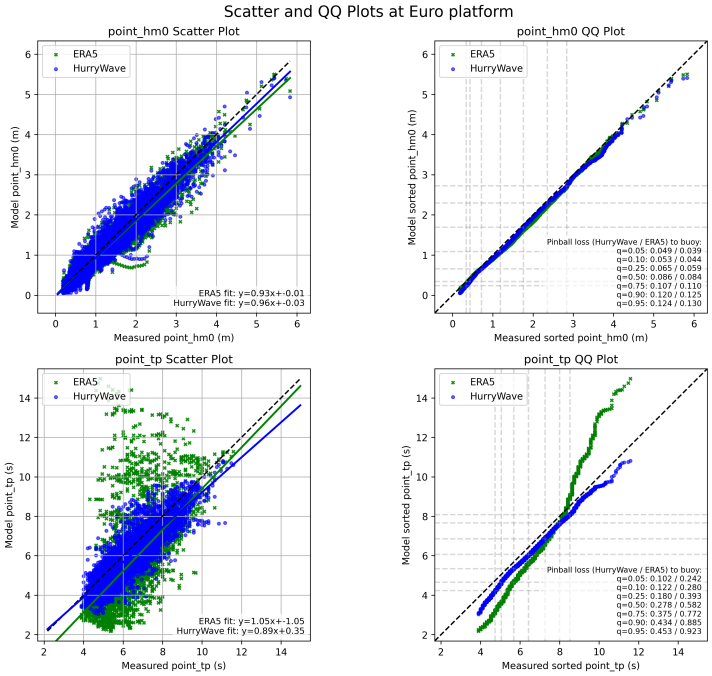

Figure Appendix D.1: Combined Scatter QQ in 2023 at Euro platform.

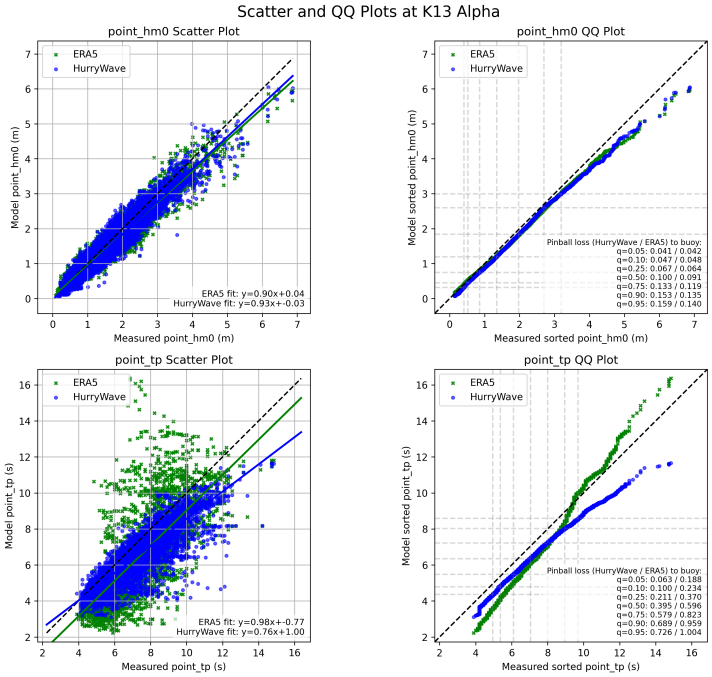

Figure Appendix D.2: Combined Scatter QQ in 2023 at K13 Alpha.

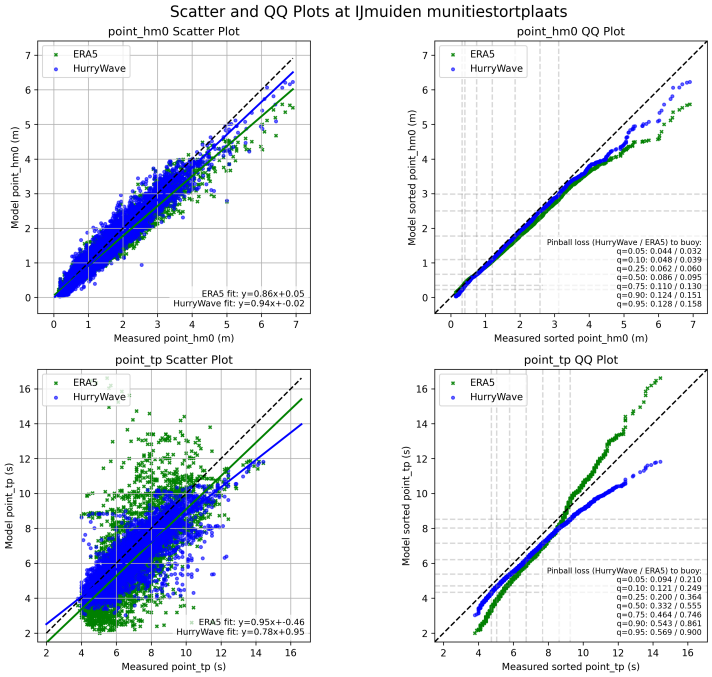

Figure Appendix D.3: Combined Scatter QQ in 2023 at Ijmuiden munitiestortplaats.

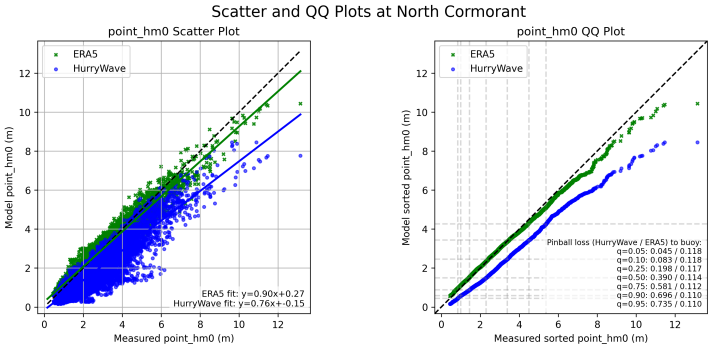

Figure Appendix D.4: Combined Scatter QQ in 2023 at North Cormorant.

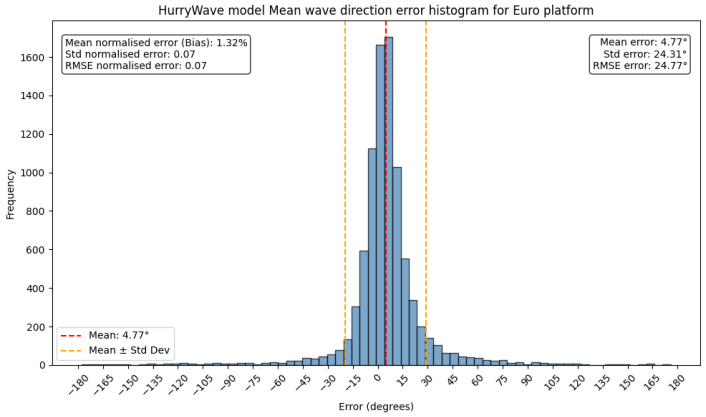

Figure Appendix D.5: Direction error histogram in 2023 at Euro platform.

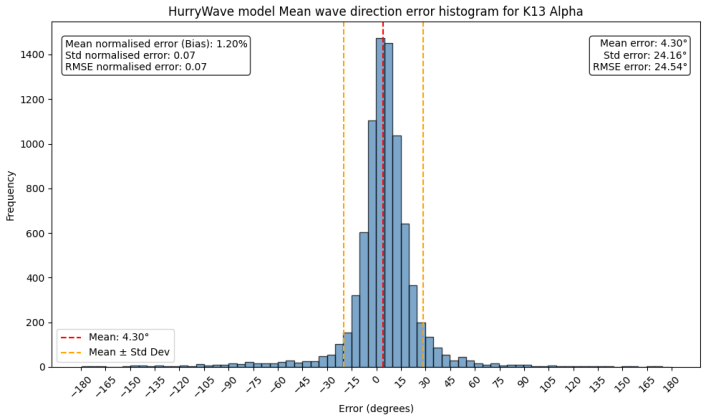

Figure Appendix D.6: Direction error histogram in 2023 at K13 Alpha.

