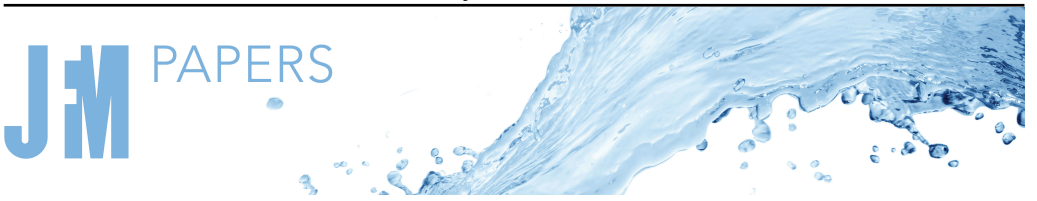
J. Fluid Mech. (2025), *vol.* 1007, A47, doi:10.1017/jfm.2025.84



Assessing the predictability of random ocean waves

Alexis Mérigaud[1], **Jiamin Zhu**[1] **and Paolino Tona**[1]

[1]IFP Énergies Nouvelles (IFPEN), 92500 Rueil-Malmaison, France

Corresponding author: Alexis Mérigaud, alexis.merigaud@ifpen.fr

(Received 23 July 2024; revised 3 January 2025; accepted 16 January 2025)

Real-time wave forecasting (RTWF) consists in predicting ocean wave motion or forces, from seconds to minutes in advance, using real-time measurements. For the successful development of RTWF, understanding wave predictability is essential. Usually, a deterministic 'predictable zone' (DPZ) is geometrically constructed from the wave group velocities and directions present in the spectrum. DPZs have little experimental evidence, and suffer ambiguities regarding the choice of cutoff frequencies and directions – since actual ocean waves are not band-limited. The present study addresses those shortcomings, by defining probabilistic predictable zones (PPZs) with respect to chosen uncertainty thresholds, using a rigorous statistical framework restricted to near-Gaussian sea states (precisely those where RTWF would be employed). PPZs are examined in idealised spectra and in a stereo dataset of a real wave field. It is shown that the PPZ geometry is quantitatively related to the sea state characteristics, through three physical parameters: two limiting group velocities (similar to the deterministic theory), and a directional spreading effect, which also limits the PPZ extent. While the lower group velocity depends on the chosen uncertainty threshold, the upper group velocity is better approximated by that of the spectrum peak frequency, which is a novel finding. The empirical data support the validity of the present PPZ theory. In contrast, both theoretical and empirical results contradict the fan-shaped predictable zones, constructed in the three-dimensional deterministic theory, thus highlighting the importance of considering stochasticity to understand the predictability of actual ocean waves.

Key words: surface gravity waves

1. Introduction

Real-time wave forecasting (RTWF), also often termed deterministic or phase-resolved forecasting, consists in predicting ocean wave motion or forces, from a few seconds to several minutes in advance, using measurements updated in real time. RTWF could show

© The Author(s), 2025. Published by Cambridge University Press.

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

substantial benefits in many coastal and marine applications, through the implementation of predictive control strategies, with the potential to improve the performance of wave energy conversion (Ringwood *et al.* 2014), dynamic positioning (Sørensen 2011), heave compensation (Woodacre *et al.* 2015) or vessel-mounted robotic manipulator (From *et al.* 2011) systems. Another example concerns floating wind turbines, of which the blade pitch angle can be controlled in real time, not only to maximise power and regulate rotor speed (which are 'classical' wind turbine control objectives), but also to reduce the floater roll and pitch (Raach *et al.* 2014; Skaare *et al.* 2015) or wave-induced mechanical loads (Ma *et al.* 2018). For the latter two objectives, the application of predictive control is a promising approach, which requires the prediction of wave forces onto the floater (Raach *et al.* 2014; Ma *et al.* 2018), see also the Floatech project.

Besides predictive control, RTWF would be a precious tool to improve decision-making in critical operations at sea, which require short quiescent periods, i.e. periods with very little wave motion in otherwise larger seas. Those include lifting operations, helicopter landing on ships, small submersible or surface vehicle recovery, crew transfer, etc. (Belmont *et al.* 2019). The possibility of predicting quiescent periods, or, conversely, dangerous wave trains, a few tens of seconds in advance, could significantly extend the range of sea conditions in which such operations can be safely carried out, thus mitigating risks and costs in marine renewable energy operations and maintenance.

RTWF is, however, a scientific and technical challenge, which has so far prevented it from becoming an industry standard (although at least two companies, FutureWaves and NextOcean, offer such services), while arousing a growing body of academic works. An RTWF system requires, at least, a sensing system which collects wave measurements in real time, and a prediction method, i.e. a mathematical and numerical procedure which encapsulates the physical laws of wave propagation to predict a wave signal from the collected measurements, at a point of interest in time and space.

The bulk of the RTWF literature focuses on the wave prediction step *per se*, typically assuming that the free-surface elevation (FSE) is directly measured at required locations. Although recent years have witnessed the use of machine learning for RTWF, with the ability to model nonlinear wave propagation features at low computational cost (Law *et al.* 2020; Zhang *et al.* 2022; Li *et al.* 2023), as yet the vast majority of wave prediction methods are based on deterministic numerical models, which express the physical laws of wave propagation. Most often, using available measurements, the wave field is decomposed into a finite number of discrete wave components, i.e. particular solutions of the wave equations, which are then propagated in space and time. When linear waves are assumed, those wave components are sinusoidal, and they are combined through a simple summation (Belmont *et al.* 2014; Naaijen *et al.* 2014; Hlophe *et al.* 2021, 2022, 2023), but much of the recent research effort has been dedicated to nonlinear wave models, capable of accommodating steeper wave fields, such as linear theory with nonlinear dispersion relation corrections (Stuhlmeier & Stiassnie 2021; Desmars *et al.* 2023; Kim *et al.* 2023; Meisner *et al.* 2023), Lagrangian models (Desmars *et al.* 2023; Kim *et al.* 2023) and higher-order spectral methods (Köllisch *et al.* 2018; Qi *et al.* 2018*a*; Klein *et al.* 2020; Desmars *et al.* 2023). In such models, the data assimilation step, whereby sparse, sometimes irregularly located measurements are used to infer underlying wave components, is critical (Köllisch *et al.* 2018; Al-Ani *et al.* 2019). Other avenues include the nonlinear Schrödinger (Simanesew *et al.* 2017; Klein *et al.* 2020) and Zakharov (Stuhlmeier & Stiassnie 2021) equations. Also note, for unidirectional waves, the original path chosen by Huchet *et al.* (2023), which consists in using ADCP velocity measurements to derive a complete set of initial conditions at the computational domain boundaries,

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

which are then propagated through numerical integration, using a higher-order spectral implementation dedicated to numerical wave tank modelling.

Apart from the machine learning and deterministic models succinctly reviewed above, RTWF can also be carried out using a probabilistic model. Indeed, in most sea conditions of operational interest, the wave field can be well-approximated by a Gaussian random field (Ochi 2005). As such, in contrast to the deterministic physical models, all measured and predicted variables can be modelled as Gaussian random variables, and their covariance structure can be exploited to derive optimal wave predictions in a probabilistic sense (Simanesew *et al.* 2017; Mérigaud & Ringwood 2018*a*; Mérigaud & Tona 2022). In fact, Simanesew *et al.* (2017) and Mérigaud & Ringwood (2018*a*) highlighted that the covariance structure can be calculated from the wave spectrum and the measurement layout. Most importantly, the probabilistic approach provides the uncertainty associated with the prediction, quantified for example in terms of mean-square error for each time horizon, which will be instrumental in the present study.

Indeed, regardless of the prediction method, the successful development of RTWF requires an accurate understanding of wave predictability, that is, the relation between the available measurements, and the prediction performance which can be expected. Ultimately, wave predictability governs the choice of an appropriate measurement system. In an operational context, it provides information on the prediction confidence, which is certainly an essential complement to the prediction itself, as discussed by Fucile *et al.* (2018).

Wave predictability is usually characterised by a predictable zone, or wave prediction diagram (see e.g. Abusedra & Belmont 2011; Naaijen *et al.* 2014), described as a domain in space and time, where the waves are assumed accurately predictable. In two dimensions (unidirectional waves), the predictable zone is geometrically constructed from the group velocities present in the wave spectrum (Naaijen *et al.* 2014), thus obtaining triangular zones with slopes governed by two limiting velocities, as illustrated in figure 1(*a*). The wave prediction diagram has been extended to three-dimensional (i.e. multidirectional) wave fields by Qi *et al.* (2018*b*), which requires the assumption that the totality of the wave energy is directionally restricted to a given angular sector narrower than 180°. Fan-shaped zones in the horizontal $(x - y)$ plane are obtained, which gradually fade out with the prediction horizon, as shown in figure 1(*b*).

A limitation of those theoretical predictable zones is that they are not usually specified with respect to a certain accuracy threshold, quantified, for example, through the mean-square error or correlation. Yet, the application of wave prediction methods, in simulated or experimental data, typically yields continuous maps of error levels (Abusedra & Belmont 2011; Naaijen *et al.* 2014; Al-Ani *et al.* 2020; Law *et al.* 2020; Hlophe *et al.* 2021; Huchet *et al.* 2023; Kim *et al.* 2023). In two dimensions, contour levels of those maps exhibit triangular shapes, that resemble those predicted by the theory, but a difficulty consists in relating them to group velocities at specific frequencies, since actual ocean wave spectra are not band-limited. It is particularly the case for the low-frequency cutoff, which determines the largest limiting group velocity: Law *et al.* (2020) choose their lower cutoff frequency such that 85 % of the spectrum energy is above it; instead, Hlophe *et al.* (2021) retain a fraction of 0.7 times the peak frequency; finally Huchet *et al.* (2023) propose a criterion based on a percentage (5 %, 25 % and 75 %) of the spectrum value at its peak frequency. In those three references, however, the empirical predictable zone area appears, in general, larger than the theoretical one, suggesting that the maximum limiting group velocity is overestimated.