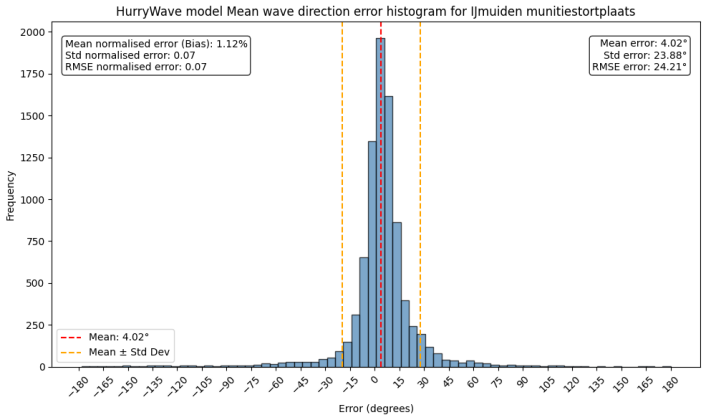

Figure Appendix D.7: Direction error histogram in 2023 at Ijmuiden munitiestortplaats.

Appendix E
Additional single storm results

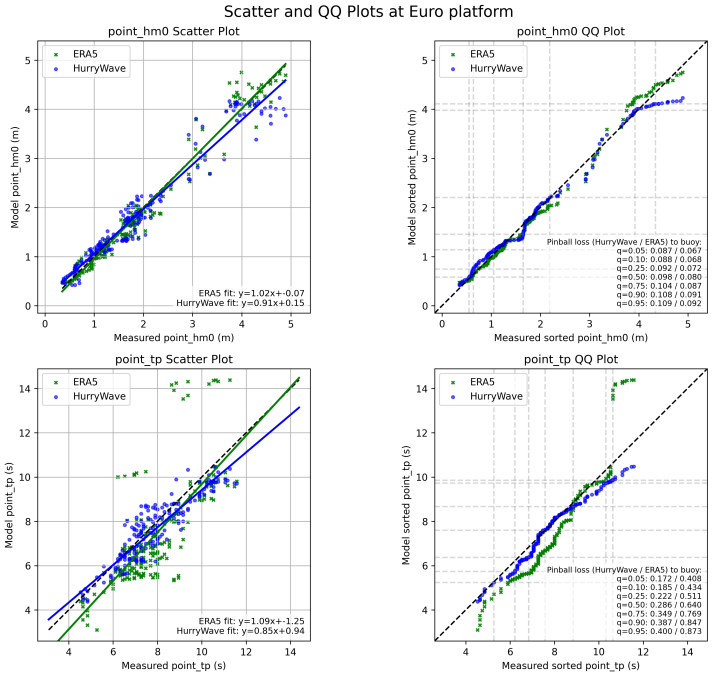

Figure Appendix E.1: Combined Scatter QQ during the Sinterklaas storm at Euro platform.

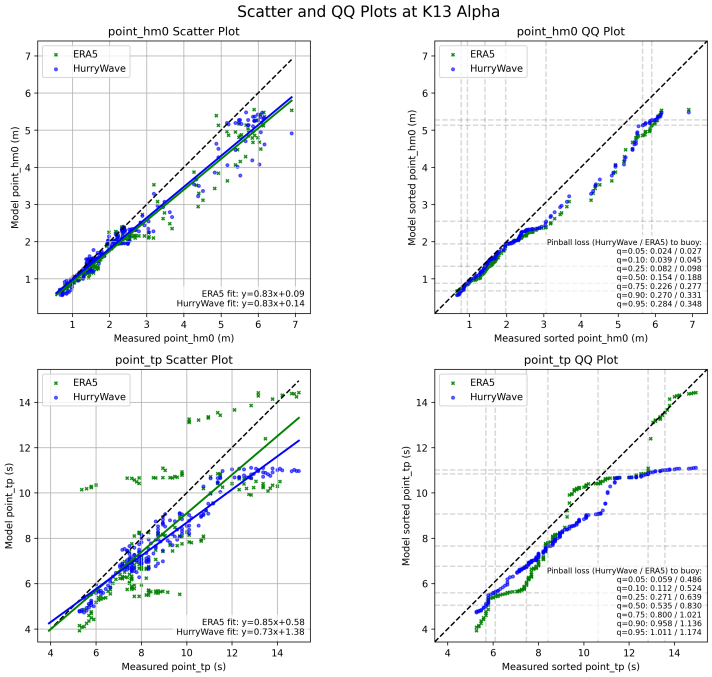

Figure Appendix E.2: Combined Scatter QQ during the Sinterklaas storm at K13 Alpha.

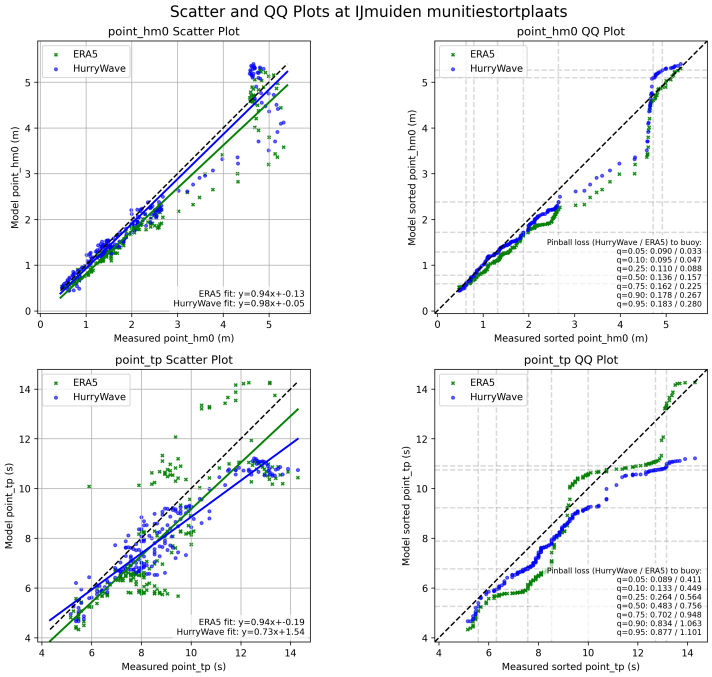

Figure Appendix E.3: Combined Scatter QQ during the Sinterklaas storm at Ijmuiden munitiestortplaats.

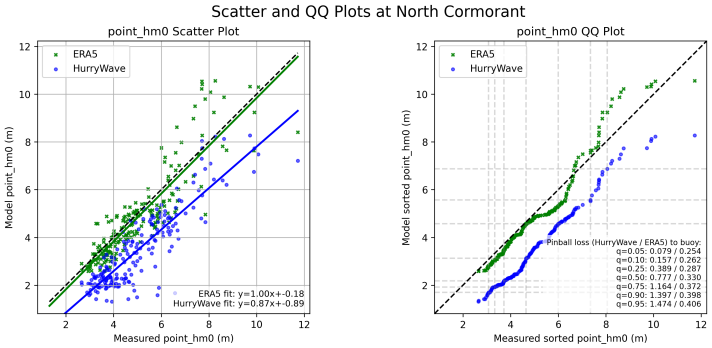

Figure Appendix E.4: Combined Scatter QQ during the Sinterklaas storm at North Cormorant.

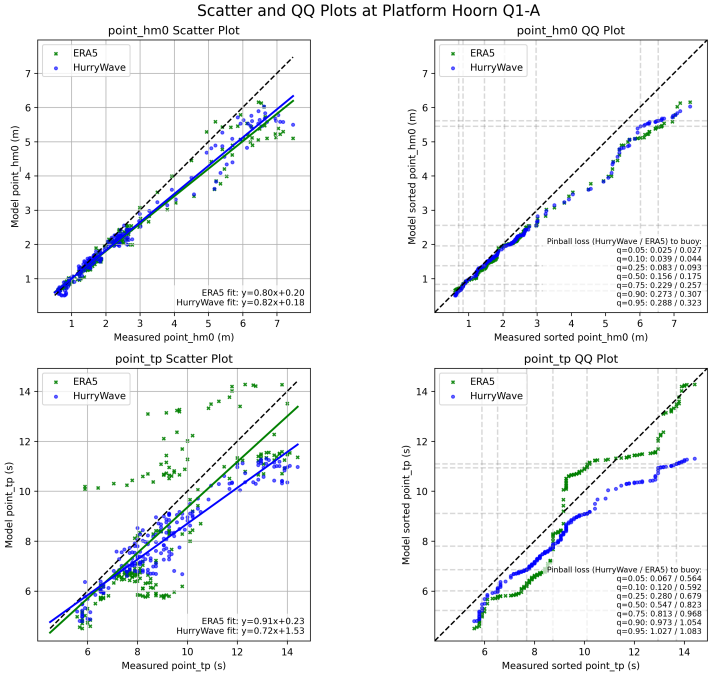

Figure Appendix E.5: Combined Scatter QQ during the Sinterklaas storm at Platform Hoorn Q1-A.

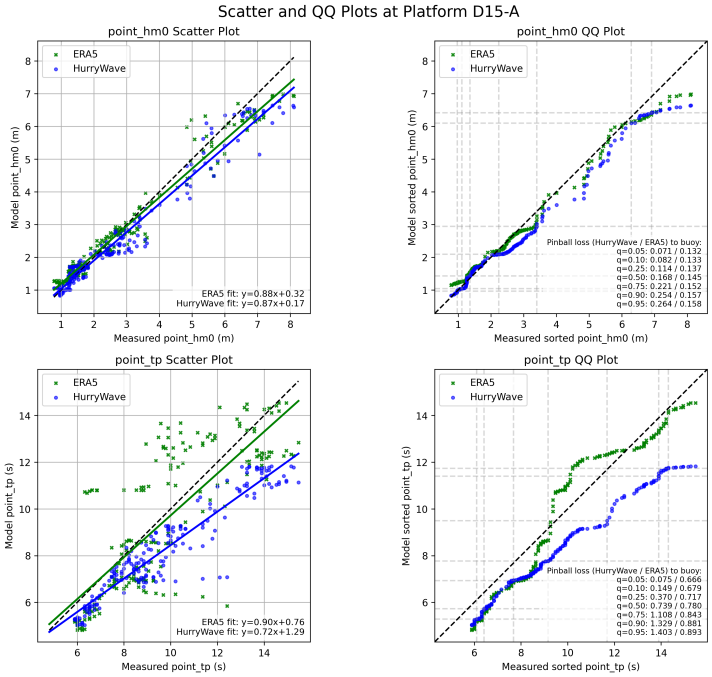

Figure Appendix E.6: Combined Scatter QQ during the Sinterklaas storm at Platform D15-A.

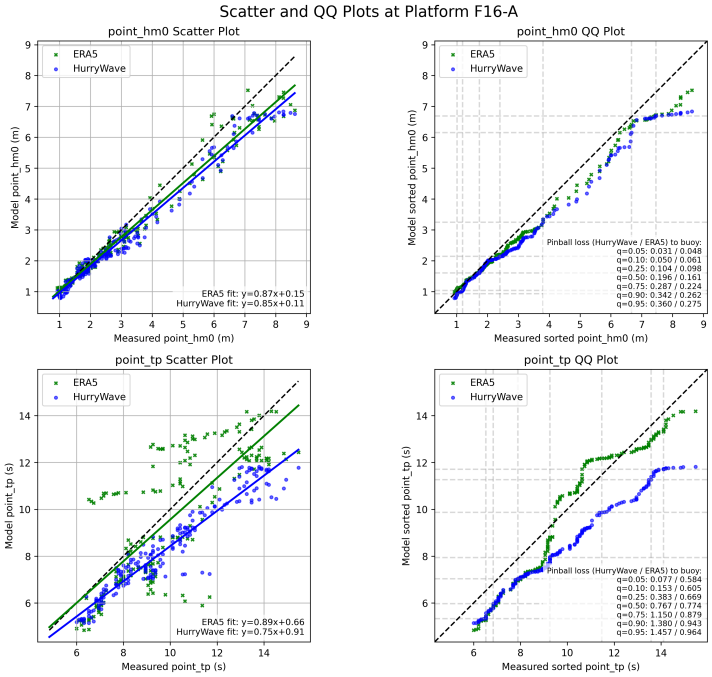

Figure Appendix E.7: Combined Scatter QQ during the Sinterklaas storm at Platform F16-A.

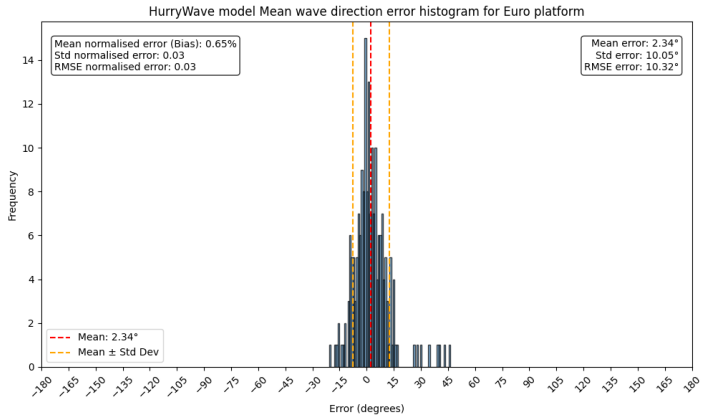

Figure Appendix E.8: Direction error histogram during the Sinterklaas storm at Euro platform.

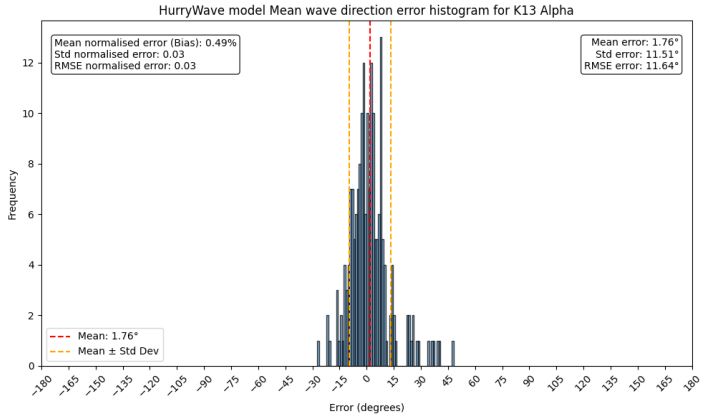

Figure Appendix E.9: Direction error histogram during the Sinterklaas storm at K13 Alpha.

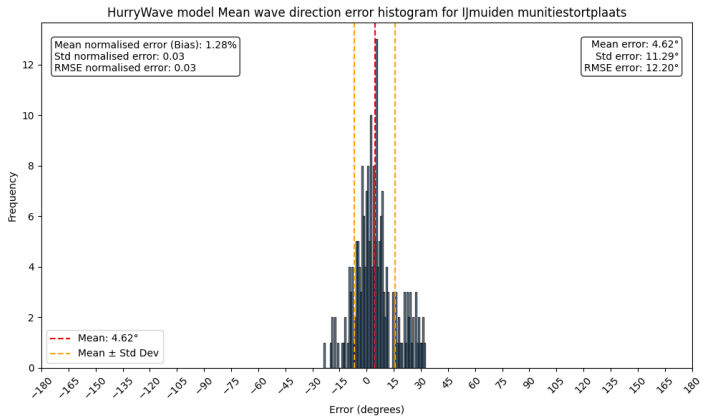

Figure Appendix E.10: Direction error histogram during the Sinterklaas storm at Ijmuiden munitiestortplaats.

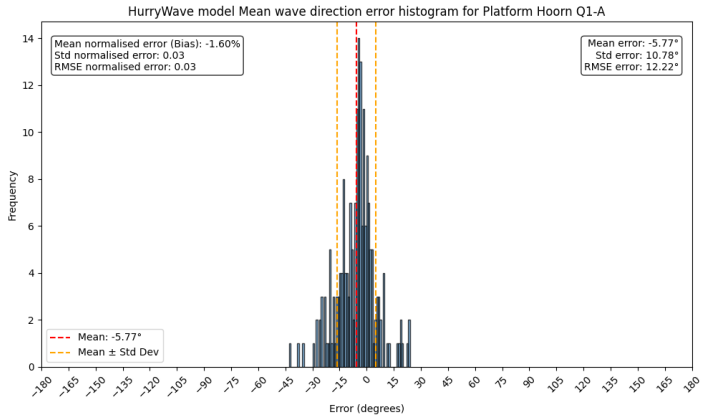

Figure Appendix E.11: Direction error histogram during the Sinterklaas storm at Platform Hoorn Q1-A.