Except for the exact specification of limiting group velocities, two-dimensional prediction zones are well-supported by existing simulated and experimental results. The

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

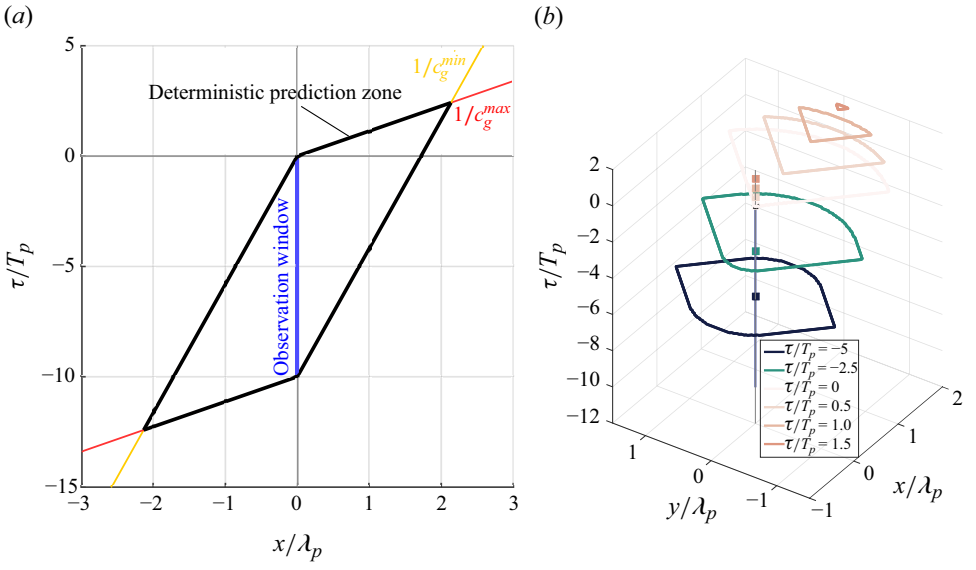


Figure 1. Two-dimensional (*a*) and three-dimensional (*b*) deterministic wave prediction diagrams. Time (relative to present) is represented by the vertical axis. In this example, waves are observed in $x = 0$ for a duration of 10 peak wave periods into the past, represented by a thick, blue line. Cut-in and cutoff frequencies are $0.7/T_p$ and $3/T_p$ Hz, respectively. Those provide the two limiting group velocities in both (*a*) and (*b*). In *(b),* the fan-shaped predictable zones are shown for six different time horizons, marked by square markers on the vertical axis. In this example, all wave energy content propagates within a $\pm 45°$ angular sector centred around the x axis. Refer to Qi *et al.* (2018*b*) for more detail on the diagram construction.

situation is different in three dimensions, partly due to the fact that three-dimensional wave field data are extremely rare. In the sea trial of Al-Ani *et al.* (2020), based on waves measured through the WAMOS system, the correlation between predictions and observations is calculated within the three-dimensional wave predictable zone, constructed following Qi *et al.* (2018*b*), but the empirical correlation decreases to low values as the time horizon increases, much more rapidly than the theoretical predictable zone vanishes. In the three-dimensional wave tank experiments of Kim *et al.* (2023), the prediction performance at the probe locations is shown for several time horizons, with the theoretical predictable zones superimposed, but the limited number of probes does not support conclusive evidence that the fan shape of the predictable zones is verified. Another notable study based on three-dimensional real wave data is that of Simanesew *et al.* (2017), already mentioned, where waves are measured through a 10×10 probe arrangement in wave tank experiments, although the results of that work do not provide indications of the three-dimensional aspect of wave predictability.

Apart from the Gaussian probabilistic approach of Mérigaud & Ringwood (2018*a*), which provides achievable prediction uncertainty along with the optimal prediction, few other works are concerned with a statistical characterisation of wave predictability. Fucile *et al.* (2018) propose a methodology to quantify prediction uncertainty, also based on a Gaussian wave field model, with the additional assumption that predictions are obtained by fitting a number of sinusoidal wave components to the observations. Like in Mérigaud & Ringwood (2018*a*), it is shown that the uncertainty depends on the wave spectrum. Finally, among machine learning approaches, it is worth mentioning the work of Zhang *et al.* (2022), who employed variational Bayesian machine learning to provide RTWF with quantified uncertainty, in a two-dimensional laboratory wave flume.

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

The main objective of the present work is to specify further the notion of wave predictability, which is revisited in the probabilistic framework of Simanesew *et al.* (2017) and Mérigaud & Ringwood (2018*a*). In both two- and three-dimensional wave fields, this framework allows the definition of probabilistic predictable zones (PPZs) with respect to a chosen uncertainty threshold, which we choose to quantify through the normalised mean-square error (NMSE). The PPZs are thoroughly compared with their deterministic counterparts found in the literature. It is shown that, although the physical reasoning may be partly obscured by the stochastic calculations, the resulting PPZ geometry can be approximated from key characteristics of the wave spectrum, using relatively simple physical interpretations – at least in simple wave spectra and measurement layouts. Finally, empirical evidence of the calculated predictable zones is provided, using a real, three-dimensional sea wave dataset, recorded by means of a stereo camera system.

The remainder of this article is organised as follows: § 2 introduces conventions regarding observation and prediction configurations, which are used throughout the rest of this work. Section 3 presents the Gaussian probabilistic approach to RTWF, which provides the basis for defining PPZs. In §§ 4 and 5, wave predictability is studied in unidirectional and directional idealised wave spectra, respectively. Particular attention is given to the physical interpretation of the PPZ geometry, and to the similarities and differences with the usual deterministic theory. The present probabilistic theory is tested against real-sea experimental data in § 6. Finally, the main results of this study are summarised and discussed in § 7.

2. Observation and prediction configuration

Like the vast majority of the RTWF literature, this work assumes that the FSE, denoted as η, can be directly measured at any desired set of locations of the horizontal plane. In addition, the predicted wave signal also consists in the FSE.

More specifically, η is measured at a number M_o of 'observation' locations, to predict its values at a number M_p of 'prediction' locations. Denote as $\mathbf{z}_o(t) \in \mathbb{R}^{M_o}$ the observation vector time series, formed by the FSE at the observation locations, and $\mathbf{z}_p(t) \in \mathbb{R}^{N_p}$ the prediction vector time series, formed by the FSE at the prediction locations, that is, $\mathbf{z}_o(t) = \big(\eta(\mathbf{x}_1, t), \cdots \eta(\mathbf{x}_{M_o}, t)\big)^\mathsf{T}$ and $\mathbf{z}_p(t) = \Big(\eta(\mathbf{x}'_1, t), \cdots \eta(\mathbf{x}'_{M_p}, t)\Big)^\mathsf{T}$. For example, in figure 2, η is observed at two locations $\mathbf{x}_1 = (x_1, y_1)$ and $\mathbf{x}_2 = (x_2, y_2)$, and predicted at a location $\mathbf{x}'_1 = (x'_1, y'_1)$, so that $M_o = 2$ and $\mathbf{z}_o(t) = \big(\eta(\mathbf{x}_1, t), \eta(\mathbf{x}_2, t)\big)^\mathsf{T}$, while $M_p = 1$ and $\mathbf{z}_p(t) = \eta(\mathbf{x}'_1, t)$.

Typically, the prediction task is implemented in a sliding-window fashion. More specifically, let t denote the present time, at which the measurements and predictions are updated. Denote $\tau_o = \big(\underline{\tau}_o \cdots \overline{\tau}_o\big)^\mathsf{T} \in \mathbb{R}^{N_o}$ and $\tau_p = \big(\underline{\tau}_p \cdots \overline{\tau}_p\big)^\mathsf{T} \in \mathbb{R}^{N_p}$, the sets of observation and prediction instants relatively to present time t, where the constants $\underline{\tau}_o \leqslant \overline{\tau}_o \leqslant 0$ and $\underline{\tau}_p \leqslant \overline{\tau}_p$. For the sake of simplicity, the observation instants (respectively, prediction instants) are assumed regularly sampled between $t + \underline{\tau}_o$ and $t + \overline{\tau}_o$ (respectively $t + \underline{\tau}_p$ and $t + \overline{\tau}_p$). At time t, the available observations, which are inputs to the prediction, are concatenated into a vector $\mathbf{Z}_o(t)$, defined as follows:

$$\mathbf{Z}_o(t) = \begin{pmatrix} \mathbf{z}_o(t + \underline{\tau}_o) \\ \vdots \\ \mathbf{z}_o(t + \overline{\tau}_o) \end{pmatrix}. \tag{2.1}$$

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

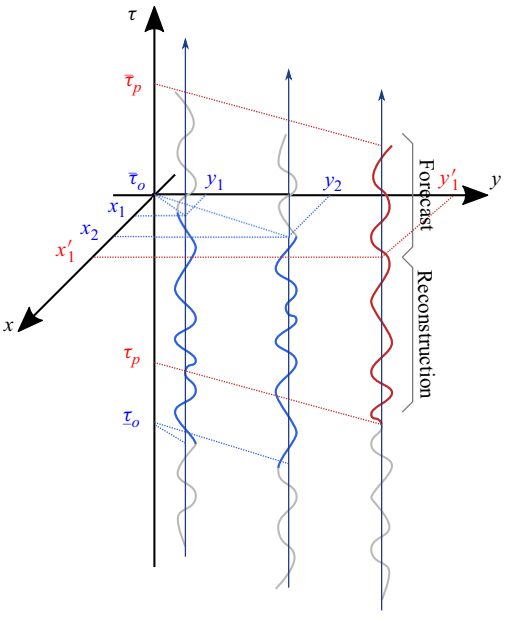


Figure 2. An example of observation and prediction configuration, with two measurement points $\mathbf{x}_1 = (x_1, y_1)$ and $\mathbf{x}_2 = (x_2, y_2)$ and one prediction point $\mathbf{x}'_1 = (x'_1, y'_1)$.

Similarly, the predicted vector is defined as follows:

$$\mathbf{Z}_p(t) = \begin{pmatrix} \mathbf{z}_p(t + \underline{\tau}_p) \\ \vdots \\ \mathbf{z}_p(t + \overline{\tau}_p) \end{pmatrix}. \tag{2.2}$$

The spatial locations of the measured and predicted FSE (i.e. $\mathbf{x}_{i=1,\ldots,M_o}$ and $\mathbf{x}'_{j=1,\ldots,M_p}$), the choice of $\underline{\tau}_o$, $\overline{\tau}_o$, $\underline{\tau}_p$ and $\overline{\tau}_p$, as well as the sampling time, together define the observation and prediction configuration. In the example shown in figure 2, $\overline{\tau}_o = 0$, implying that observations are taken into account up to present time $\tau = 0$. Regarding predictions, it is useful to distinguish further forecasts *per se* from reconstructions, which are predictions for past time horizons.

Typically, when studying wave predictability, the observation configuration (defined by locations $\mathbf{x}_{i=1\ldots M_o}$ and limiting observation times $\underline{\tau}_o$ and $\overline{\tau}_o$) is fixed, while the prediction locations and instants are varied across the spatial and temporal domain, to determine where, and when, the chosen measurement configuration allows accurate wave forecasts, such as in the example diagrams of figure 1.