Appendix F
Additional multiple year results

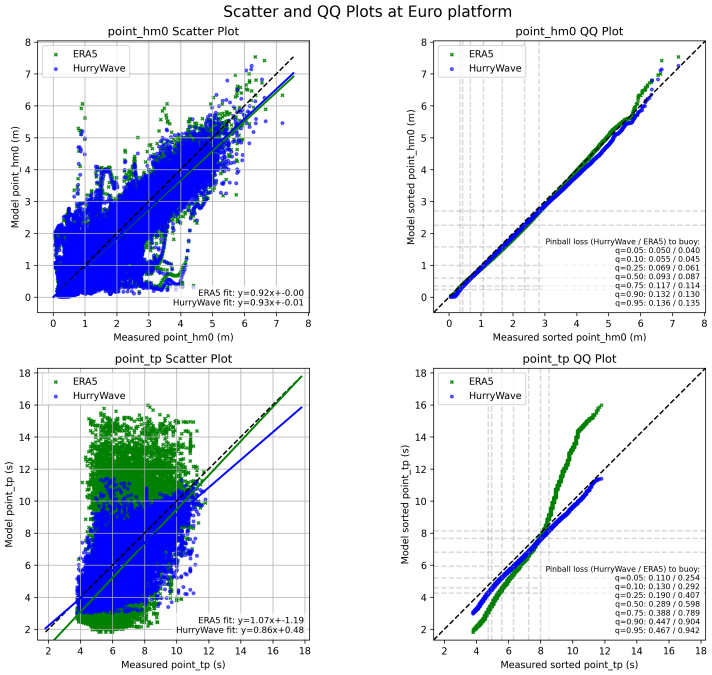

Figure Appendix F.1: Combined Scatter QQ for all years studied storm at Euro platform.

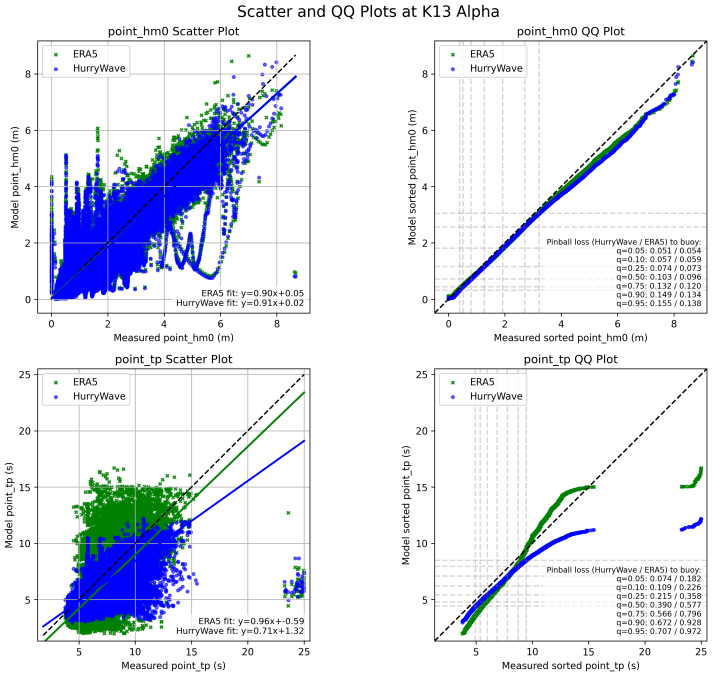

Figure Appendix F.2: Combined Scatter QQ for all years studied storm at K13 Alpha.

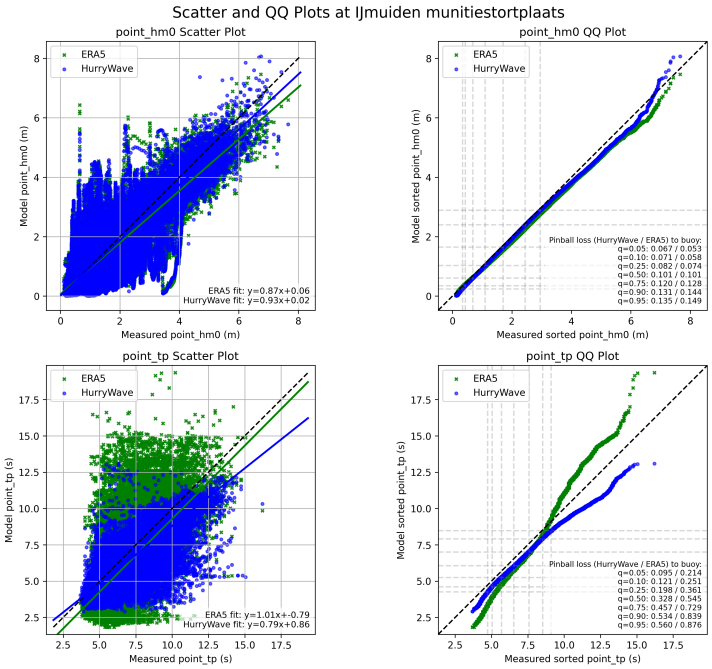

Figure Appendix F.3: Combined Scatter QQ for all years studied storm at Ijmuiden munitiestortplaats.

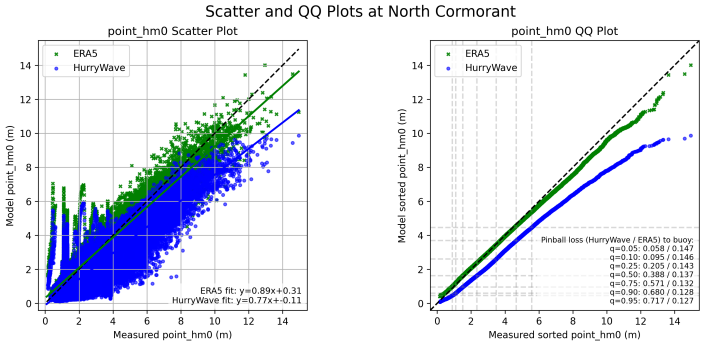

Figure Appendix F.4: Combined Scatter QQ for all years studied storm at North Cormorant.

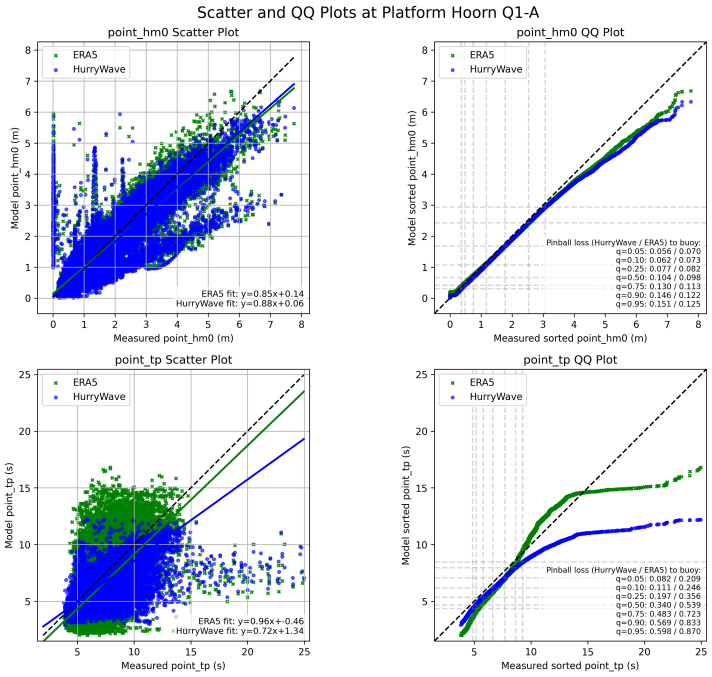

Figure Appendix F.5: Combined Scatter QQ for all years studied storm at Platform Hoorn Q1-A.

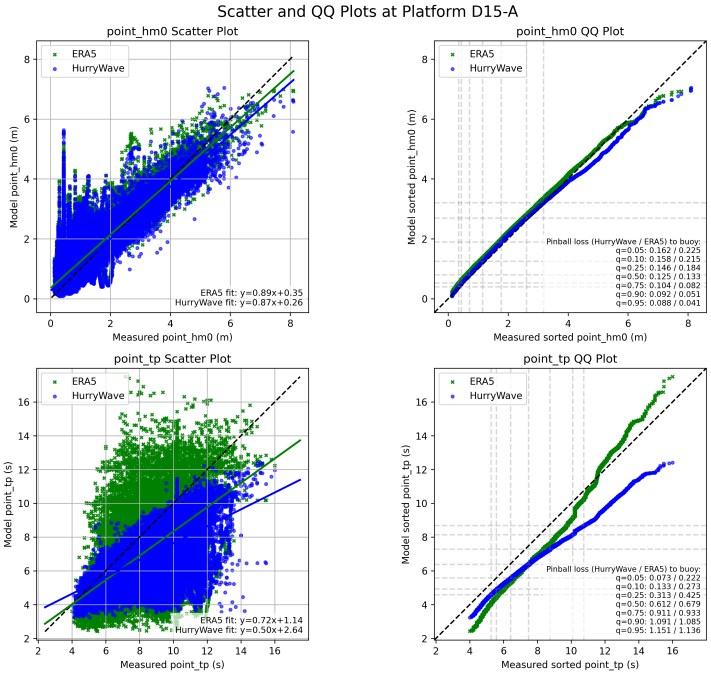

Figure Appendix F.6: Combined Scatter QQ for all years studied storm at Platform D15-A.

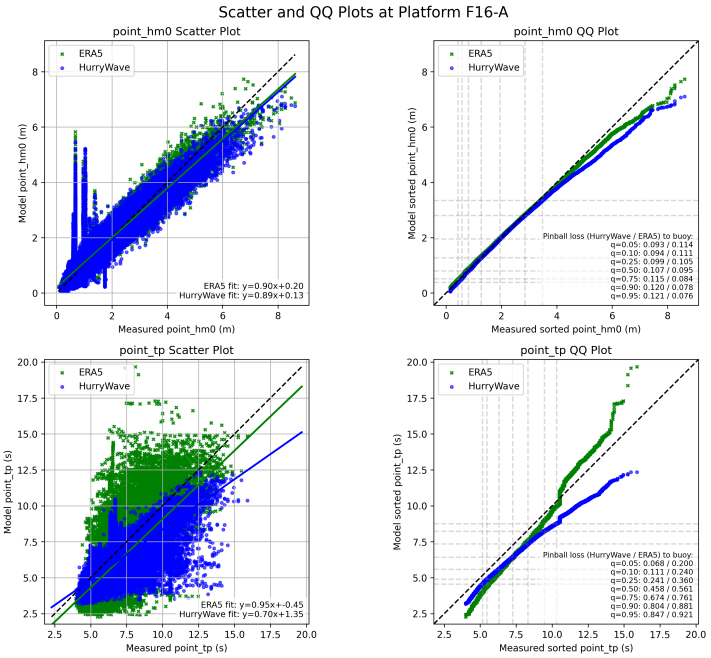

Figure Appendix F.7: Combined Scatter QQ for all years studied storm at Platform F16-A.

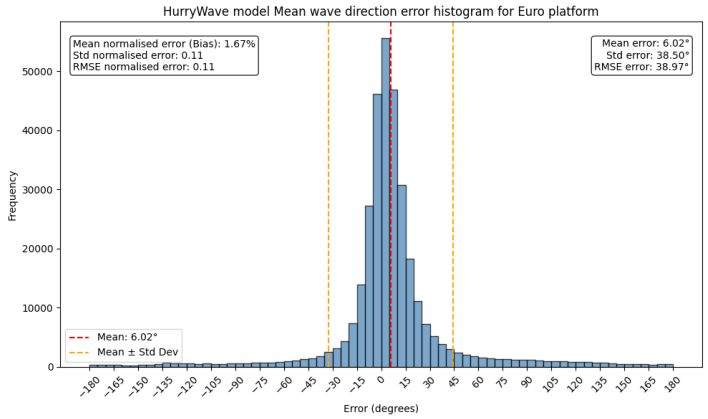

Figure Appendix F.8: Direction error histogram for all years studied at Euro platform.

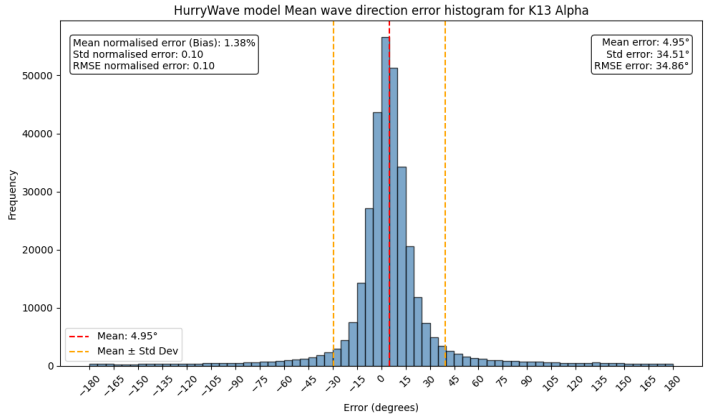

Figure Appendix F.9: Direction error histogram for all years studied at K13 Alpha.

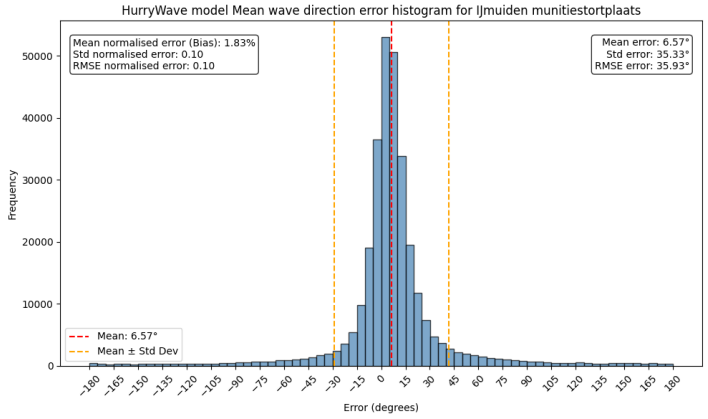

Figure Appendix F.10: Direction error histogram for all years studied at Ijmuiden munitiestortplaats.

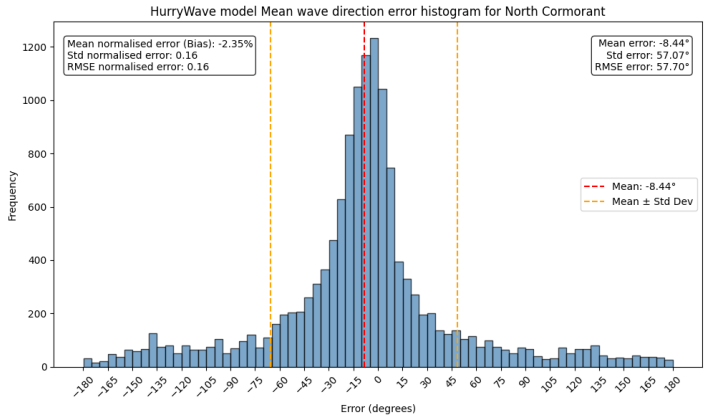

Figure Appendix F.11: Direction error histogram for all years studied at North Cormorant.