3. A Gaussian probabilistic wave forecasting framework

3.1. *Assumptions*

Considering the wave field as random, stationary and Gaussian, the wave prediction problem can be formulated in a rigorous statistical framework, as shown by Simanesew *et al.* (2017) and Mérigaud & Ringwood (2018*a*). The assumptions underlying the approach are commonplace in oceanography and marine engineering since the 1950s (Pierson 1952).

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

(i) Assumption 1. The FSE η is a zero-mean and Gaussian random field, which means that any finite set of wave elevation samples $\eta(\mathbf{x}_1, t_1)$, $\eta(\mathbf{x}_2, t_2)$, $\cdots$ $\eta(\mathbf{x}_N, t_N)$, taken at N positions and instants, follows a multivariate Gaussian distribution. The Gaussian nature of ocean waves is verified, at least to a good approximation, in the majority of sea conditions of interest for marine operations (Ochi 2005), which excludes major storms and shallow water waves (with marked peak-to-trough asymmetry), as well as specific phenomena such as wave breaking, whitecapping and rogue waves.

(ii) Assumption 2a. Examined over the time scales of interest (typically, a few seconds to 10 min), the wave field can be considered stationary, i.e. its statistical properties do not change substantially. Since the waves are also Gaussian, the field stationarity reduces to the stationarity of its first- and second-order moments (i.e. the wave process is wide-sense stationary).

(iii) Assumption 2b. The wave field is statistically homogeneous over the spatial region considered.

Although only FSE observations and predictions are considered in this work, the approach below applies to any other variable, related to the FSE through a linear transfer function, such as pressure, particle or surface velocity, buoy or ship motion in any degree of freedom, or any combination of the above.

Under the Gaussian assumption 1, the wave field is entirely characterised by its covariance function, defined as $\mathbf{r}_{\eta\eta}(\mathbf{x}_1, \mathbf{x}_2, t_1, t_2) := \langle \eta(\mathbf{x}_1, t_1)\eta(\mathbf{x}_2, t_2)\rangle$, where $\langle . \rangle$ denotes the mathematical expectation. Because of the stationarity and homogeneity properties of assumptions 2a and 2b, the covariance function depends only on the relative position and relative instant (lag), and therefore can be written as $\mathbf{r}_{\eta\eta}(\Delta\mathbf{x}, \tau)$ where $\Delta\mathbf{x} = \mathbf{x}_2 - \mathbf{x}_1$ and $\tau = t_2 - t_1$. The covariance function can be calculated from the wave spectrum through an inverse Fourier transform, the Wiener–Khinchin relation (Ochi 2005), which, for a directional sea, takes the following form:

$$\mathbf{r}_{\eta\eta}(\Delta\mathbf{x}, \tau) = \frac{1}{2\pi} \int_0^\infty \int_{\theta=0}^{2\pi} S_{\eta\eta}(\omega, \theta) e^{j[\mathbf{k}(\theta,\omega).\Delta\mathbf{x} - \omega\tau]} \mathrm{d}\theta \mathrm{d}\omega \tag{3.1}$$

where $\mathbf{k}(\theta, \omega)$ is a vector with direction θ and magnitude k, related to ω through a dispersion relation, and $S_{\eta\eta}(\omega, \theta)$ is the frequency-direction wave spectrum.

3.2. *Optimal predictor calculation*

From assumptions 1 and 2, the observed and predicted vector time series $\mathbf{z}_o(t)$ and $\mathbf{z}_p(t)$, defined in § 2, jointly form a stationary, Gaussian, multi-dimensional process, characterised by the autocovariance matrix-valued functions $\mathbf{r}_{oo}(\tau) := \langle \mathbf{z}_o(t)\mathbf{z}_o^\mathsf{T}(t+\tau)\rangle$ and $\mathbf{r}_{pp}(\tau) := \langle \mathbf{z}_p(t)\mathbf{z}_p^\mathsf{T}(t+\tau)\rangle$, and the cross-covariance function $\mathbf{r}_{op}(\tau) := \langle \mathbf{z}_o(t)\mathbf{z}_p^\mathsf{T}(t+\tau)\rangle = \mathbf{r}_{po}^\mathsf{T}(-\tau)$ (Papoulis & Pillai 2002). Therefore, at any time t, the observed values $\mathbf{Z}_o(t)$ from (2.1) and the values to be predicted, $\mathbf{Z}_p(t)$ from (2.2), are zero-mean, jointly Gaussian vectors, entirely characterised by their covariance matrices, denoted as Σ_{oo}, Σ_{pp} and $\Sigma_{op} = \Sigma_{po}^\mathsf{T}$.

The covariance matrices Σ_{oo}, Σ_{pp} and $\Sigma_{op} = \Sigma_{po}^\mathsf{T}$ do not depend on t, and they are block-Toeplitz, constructed as follows:

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

$$\Sigma_{oo} = \begin{pmatrix} \mathbf{A}_{1,1} & \cdots & \mathbf{A}_{1,N_o} \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{N_o,1} & \cdots & \mathbf{A}_{N_o,N_o} \end{pmatrix} \in \mathbb{R}^{M_o N_o \times M_o N_o}$$
$$\Sigma_{pp} = \begin{pmatrix} \mathbf{B}_{1,1} & \cdots & \mathbf{B}_{1,N_p} \\ \vdots & \ddots & \vdots \\ \mathbf{B}_{N_p,1} & \cdots & \mathbf{B}_{N_p,N_p} \end{pmatrix} \in \mathbb{R}^{M_p N_p \times M_p N_p} \tag{3.2}$$
$$\Sigma_{op} = \begin{pmatrix} \mathbf{C}_{1,1} & \cdots & \mathbf{C}_{1,N_p} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{N_o,1} & \cdots & \mathbf{C}_{N_o,N_p} \end{pmatrix} \in \mathbb{R}^{M_o N_o \times M_p N_p}$$
$$\Sigma_{po} = \Sigma_{op}^{\mathsf{T}}$$

with blocks $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ defined as

$$\begin{cases} \mathbf{A}_{i,j} = \mathbf{r}_{oo}(\tau_o(j) - \tau_o(i)), \ \forall (i,j) \in [1 \,..\, N_o]^2 \\ \mathbf{B}_{i,j} = \mathbf{r}_{pp}(\tau_p(j) - \tau_p(i)), \ \forall (i,j) \in [1 \,..\, N_p]^2 \\ \mathbf{C}_{i,j} = \mathbf{r}_{op}(\tau_p(j) - \tau_o(i)), \ \forall (i,j) \in [1 \,..\, N_o] \times [1 \,..\, N_p] \end{cases}.$$

Under such assumptions, the mean value of $\mathbf{Z}_p(t)$ given the available observations $\mathbf{Z}_o(t)$, denoted $\mu_{p|o}(t)$, constitutes the optimal predictor for $\mathbf{Z}_p(t)$, in a mean-square error sense, given the observations $\mathbf{Z}_o(t)$. Here $\mu_{p|o}(t)$ is calculated as follows (Papoulis & Pillai 2002):

$$\mu_{p|o}(t) = \mathbf{P}\mathbf{Z}_o(t). \tag{3.3}$$

Here $\mathbf{P}$ is called the optimal prediction matrix, and is calculated as follows:

$$\mathbf{P} = \Sigma_{po}\Sigma_{oo}^{\dagger} \tag{3.4}$$

and the $\dagger$ (dagger) superscript denotes the inverse – if Σ_{oo} is full-rank – or a pseudoinverse – if Σ_{oo} is not full-rank, which indicates that the information in $\mathbf{Z}_o(t)$ is statistically redundant, as shown by Mérigaud & Ringwood (2018*a*).

The corresponding prediction error, $\mathbf{E}(t) := \mu_{p|o}(t) - \mathbf{Z}_p(t)$, is zero-mean Gaussian, characterised by the optimal prediction error covariance matrix $\Sigma_{p|o}$ calculated as follows:

$$\Sigma_{p|o} = \Sigma_{pp} - \Sigma_{po}\Sigma_{oo}^{\dagger}\Sigma_{op}. \tag{3.5}$$

Hence, at each time step t, the optimal prediction for times $(t + \underline{\tau}_p \ldots t + \overline{\tau}_p)$ is obtained through a simple multiplication, by $\mathbf{P}$, of the observed data at times $(t + \underline{\tau}_o \ldots t + \overline{\tau}_o)$.

The only remaining challenge is the calculation of the covariance matrices (3.2) at discrete lag values corresponding to the prediction configuration, but from § 3.1, the entries of $\mathbf{r}_{oo}$, $\mathbf{r}_{pp}$ and $\mathbf{r}_{op}$ may be computed for any lag τ using the wave spectrum information, through (3.1). For example, entry i, j of $\mathbf{r}_{oo}(\tau)$ is $\mathbf{r}_{oo,i,j}(\tau) = \mathbf{r}_{\eta\eta}(\mathbf{x}_j - \mathbf{x}_i, \tau)$, where $\mathbf{r}_{\eta\eta}$ is given in (3.1). Therefore, over short time scales below a few tens of minutes, the optimal prediction matrix $\mathbf{P}$ and the optimal prediction error covariance matrix $\Sigma_{p|o}$ are both time-invariant, and only need to be updated as the wave condition (i.e. the spectrum) evolves, thus slowly modifying the statistical properties of all measured and predicted processes. The prediction error covariance matrix provides the expected prediction performance for each time horizon. Finally, although the physics of wave propagation is not explicit in

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

the prediction (3.4), it is implicitly modelled when calculating the covariance structure, through the dispersion relation which appears in (3.1).