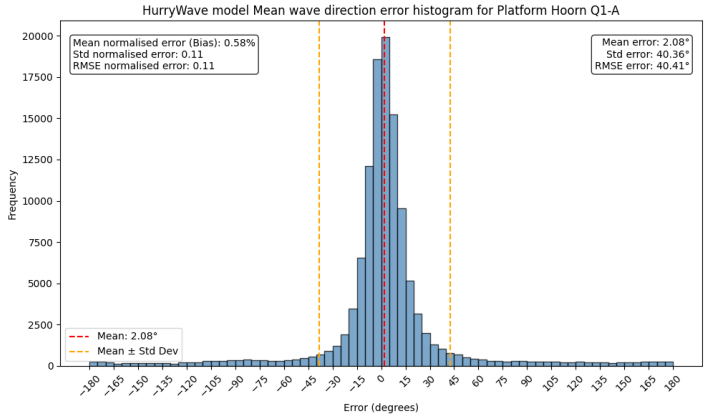

Figure Appendix F.12: Direction error histogram for all years studied at Platform Hoorn Q1-A.

Appendix G
Additional wind analysis results

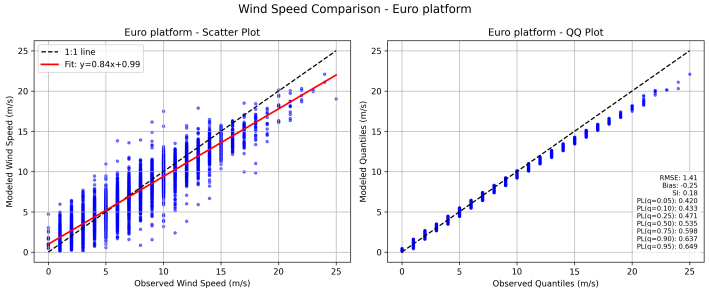

Figure Appendix G.1: Wind speed comparison in 2023 at Euro platform.

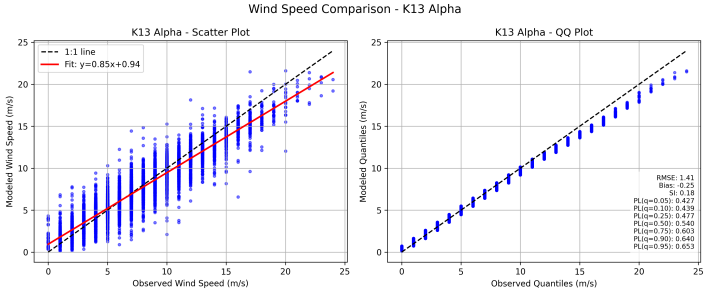

Figure Appendix G.2: Wind speed comparison in 2023 at K13 Alpha.

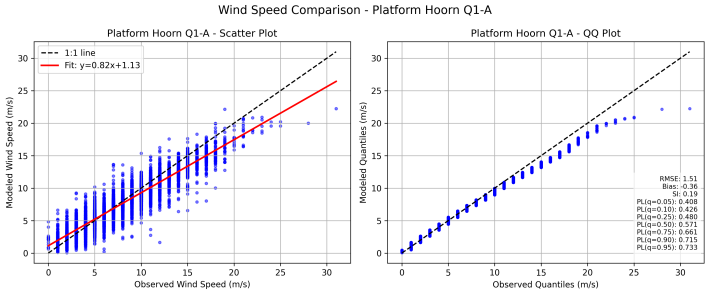

Figure Appendix G.3: Wind speed comparison in 2023 at Platform Hoorn Q1-A.

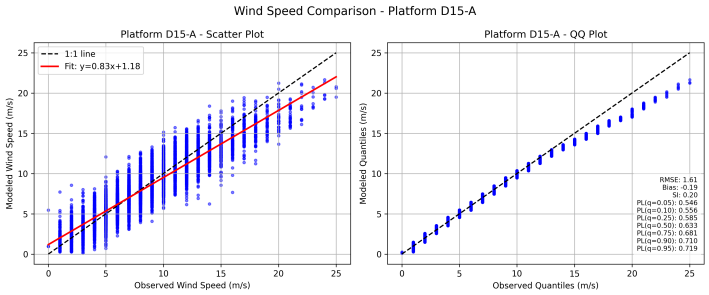

Figure Appendix G.4: Wind speed comparison in 2023 at Platform D15-A.

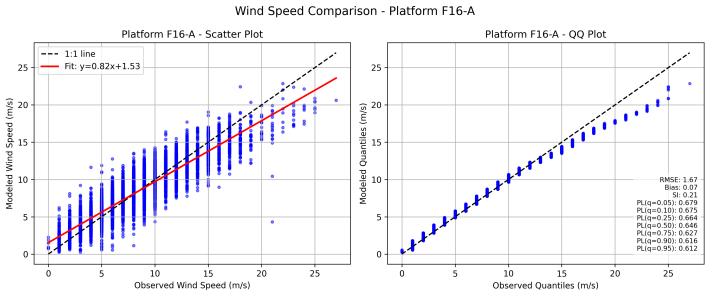

Figure Appendix G.5: Wind speed comparison in 2023 at Platform F16-A.

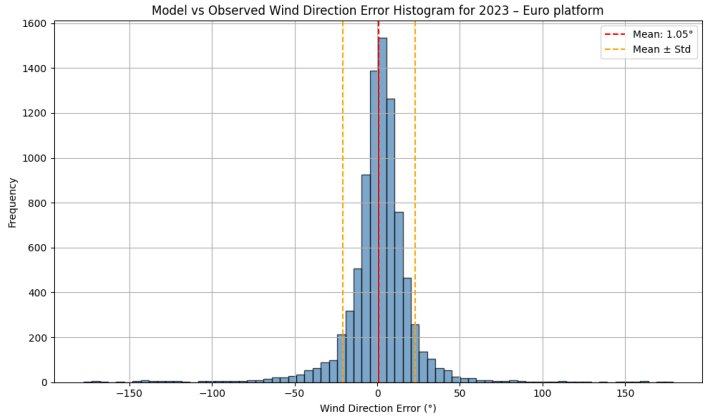

Figure Appendix G.6: Wind direction error histogram in 2023 at Euro platform.

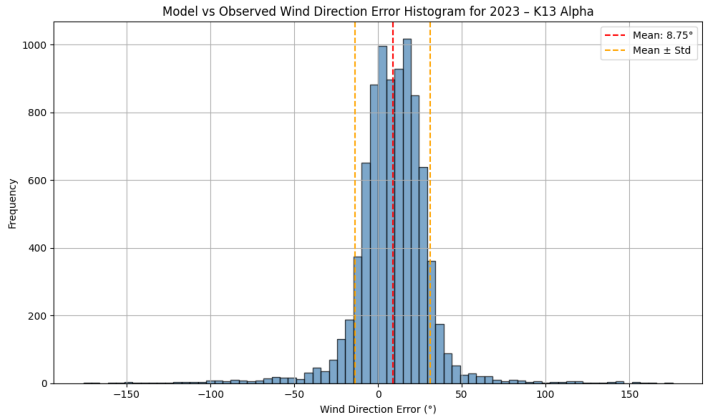

Figure Appendix G.7: Wind direction error histogram in 2023 at K13 Alpha.

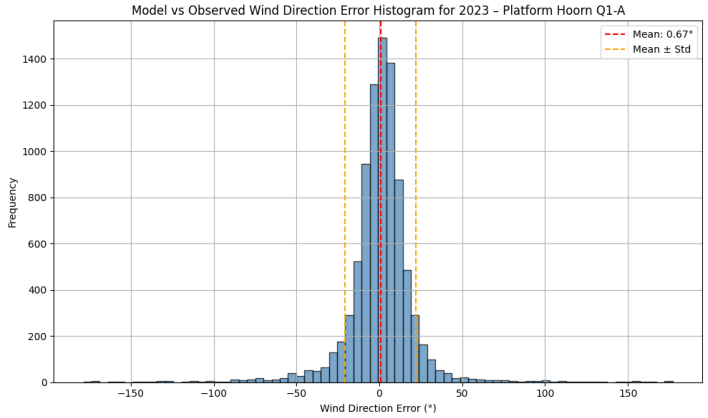

Figure Appendix G.8: Wind direction error histogram in 2023 at Platform Hoorn Q1-A.

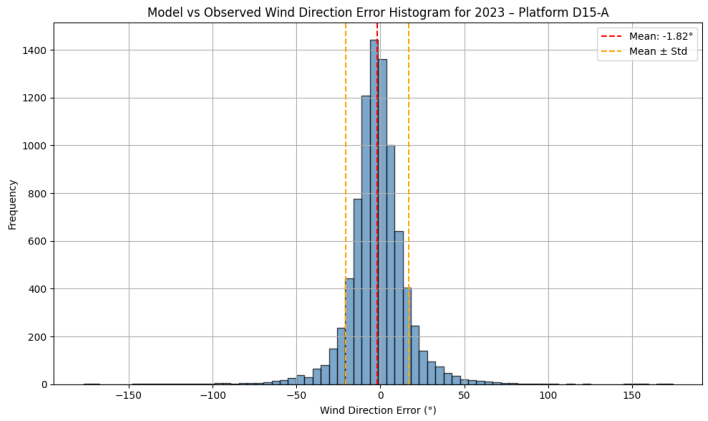

Figure Appendix G.9: Wind direction error histogram in 2023 at Platform D15-A.

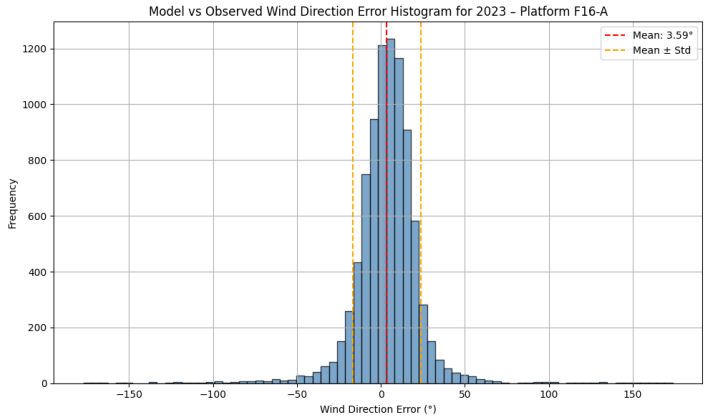

Figure Appendix G.10: Wind direction error histogram in 2023 at Platform F16-A.

Appendix H
Additional larger domain results

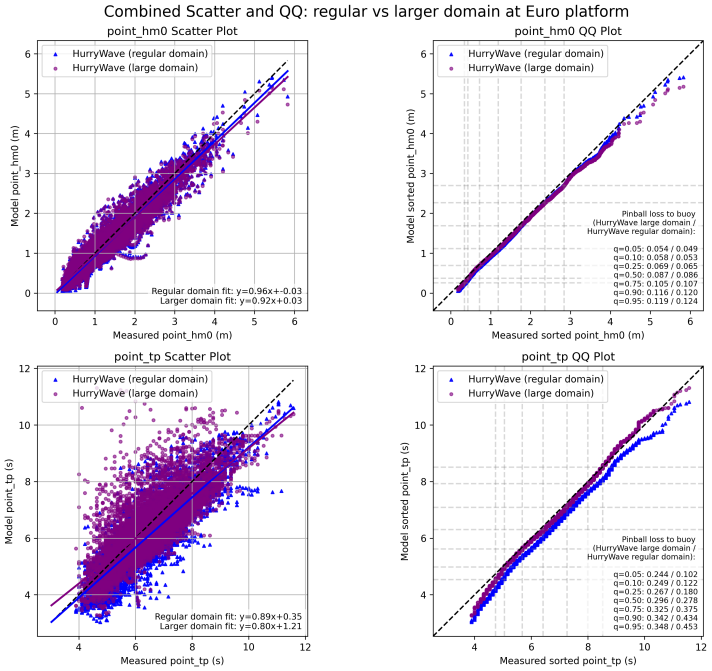

Figure Appendix H.1: Comparison of results between the regular and extended domain at Euro platform.

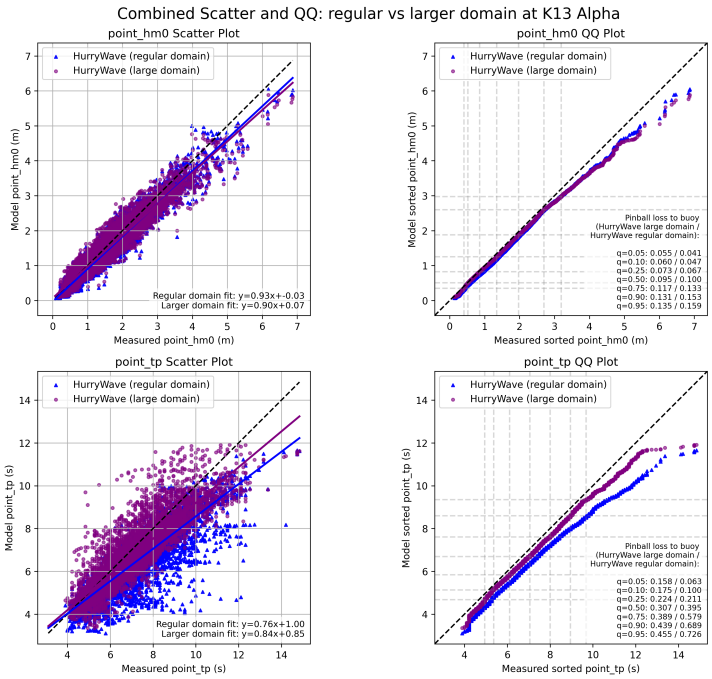

Figure Appendix H.2: Comparison of results between the regular and extended domain at K13 Alpha.

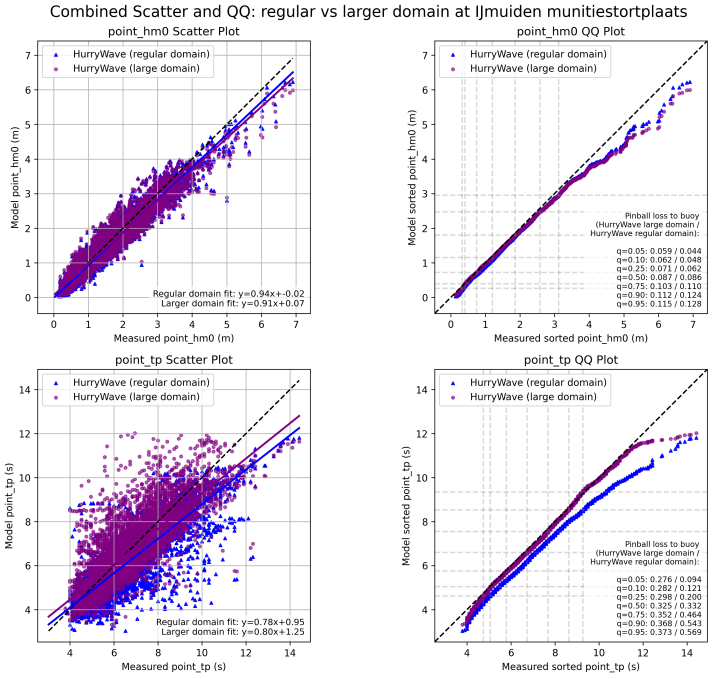

Figure Appendix H.3: Comparison of results between the regular and extended domain at Ijmuiden munitiestortplaats.

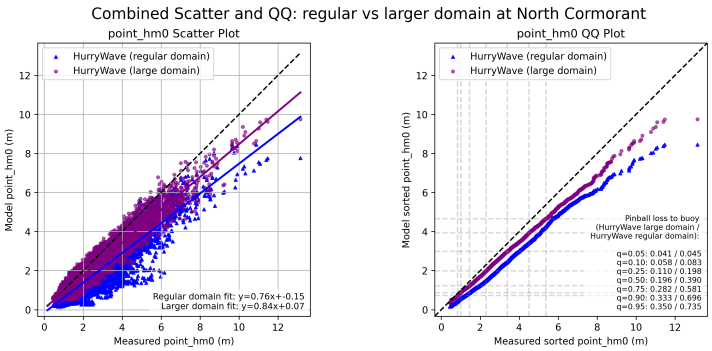

Figure Appendix H.4: Comparison of results between the regular and extended domain at North Cormorant.