3.3. *Accounting for current*

The presence of current, or, equivalently, of a non-zero forward speed if the measurement system is mounted on a moving vessel, result in a Doppler shift, which implies that waves with a given wavelength and direction have an apparent frequency which deviates from their intrinsic frequency (i.e. their frequency in the absence of current, typically determined by the linear dispersion relation). For example, a single wave component is written as $\eta(\mathbf{x}, t) = \Re\{\hat{\eta} e^{j[\mathbf{k}.(\mathbf{x}-\mathbf{U}t)-\omega t]}\}$, where $\mathbf{U}$ is the current vector, so that an observer moving at speed $\mathbf{U}$ would see a 'usual' progressive wave $\eta(\mathbf{x}, t) = \Re\{\hat{\eta} e^{j[\mathbf{k}.\mathbf{x}-\omega t]}\}$. Accordingly, the Wiener–Khinchine relation (3.1) must be amended as follows to account for current:

$$\mathbf{r}_{\eta\eta}(\Delta\mathbf{x}, \tau) = \frac{1}{2\pi} \int_0^{\infty} \int_{\theta=0}^{2\pi} S_{\eta\eta}(\omega, \theta) e^{j[\mathbf{k}(\theta,\omega).\Delta\mathbf{x} - [\mathbf{k}(\theta,\omega).\mathbf{U}+\omega]\tau]} \mathrm{d}\theta \mathrm{d}\omega. \tag{3.6}$$

Note that in (3.6), the spectrum is the intrinsic spectrum, i.e. the spectrum which would be measured if there were no current. It can be found from (3.6) that wave measurements at different locations still form together a stationary, Gaussian multidimensional process, with covariance depending only on the lag between measurement instants and on relative positions. Only the computation of the process covariance, by using (3.6) instead of (3.1), is affected.

The above formulations (3.1) and (3.6), to obtain point-to-point covariance values, rely on the assumption of homogeneous current and wave fields. Such a model could be reasonable at scales below 10 km in the open sea Ardhuin *et al.* (2017), but definitely not in intermediate and shallow depths with locally complex bathymetry or topography, which can be found in coastal areas. In the latter cases, the current, the dispersion relation, as well as the wave spectrum, could be varying within the spatial domain considered, which would considerably complicate the calculation of point-to-point covariance. In particular, the fact would no longer hold, that the spatial dependency of the covariance function $\mathbf{r}_{\eta\eta}$ is only in the relative position $\Delta\mathbf{x}$, thus requiring a formulation depending on absolute positions, i.e. $\mathbf{r}_{\eta\eta}(\mathbf{x}_1, \mathbf{x}_2, \tau)$. Note, however, that if the covariance $\mathbf{r}_{\eta\eta}(\mathbf{x}_1, \mathbf{x}_2, \tau)$ can still be computed, either from direct calculations, or from observations, then the rest of the statistical approach, in (3.2) to (3.6), remains unchanged.

4. Two-dimensional wave predictability

In this section, unidirectional waves are considered, whereby all the wave energy is assumed to propagate along the x direction. Based on an idealised wave spectrum, and using the approach outlined in § 3, predictability is studied in a rigorous statistical framework, along with a critical comparison with the traditional deterministic prediction diagram. The waves are assumed characterised by a JONSWAP spectrum (Hasselmann *et al.* 1973) with peak period $T_p = 10$ s, significant wave height $H_{m_0} = 1$ m and peak enhancement factor $\gamma = 1$. The water depth is 50 m (water depth is used to calculate the group velocity, as well as the dispersion relation in the calculations of (3.1) and (3.6)). Throughout this section and the following one, space and time coordinates are normalised by λ_p (the peak wavelength) and T_p, respectively. Additionally, because all prediction performance metrics are normalised by the FSE variance, and because linear waves are

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

assumed, the value of H_{m_0} has no influence on the theoretical results presented in this section as well as § 5.

Note that, in this section and the following one, all results are directly computed from the wave spectrum, using the Wiener–Khinchine relation (3.1), observation and prediction covariance matrices (3.2), and optimal error covariance matrix (3.6), without the need for any numerical simulation.

4.1. *Temporal observation interval*

First, FSE measurements are taken in a single observation location $x = 0$, for five peak wave periods into the past, i.e. $\underline{\tau}_o = -5T_p$ and $\overline{\tau}_o = 0$, with a sampling time $T_p/10$. The probabilistic approach of § 3 is applied for many prediction locations downwave along the x axis, with a prediction time interval defined by $\underline{\tau}_p = -5T_p$ and $\overline{\tau}_p = 10T_p$. For each prediction location, the prediction uncertainty is quantified in terms of NMSE, obtained by dividing the diagonal terms of the prediction covariance matrix $\Sigma_{p|o}$ (3.6) by m_0, the variance of η.

The NMSE map, resulting from the above procedure, is first visualised in figure 3(a) in the distance–time continuum. By drawing the NMSE contour lines for specified thresholds as in figure 3(b), the probabilistic analogues to deterministic 'predictable zones' (DPZs) are obtained. From this point on, PPZs shall always be defined with respect to an NMSE threshold, expressed in %, for example 'the 10% PPZ'. Those PPZ contour lines, as can be observed in figure 3(b), form triangular zones, similarly to the deterministic theory illustrated in figure 1(a). There remains, like in DPZs, to assess whether each PPZ can be meaningfully related to two limiting group velocities, representative of the wave spectrum.

Beginning with the slowest group velocity, which determines the lower limit of each triangular predictable zone, it is found that a simple energy criterion, illustrated in figures 3(b) (dashed lines) and 3(c), provides a satisfactory agreement: considering the z % PPZ, the lower limiting group velocity $c_g^{z\,\%}$ is found at the frequency $f_{z\,\%}$, above which z % of the spectrum integral area is confined. To the best of the authors\ knowledge, this is the first time a quantitative criterion is proposed, to relate a particular cutoff frequency to the predictable zone.

Now examining the fastest group velocity, which should determine the upper limit of each triangular predictable zone, applying a similar energy criterion does not provide convincing results, as can be appreciated from the dashed–dotted lines in figure 3(b). In fact, the group velocities, which can be calculated from the lowest (subpeak) frequency content of the spectrum, provide exceedingly conservative estimates of the predictable zones, which is a result similar to those of Law *et al.* (2020), Hlophe *et al.* (2021) and Huchet *et al.* (2023), discussed in the Introduction. In contrast, taking the peak group velocity to determine the upper limit of the triangular zones yields satisfactory agreement for all threshold values considered here, as can be appreciated in figure 3(d) (cyan, solid line). The latter result, which is novel and provides more optimistic prediction horizons, also differs slightly from those of Naaijen *et al.* (2014), where the limiting upper slopes at different error thresholds appear significantly different from each other.

This can be understood by going back to the line of reasoning which leads to deterministic prediction diagrams. For a wave group to be predicted, it must have travelled through the observation location $x = 0$ during the observation time interval. Therefore, for a point (x, t) to belong to the predictable zone, all wave groups, which together contribute to the wave signal (x, t), must have travelled through the observation location during the observed time interval. Such a train of thoughts implies that the wave process is decomposed into many wave groups, or equivalently, that the wave spectrum has been

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

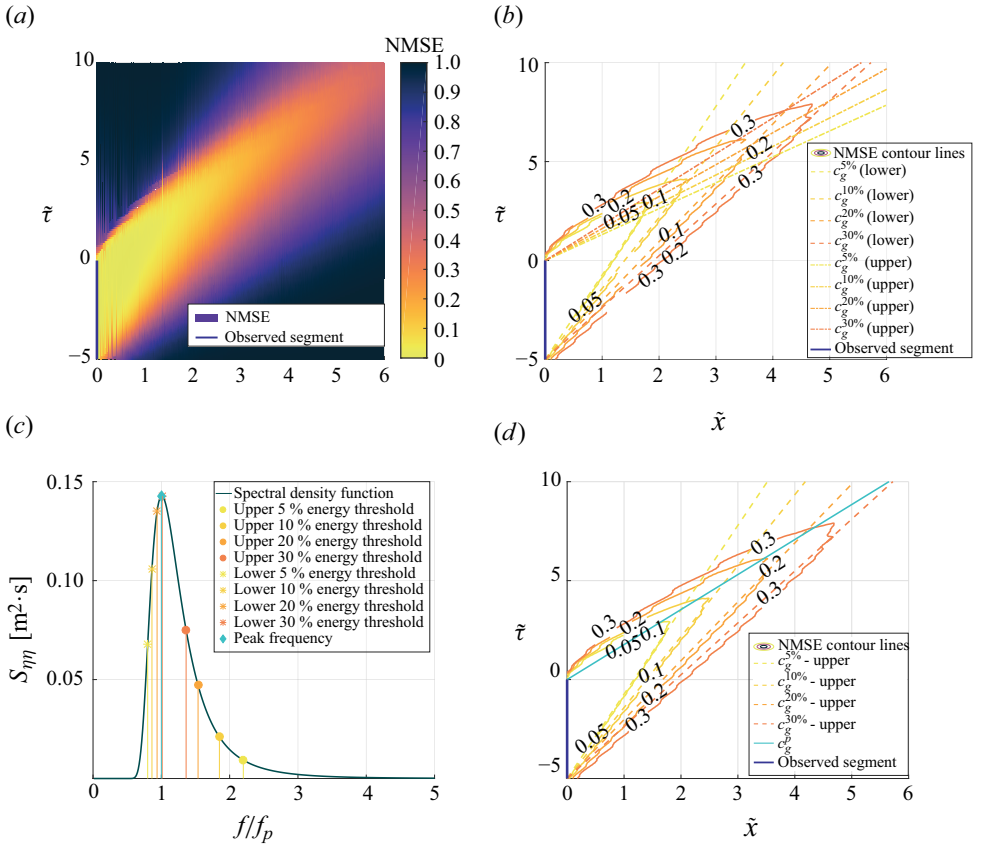


Figure 3. Two-dimensional PPZs in the distance–time continuum, where $\tilde{x} = x/\lambda_p$ and $\tilde{\tau} = \tau/T_p$: (*a*) NMSE map; (*b,d*) PPZs defined by NMSE contour lines. The PPZ lower slopes are retrieved from an energy criterion based on the upper frequency range: (*c*) let $f_{z\%}$ be the frequency above which z % of the spectrum energy are confined, and $c_g^{z\%}$ the corresponding group velocity. Then the z % PPZ is bounded from below by a slope $\approx 1/c_g^{z\%}$, see (*b*) and (*d*). In contrast, the PPZ upper limit is poorly approximated by (*b*) a similar energy criterion in the lower frequency range, but well approximated by (*d*) the group velocity at peak frequency.

partitioned into many, sharply concentrated spectra (each of which corresponding to a random wave group). In the upper frequency region, the spectrum decreases gradually, and therefore such a partitioning can be seen as a continuum of many, small wave groups, which would explain why the criterion based on the energy integral is satisfactory. In contrast, because the spectrum decreases abruptly below the peak frequency, the spectrum peak can be seen as the spectrum of only one random wave group, centred around the peak frequency, and accounting for a large fraction of the total spectrum energy. Such an interpretation is confirmed in appendix A, by constructing a fictitious spectrum, with a more gradual decrease of the energy content below the peak frequency: In the latter case, the upper PPZ limits can also be inferred from an energy criterion in the lower frequency range, similarly to the lower PPZ limit.

4.2. *Spatial observation interval*

In § 4.1, the prediction is initialised using wave history at a single point. However, many RTWF methods rather rely on a spatial initialisation of the prediction, whereby a snapshot of the sea surface, at a single instant, is used as an input to the forecasting model.

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

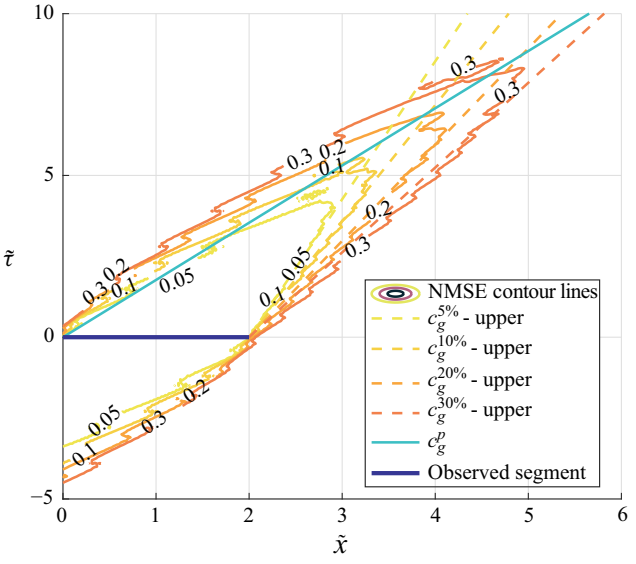


Figure 4. Two-dimensional PPZs for a \snapshot\ measurement configuration. Lower and upper limits of PPZs can be retrieved from the same group velocities as in the configuration of figure 3. See figure 3(*c*) for colour interpretation of the straight lines.

The very same methodology outlined in §§ 2 and 3 is followed to study such spatially distributed observations. More specifically, the observation time interval reduces to a single instant, i.e. $\underline{\tau}_o = \overline{\tau}_o = 0$ in the notations of § 2, while the list of observation locations $\mathbf{x}_1, \ldots \mathbf{x}_{M_o}$ spans the spatial interval of the observation snapshot. figure 4 shows the results obtained for observation points spanning two peak wavelengths, with 20 points per peak wavelength.

As can be appreciated in the figure, in the spatially distributed scenario, the PPZs can be approximated using the same two group velocities as in the 'time-history' scenario: the peak group velocity, together with the group velocity of the upper cutoff frequency, associated with the prescribed NMSE threshold.

In summary, the probabilistic approach is consistent with the two-dimensional deterministic prediction diagram, which the results of Mérigaud & Ringwood (2018*b*) already suggested in a qualitative way. More specifically, the present results confirm that, in a two-dimensional scenario, predictable zones can be constructed from representative group velocities in the wave spectrum, with the novel clarifications that the fastest group velocity to take into account should be the peak group velocity, and that the slowest group velocity is determined quantitatively, based on an energy criterion, depending on the threshold retained to define the predictable zone.

5. Three-dimensional wave predictability

For the three-dimensional (omnidirectional) case, the spectrum frequency distribution remains identical to that of § 4, while the directional distribution is given by the Mitsuyasu spreading function (Goda 1975), with the mean wave propagation direction θ_p aligned with the x axis ($\theta_p = 0$), and a parameter s_{max} governing the directional spreading, with smaller values of s_{max} representing widely spread energy content. Two representative, realistic values of s_{max} are studied: 10 (wind waves) and 75 (swell with long decay distance).

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

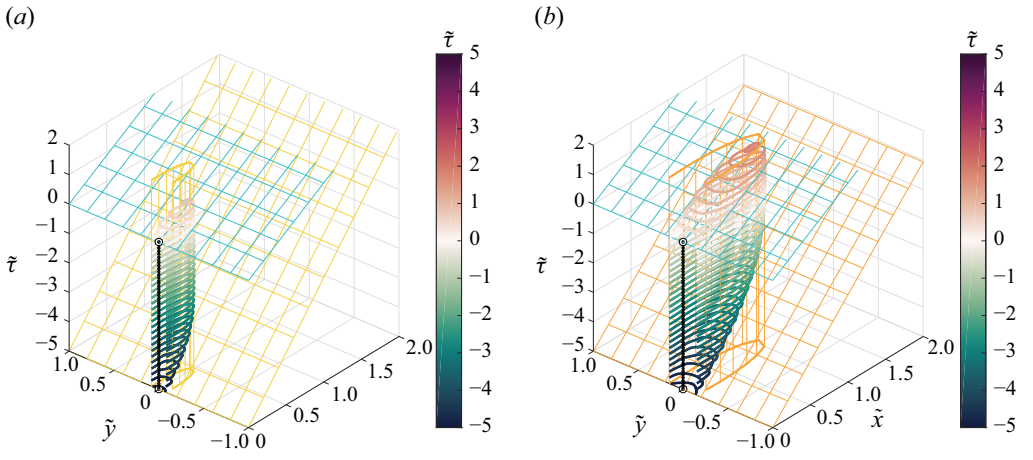


Figure 5. 10 % and 20 % PPZs with $s_{max} = 75$ and a single observation point in $(x, y) = (0, 0)$. Waves are observed for five peak periods into the past. Vertical spacing between $(\tilde{x}, \tilde{y})$ contour lines: $\Delta\tilde{\tau} = 1/5$. The vertical cylindrical domains, generated using 10 % and 20 % contour lines of figure 6(*c*), are also represented. (*a*)10% probabilistic predictable zones. Upper and lower limiting slopes of the two-dimensional case are represented by planes with slopes $1/c_g^p$ and $1/c_g^{10\%\text{upper}}$, respectively. See Figure 3c for color interpretation (*b*) 20\% probabilistic predictable zones. Upper and lower limiting slopes of the two-dimensional case are represented by planes with slopes $1/c_g^p$ and $1/c_g^{20\%\text{upper}}$, respectively. See Figure 3c for color interpretation.

The methodology is identical to that of the two-dimensional case, but the prediction point location must vary in the two horizontal directions, to build a three-dimensional map of the NMSE. Since the prediction performance now varies along three dimensions, the predictable zones are volumes, which makes their visual representation more complex. For that purpose, it is chosen to represent the $(\tilde{x}, \tilde{y})$ contour lines of the predictable zone, for a set of prediction horizons $\tilde{\tau}$, where $\tilde{x} = x/\lambda_p$, $\tilde{y} = y/\lambda_p$ and $\tilde{\tau} = \tau/T_p$ are normalised spatial and temporal coordinates.

5.1. *Single observation point*

It is first examined how the waves can be predicted using a single measurement location, like in the two-dimensional case of § 4.1. Considering the spectrum with the least directional spreading ($s_{max} = 75$), and observations of five peak wave periods into the past, figures 5(*a*) and 5(*b*) represent the 10 % and 20 % PPZs. Also shown on the figures, are the two planes which would limit the predictable zone, if the waves were unidirectional (see figure 3*d* for colour interpretation), as well as a vertical, cylindrical envelope, of which the meaning is discussed next.

As can be appreciated in figure 5, the spatial and temporal extent of the predictable zone is dramatically reduced, with respect to the unidirectional case. Admittedly, at very short distances from the observation point (less than 0.5 λ_p in this case), the upper and lower limits of the predictable zone are well approximated by the two limiting planes. However, the predictable zone rapidly shrinks with the distance from the prediction point, so that the maximum time horizon for the 10 % NMSE criterion is smaller than T_p. By loosening the prediction quality criterion to 20 % (figure 5*b*), the predictable zone is expanded, but still collapses to zero more rapidly than the two limiting planes would predict.

It is thus clear from figure 5 that directional spreading has a strong adverse effect on wave predictability, when a single measurement location is employed. In order to better isolate the effect of directional spreading from that of wave dispersion, we can consider the

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

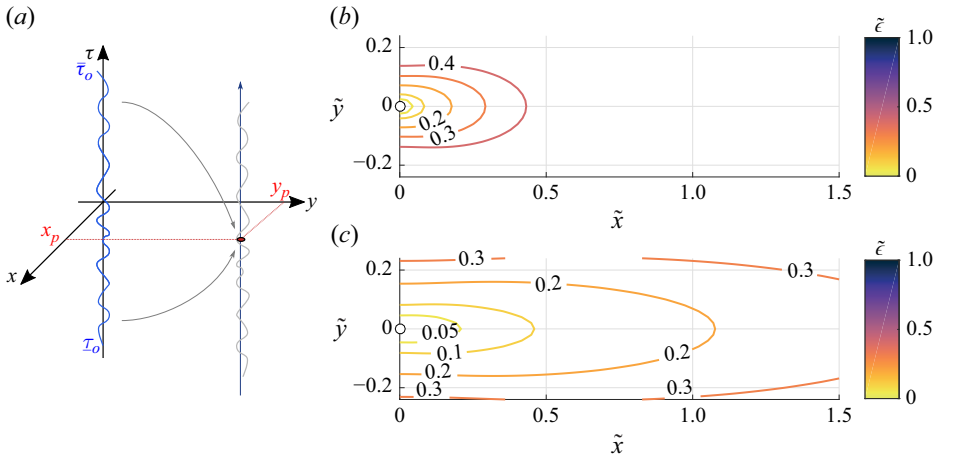


Figure 6. Schematic diagram of wave prediction at the present time in (x_p, y_p), using both past and future observations in $x = y = 0$ (figure 6*a*); corresponding achievable NMSE contour lines in the horizontal plane when the observation time interval grows to infinity, for two different spreading parameters s_{max} (figures 6*b* and 6*c*). (*b*) NMSE contour lines, $s_{max} = 10$ (*c*) NMSE contour lines, $s_{max} = 75$.

following hypothetical scenario, whereby waves are predicted at a given instant (without loss of generality, the present instant $\tau = 0$), at a given location (x_p, y_p) of the horizontal plane, using waves observed in $x = y = 0$ measured both in the past and in the future, as sketched in figure 6(*a*). Furthermore, we can make the observation duration grow to infinity, i.e. $\underline{\tau}_o \to -\infty$ and $\overline{\tau}_o \to \infty$. In practice, $\overline{\tau}_o$ and $\underline{\tau}_o$ are set to large enough, but finite values $\overline{\tau}_o = -\underline{\tau}_o = 20\ T_p$: beyond $\overline{\tau}_o = -\underline{\tau}_o = 10\ T_p$ there is little or no change in the results, within the spatial domain considered. In other words, we quantify how well the wave signal at a location (x_p, y_p) can be reconstructed, knowing the whole time history, both past and future, of the wave signal in (0, 0). We also hereby suppress the time dependency of the problem.