Appendix I

Additional triad investigation results

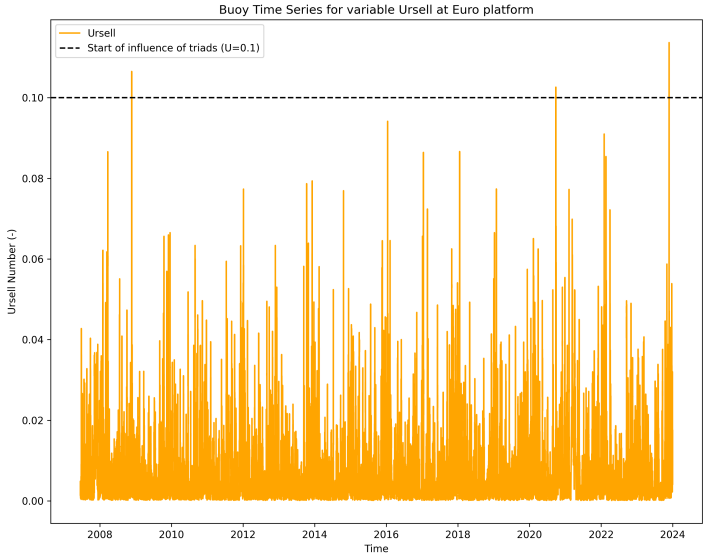

Figure Appendix I.1: Buoy data Ursell number plot at Euro platform.

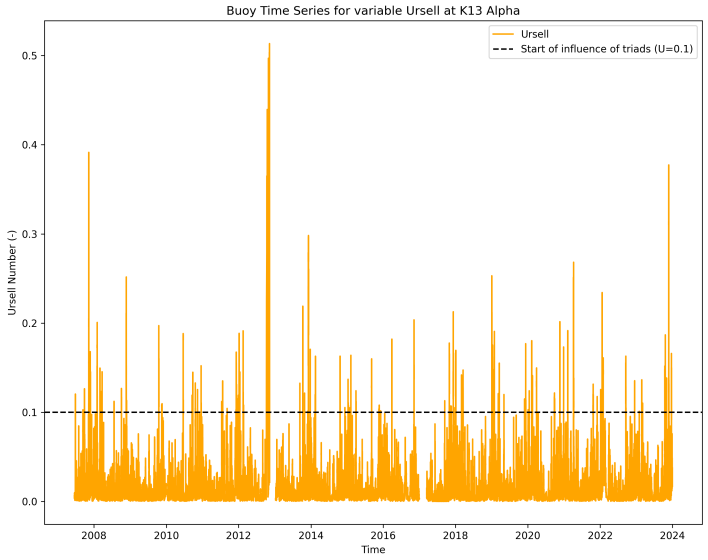

Figure Appendix I.2: Buoy data Ursell number plot at K13 Alpha.

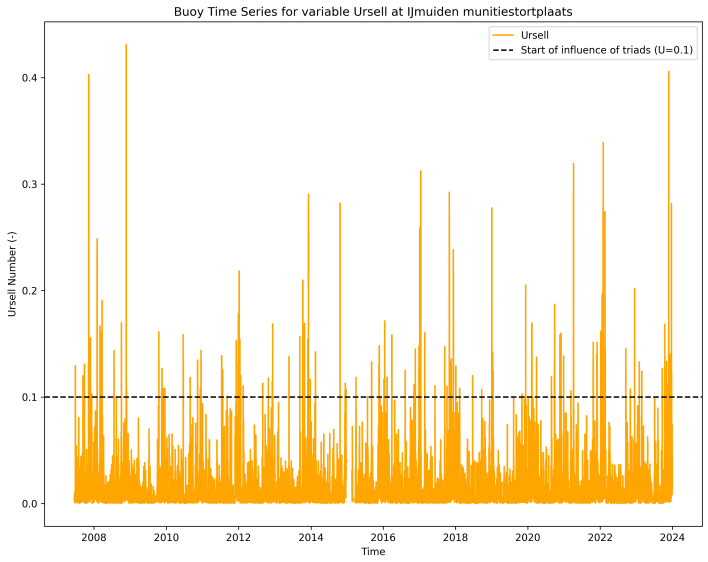

Figure Appendix I.3: Buoy data Ursell number plot at Ijmuiden munitiestortplaats.

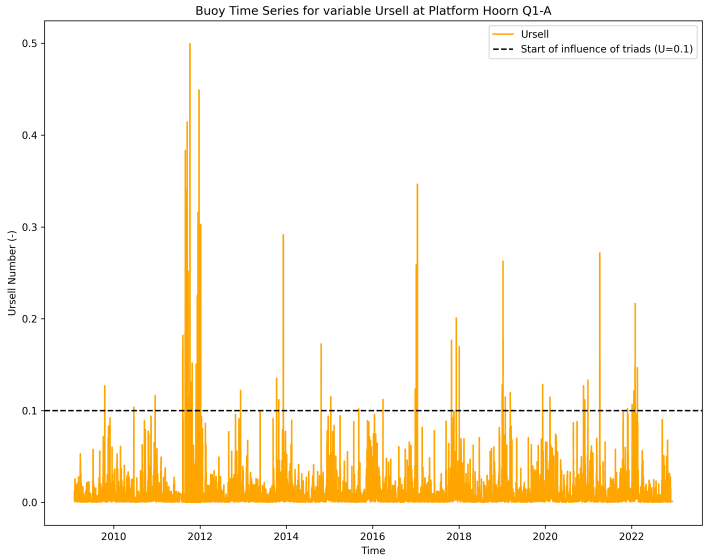

Figure Appendix I.4: Buoy data Ursell number plot at Platform Hoorn Q1-A.

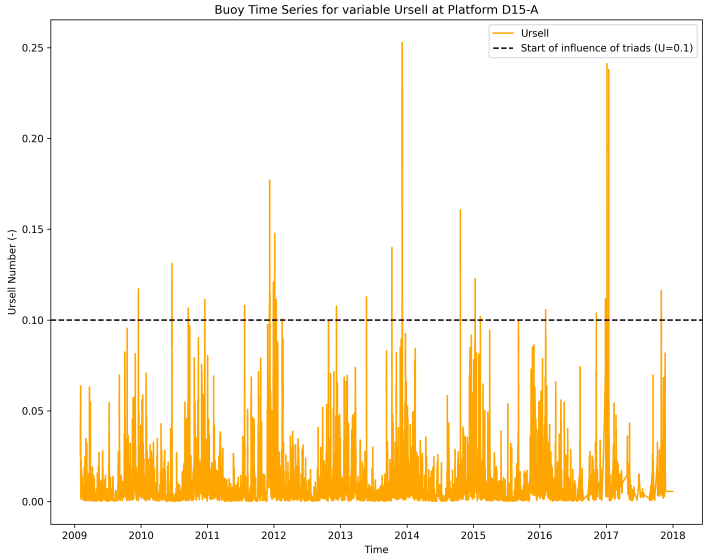

Figure Appendix I.5: Buoy data Ursell number plot at Platform D15-A.

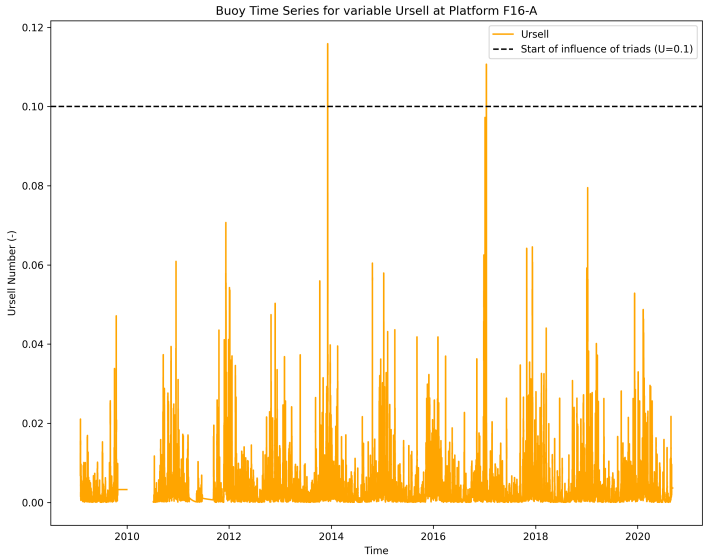

Figure Appendix I.6: Buoy data Ursell number plot at Platform F16-A.