By making (x_p, y_p) vary across the horizontal plane and calculating the corresponding NMSE, the contour lines of figures 6(*b*) and 6(*c*) are obtained. Note that, because the results are symmetric with respect to the y axis, only positive x values are shown.

figure 6 shows that the achievable prediction performance decreases with the distance to the observation location, most rapidly in the direction transverse to the main wave direction, and least rapidly in the main wave direction. The decrease is also stronger for widely spread seas (figure 6*b*) than for long-crested seas (figure 6*c*). Clearly, the prediction NMSE in figure 6, obtained with an infinitely long observation time interval, should be smaller than the NMSE obtained with observations, limited to a finite time interval. Therefore, each contour line in figures 6(*b*) and 6(*c*) defines a spatial domain, which limits the extent of the corresponding PPZ for a single observation location, regardless of the observed time segment.

To visualise the latter statement, figures 5(*a*) and 5(*b*) also show a representation of the 10 % (respectively, 20 %) contour lines of figure 6(*c*), extended vertically to form a cylindrical domain. As evidenced in the figure, the three-dimensional predictable zone for a single observation location can be constructed from three ingredients: two limiting group velocities, as in the unidirectional case, and a cylindrical envelope, centred around the observation location, which limits the PPZ spatial extent under the effect of directional spreading.

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

Considering that the results of figure 5 are obtained in the most favourable scenario $s_{max} = 75$, it seems unlikely that forecasts of any practical interest can be achieved beyond one wave period into the future, when only one observation location is used. Another striking feature of those results is that they do not resemble the fan-shaped, DPZs, exemplified in figure 1(*b*). It is therefore important to understand the difference between the results of the present theory, and those of the deterministic approach.

5.2. *Comparison with the deterministic theory*

For the single-observation location scenario, DPZs calculated following Qi *et al.* (2018*b*) are compared with the PPZs of the present modelling framework. A first challenge is that, in the present work, unlike in the deterministic approach, PPZs are defined with respect to a prescribed NMSE level. In the following examples, without loss of generality, a 10 % NMSE threshold is chosen as a reasonable approximation to a 'perfect' prediction. Another challenge is to choose the two limiting frequencies of the deterministic approach, so as to make the deterministic and probabilistic cases as closely comparable as possible. To that end, the deterministic approach is implemented using the peak frequency f_p as the lower cutoff, while the upper cutoff frequency is set to $f_{10\%}$, accordingly to the results of § 4. This ensures that the deterministic and PPZs will be limited by the same two group velocities. Finally, the three-dimensional deterministic theory requires that all the spectrum energy propagates between two cutoff directions, $\pm\Delta_\theta$, where $\Delta_\theta \leqslant \pi/2$. Thus, for a fair comparison, the probabilistic results are also calculated based on a directionally truncated spectrum $S_{\text{trunc}}(f, \theta)$. More specifically, the JONSWAP–Mitsuyasu directional spectrum $S(f, \theta)$ of § 5.1 (with the most optimistic parameter $s_{max} = 75$) is directionally truncated,

$$S_{\text{trunc}}(f, \theta) = \begin{cases} S(f, \theta) & |\theta| < \Delta_\theta \\ 0 & |\theta| \geqslant \Delta_\theta \end{cases}. \tag{5.1}$$

The value $\Delta_\theta = \pi/6$ is chosen in the calculation of both PPZs and DPZs, and two different observation duration values 5 and 10 T_p are considered. The resulting 10 % PPZs are shown in figures 7(*a*) and 7(*c*), while the corresponding DPZs are shown in figures 7(*b*) and 7(*d*). The two limiting planes with slopes $1/c_g^p$ and $1/c_g^{10\,\%\text{upper}}$, which bound the predictable zones in both cases, are also represented. It is immediately clear that the DPZs are not only different in shape, but also significantly larger in size than their analogues defined in a probabilistic sense. Using five peak wave periods of observations, the deterministic theory predicts that waves can be forecast accurately up to 3.5 peak wave periods into the future (figure 7*b*), versus only 1.5 following the probabilistic theory (figure 7*a*). Using 10 peak wave periods of observations, instead of five, doubles the achievable forecasting horizon in the deterministic theory (figure 7*d*), while it leaves it unchanged in the probabilistic theory (figure 7*c*).

In fact, it is an inherent property of the DPZ construction method, that the achievable time horizon T_{max} is proportional to the observation duration T. The relation is

$$T_{max} = \frac{T}{\dfrac{c_g^{max}}{c_g^{min}\cos\Delta_\theta} - 1} \tag{5.2}$$

which reduces to $T_{max} = T/(c_g^{max}/c_g^{min} - 1)$ in the unidirectional case (Qi *et al.* 2018*b*; Kim *et al.* 2023). Therefore, by increasing the observation duration at a single location, the forecasting horizon could, in principle, be increased to arbitrarily large values, at a downwave location farther and farther away from the observation point. The latter

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

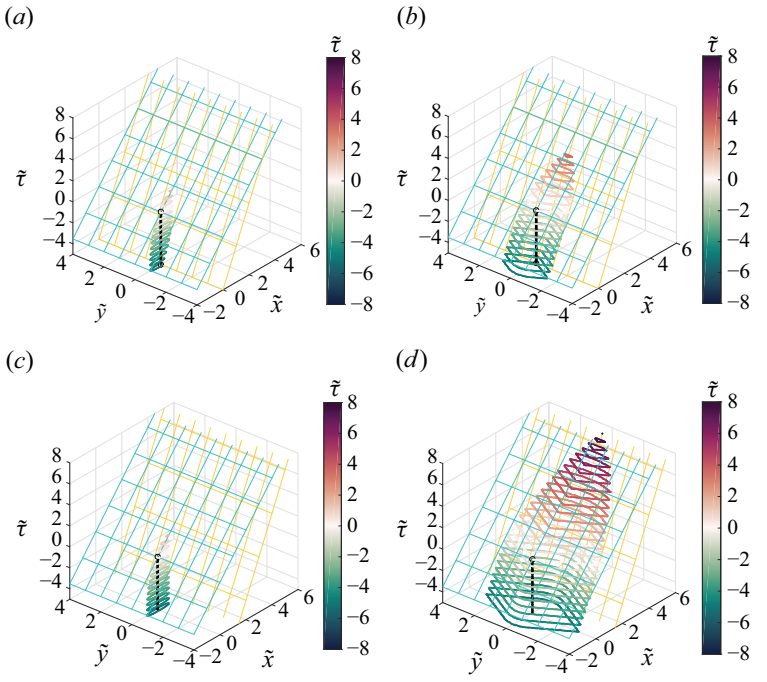


Figure 7. 10 % PPZs with $s_{max} = 75$ and a single observation point in $(x, y) = (0, 0)$. Vertical spacing between the contour lines: $\Delta\tilde{\tau} = 0.5$ $(0.5\ T_p)$. The directional wave spectrum is truncated for $|\theta| \geqslant \pi/6$. Upper and lower limiting slopes of the two-dimensional case are represented by planes with slopes $1/c_g^p$ and $1/c_g^{10\,\%\text{upper}}$, respectively. See figure 3(*c*) for colour interpretation. (*a*) 10% probabilistic predictable zones. Waves are observed for 5 peak periods into the past (*b*) Deterministic predictable zones. Waves are observed for 5 peak periods into the past (*c*) 10% probabilistic predictable zones. Waves are observed for 10 peak periods into the past (*d*) Deterministic predictable zones. Waves are observed for 10 peak periods into the past.

property is in clear contradiction with the probabilistic theory, in which predictable zones are spatially bounded under the effect of directional spreading, following the error maps shown in figure 6.

In the deterministic theory, the construction of the three-dimensional predictable zone is based on the intersection of two-dimensional predictable zones, with incident wave directions spanning the interval $[-\Delta_\theta, \Delta_\theta]$. However, for this construction to hold in a real-sea scenario, the forecasting model should be able to decompose, within the observed wave signal, the contributions from each wave direction, before propagating them in time and space. In a realistic sea, those contributions are randomly mixed at the observation location, so that it is not possible to perfectly decompose the wave signal into its directional components. Such a 'directional mixing' effect is, implicitly, modelled in the probabilistic theory, but not in the deterministic representation.

Therefore, in the present probabilistic model, with a single observation location, increasing the forecasting horizon cannot be achieved by looking further into the past. Instead, resolving for the different incoming wave directions can only be achieved by spatially extending the domain of observation, which is the topic of the next section.

5.3. *A line of observation points*

Probabilistic predictable zones are now examined for a straight line of observation points, extending in the direction transverse to the peak wave direction θ_p. The line length is set

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

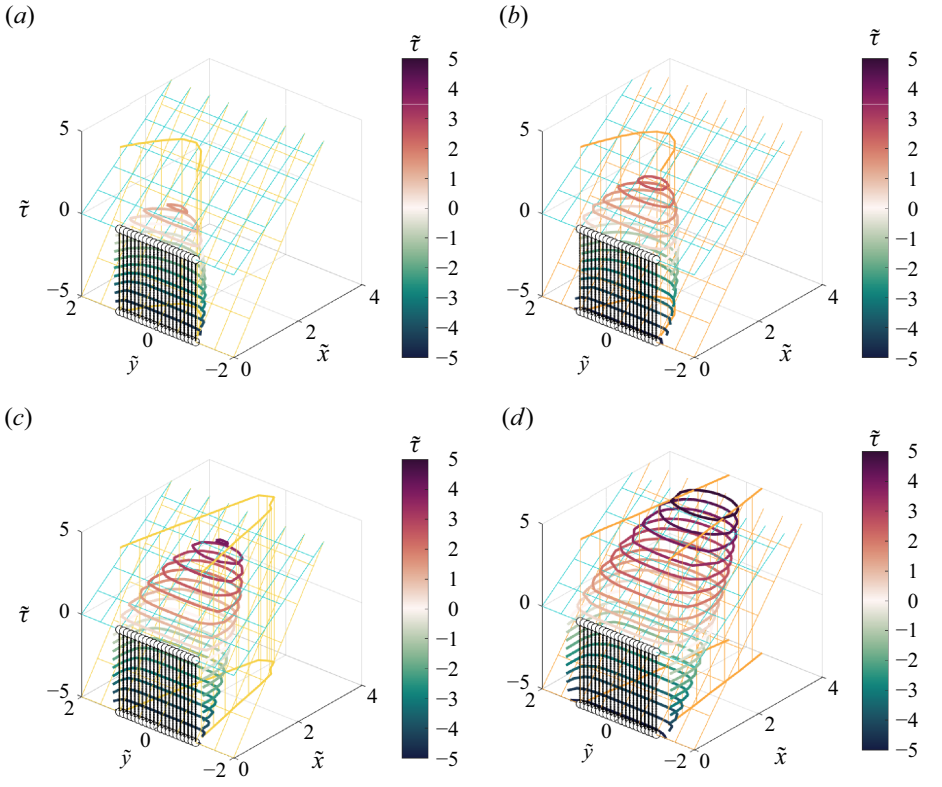


Figure 8. 10 % and 20 % PPZs with a line of observation points. Waves are measured for five peak wave periods into the past. Vertical spacing between the contour lines: $\Delta\tau = 0.5T_p$. Upper and lower limiting slopes of the two-dimensional case are represented by planes with slopes $1/c_g^p$ and $1/c_g^{10\,\%\text{upper}}$, respectively. The effect of directional spreading is approximated by cylindrical envelopes (vertically stretched 10% and 20% GOP contour lines). See figures 3(*c*) and 9 for colour interpretation. (*a*) 10 % PPZ, $s_{max} = 10$ (*b*) 20 % PPZ, $s_{max} = 10$ (*c*) 10 % PPZ, $s_{max} = 75$ (*d*) 20 % PPZ, $s_{max} = 75$.

to two peak wavelengths. Besides, it is found that no appreciable change can be observed in prediction error beyond a point density of 10 points per peak wavelength, which is therefore chosen throughout this section. The waves are observed for five peak wave periods into the past, with a time step of $T_p/10$.

figure 8 shows the 10 % and 20 % PPZs, for two different values of the spreading parameter s_{max}, along with the two limiting planes corresponding to the two-dimensional case. In comparison with the single-point configuration, the spatial information provided by the line of measurements overcomes the adverse effect of directional mixing, allowing prediction horizons ranging from 1.5 T_p (in the least favourable case $s_{max} = 10$) to 4 T_p ($s_{max} = 75$), for a 10 % NMSE criterion.

Yet, along with the two limiting wave groups, directional spreading maintains a shrinking effect onto the predictable zone geometry, although at a significantly slower rate than in the single-point case. In the same way that wave group travelling speeds govern upper and lower limits of predictable zones, it is interesting to outline a similar physical interpretation for the directional shrinking effect, based on wave group travelling directions.

To that end, we shall make the simplifying assumption that the wave field is made of random wave groups, centred around the spectrum peak frequency, and idealised as points

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

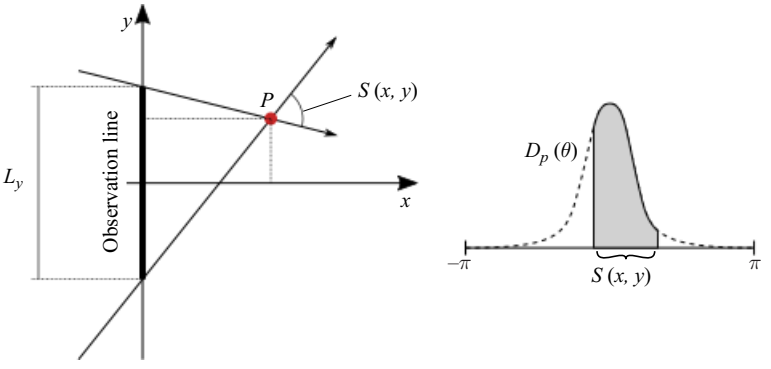


Figure 9. Geometrical calculation of the group observation probability for a point $P(x, y)$.

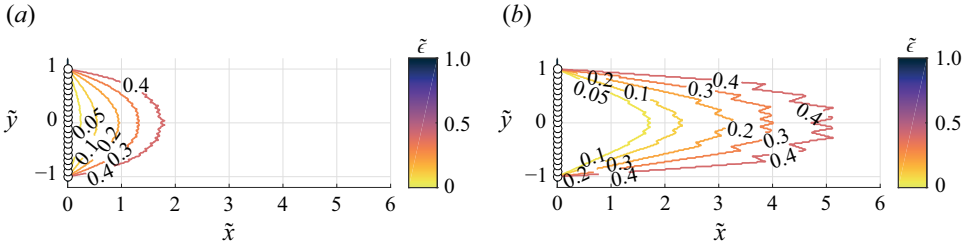


Figure 10. Group observation probability $\mathrm{GOP}(\tilde{x}, \tilde{y})$, i.e. fraction of directional energy crossing the observation line before reaching point $(\tilde{x}, \tilde{y})$, for a strip of measurements extending indefinitely in the time direction. The directional spreading function is taken at the peak frequency. (*a*) GOP contour lines, $s_{max} = 10$ (*b*) GOP contour lines, $s_{max} = 75$.

representing the wave group centre. The group travelling directions are randomly drawn in $[0; 2\pi]$, with a probability distribution represented by the directional spreading function. We also approximate the discrete measurement points as a continuous line, stretching partway in the transverse direction across a length L_y, as sketched in figure 10.

Next, we consider one such wave group, arriving at a point P of coordinates (x, y) at $t = 0$. A necessary condition, for this wave group to be predicted in (x, y), is that it has travelled through the line of observations some time in the past. The probability that it is the case can be calculated geometrically, from the directional distribution function at the peak frequency, as pictured in figure 10. With the notations of figure 10, the 'group observation probability' (GOP) is mathematically expressed as follows:

$$\mathrm{GOP}(x, y) = \int_{\mathcal{S}(x,y)} D_p(\theta)\mathrm{d}\theta \tag{5.3}$$

where $D_p(\theta)$ is the directional spreading function at the peak frequency, and $\mathcal{S}(x, y)$ is the angular sector comprised between the two extreme wave directions traversing both the observation line and the point P of coordinates (x, y). Repeating the calculation for all points (x, y) in the horizontal plane, we obtain a GOP map, representing the fraction of signal energy which has travelled through the observation line prior to reaching each point. The resulting contour lines, for selected GOP thresholds, can be visualised in figures 9(*a*) and 9(*b*) using normalised coordinates $\tilde{x} = x/\lambda_p$ and $\tilde{y} = y/\lambda_p$.

Stretching those GOP contour lines vertically, as seen in figure 8, provides, in combination with the two limiting planes, a reasonable approximation of the predictable

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

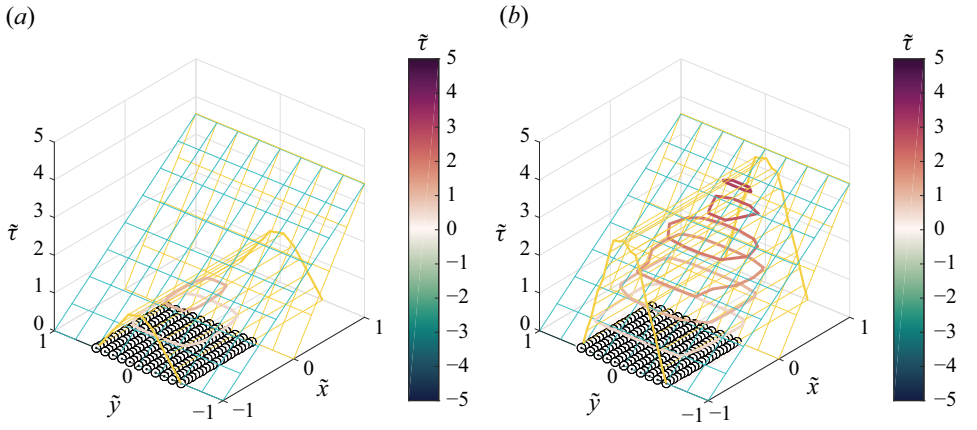


Figure 11. 10 % PPZs with a line of observation points. Waves are measured for five peak wave periods into the past. Vertical spacing between the contour lines: $\Delta\tau = 0.5T_p$. Upper and lower limiting slopes of the two-dimensional case are represented by planes with slopes $1/c_g^p$ and $1/c_g^{10\,\%\text{upper}}$, respectively. See figures 3(*c*) and 12 for colour interpretation. (*a*) 10 % PPZ, $s_{max} = 10$ (*b*) 10 % PPZ, $s_{max} = 75$.

zone satisfying the corresponding thresholds. Thus, although the GOP map certainly relies on strong approximations, it is straightforward to calculate, and it seems sufficient to provide an estimate of the spatial bounds which limit PPZs under the effect of directional spreading.

5.4. *A square snapshot of observation points*

As mentioned in § 4.2, many wave prediction methods rely on a snapshot of the sea surface to initialise the prediction, rather than on a long time history. Observations taken in a square area of the sea surface are considered here. The square side length is set to one peak wavelength. As in the previous section, the point density in the transverse direction is set to 10 points per peak wavelength, while in the longitudinal direction, a higher density of 20 points per peak wavelength is found preferable to achieve convergence of results.

Figure 11 shows the 10 % PPZs, for two different values of the spreading parameter s_{max}, along with the two limiting planes corresponding to the two-dimensional case. While being transported, and gradually shrinking in the longitudinal direction x between the two limiting planes, the PPZ also shrinks in the transverse direction y, thus taking a pyramidal aspect. The transverse shrinking rate is faster for short-crested waves, and more gradual for long-crested waves, as can be appreciated by comparing figures 11(*a*) and 11(*b*).

Similarly to the line arrangement, it is desirable to determine a simple and meaningful approximation to the directional spreading effect. To that end, we shall make the same simplifying assumptions of random wave groups, centred around the spectrum peak frequency, and idealised as points travelling at the peak group velocity. The travelling direction is randomly drawn in $[0; 2\pi]$, with a probability distribution represented by the directional spreading function. In addition, the domain of observation is simplified to a strip, extending indefinitely in the longitudinal direction, hence suppressing the x dependency of the problem, as sketched in figure 13.

We then examine a wave group reaching a point P of coordinates (y, τ), and we calculate its observation probability, i.e. the probability that it has crossed the observation domain at the measurement instant $\tau = 0$. Since the wave group travels at speed c_g^p, we know that, at $\tau = 0$, the group was located along the circle of radius τc_g^p, centred in y.

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

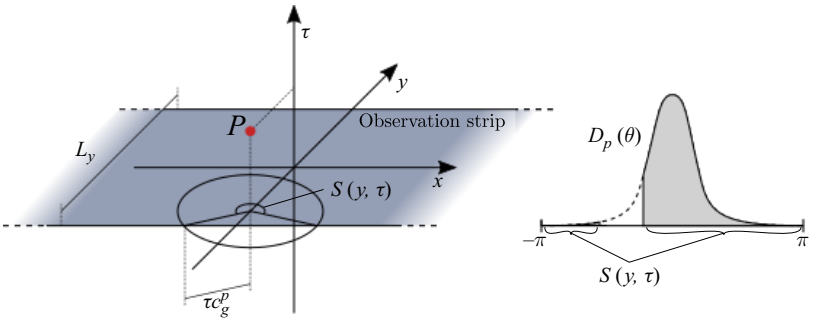


Figure 12. Geometrical calculation of the GOP for a point $P(y, \tau)$.

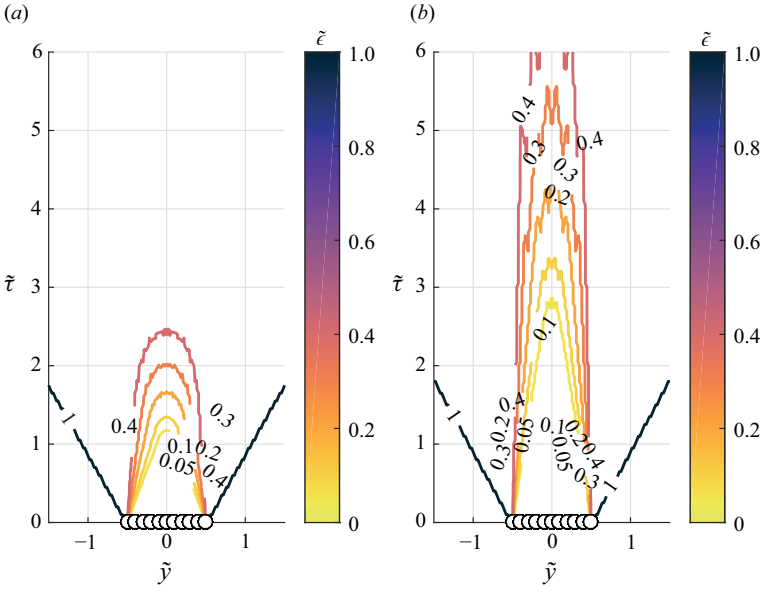


Figure 13. Group observation probability GOP($\tilde{y}$, $\tilde{\tau}$), i.e. fraction of directional energy crossing the strip of measurement at time $\tilde{\tau} = 0$ before reaching point ($\tilde{y}$, $\tilde{\tau}$), for a strip of measurements extending indefinitely in the x direction. The directional spreading function is taken at the peak frequency. (*a*) GOP contour lines, $s_{max} = 10$ (*b*) GOP contour lines, $s_{max} = 75$.

The wave group has thus been observed, if and only if its travelling direction belongs to the angular sector $\mathcal{S}$, of the circle intersection with the observation strip. The probability that it is so is calculated by integrating the directional spreading function, restricted to $\mathcal{S}$. Using the notations of figure 13, this is mathematically expressed as follows:

$$\mathrm{GOP}(y, \tau) = \int_{\mathcal{S}(y,\tau)} D_p(\theta)\mathrm{d}\theta \tag{5.4}$$

with

$$\mathcal{S}(y, \tau) = \{\theta : |\Re\{y + \tau c_g^p e^{j\theta}\}| \leqslant \frac{L_y}{2}\}. \tag{5.5}$$

Hence, similarly to the line configuration, a two-dimensional map GOP(y, τ) is obtained, representing, in each point (y, τ), the fraction of signal energy which has been

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

observed at the measurement instant. Drawing the corresponding contour lines for selected thresholds (figures 12*a* and 12*b*), and stretching them horizontally in the x direction, cylindrical domains aligned with the x axis are obtained, which are visualised in figure 11. In combination with the two limiting planes, they describe the gradual shrinking of the predictable zones with reasonable accuracy.

Like in the line observation configuration, the PPZ geometry can thus be interpreted and approximated through relatively simple physical considerations. Lines and squares are obviously not the only possible types of measurement configurations, but they are archetypal examples, which can certainly provide guidelines to determine the necessary measurement ranges in time and in the two horizontal dimensions, depending on the prediction horizon and accuracy requirements.

6. Experimental characterisation of PPZs

In this section, the present probabilistic theory of wave predictability is tested against real wave data. The general idea relies on comparing the prediction performance, obtained in a real wave field, with the prediction performance expected from the probabilistic predictability theory. More specifically, given a chosen measurement configuration, the optimal prediction matrix $\mathbf{P}$ (3.5) and the theoretical achievable NMSE are calculated using the procedure in § 3, where the covariance structure is obtained from the wave spectrum using (3.6). Then, the prediction matrix is used to actually calculate the predictions throughout all the record following (3.4), and the resulting empirical NMSE can be compared with the theoretical achievable NMSE.

6.1. *Dataset presentation*

The dataset employed in this study consists of a three-dimensional movie of a wave field, reconstructed by means of a system of two stereo cameras, from an ocean observation platform in the Yellow Sea in Korea, and thoroughly described and analysed by Benetazzo *et al.* (2018). It is part of a larger ensemble of datasets, introduced by Veras Guimãraes *et al.* (2020), recorded at various locations worldwide and made freely available online. Unlike other datasets in (Veras Guimãraes *et al.* 2020), the Yellow Sea record has a spatial coverage, in the order of two peak wavelengths, large enough for relatively accurate wave spectrum estimation through direct Fourier analysis. Furthermore, both the spatial coverage and time duration allow the empirical assessment of predictable zones. The statistical analysis reported by Benetazzo *et al.* (2018) shows skewness and kurtosis values of $m_3 = 0.14$ and $m_4 = 3.10$, respectively, indicative of detectable deviation from Gaussianity due to bound harmonics. However, as shall be observed in the rest of this section, such a deviation does not appear to have an appreciable negative impact on the applicability of the – albeit linear – probabilistic approach of § 3. Overall, this dataset is an almost-ideal playing field to assess wave predictability in a real sea scenario.

The sea state is characterised by a significant wave height H_{m_0} of approximately 2 m, and a peak wave period T_p below 6 s. In addition, as documented by Benetazzo *et al.* (2018), a strong current is present, with speed larger than 1 m s^{-1}. The water depth is approximately 50 m, which is close to deep-water conditions for the wavelengths present in the wave spectrum. Hence, bathymetry variations should have little or no impact on experimental results.

The original record is 440 s long, with a time step of 0.1 s. For the remainder of this study, a 120 m × 120 m square area is extracted, and resampled regularly in space with a 2 m resolution in both x and y directions. The wave elevation is also resampled with an

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

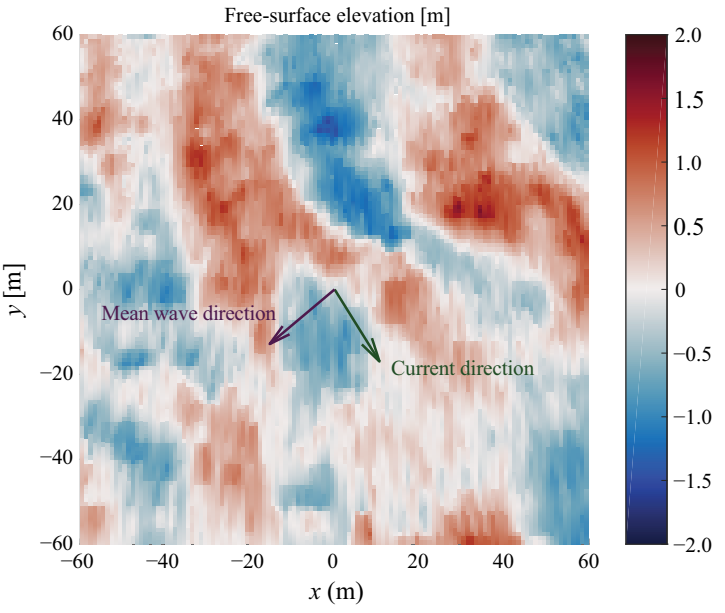


Figure 14. A 120m × 120m snapshot of the FSE record extracted from the dataset of Veras Guimãraes *et al.* (2020). The intrinsic peak wave direction and current direction are also represented.

0.25 s time step. figure 14 shows a sample snapshot of the record, in the square subdomain extracted for the present study.

The record is first used to estimate the underlying wave spectrum, which is accomplished using Fourier transform in the three dimensions $(x,\ y,\ t)$. One challenge is that, for the calculations of (3.6), the intrinsic spectrum must be used, while the direct, three-dimensional Fourier transform of the dataset provides the apparent spectrum (i.e. modified by surface current). Therefore, the current must be estimated, which allows the correction of the apparent spectrum, as detailed in appendix C. The resulting intrinsic spectrum is shown in figure 15.

The peak wave period is found to be $T_p = 5.75$ s and the peak wave direction is 220° (in the reference frame of figure 14). With a mean water depth of 50 m, the peak wavelength is 51.6 m, so that the square area covers a little more than two peak wavelengths. Finally, wave energy is widely spread around the peak wave direction. The current magnitude is $1.09\ \mathrm{m\,s^{-1}}$, with a direction of 302°.

6.2. *Experimental comparison procedure*

Using the notations of § 2, consider the observation locations $\mathbf{x}_{i=1\ldots M_o}$ and time interval $[\underline{\tau}_o, \overline{\tau}_o]$. Let the waves be predicted at a single location $\mathbf{x}' = (x', y')$, over the time interval $[\underline{\tau}_p, \overline{\tau}_p]$. Based on the intrinsic spectrum and current, and using the Wiener–Khinchine relation (3.6), the procedure of § 3 is used to obtain the theoretical NMSE from (3.5), exactly like in §§ 4 and 5, but also the prediction matrix $\mathbf{P}$ (3.4).

The time series $\mathbf{z}_o(t) = (\eta(x_1, y_1, t), \ldots, \eta(x_{M_o}, x_{M_o}, t))^T$ and $\mathbf{z}_p(t) = \eta(x', y', t)$ are extracted from the empirical record at the observation and prediction locations $\mathbf{x}_{i=1\ldots M_o}$ and $\mathbf{x}'$, and the prediction is carried out in the sliding-window fashion sketched in figure 2, using (3.4) with the prediction matrix $\mathbf{P}$. The predicted value of η at time horizon τ, calculated from the observations available at time t, is denoted as $\eta_{p|o(t)}(x', y', t+\tau)$. The empirical NMSE $\tilde{\epsilon}_{\mathrm{emp}}$ at the prediction location $\mathbf{x}' = (x', y')$ is calculated, for each

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

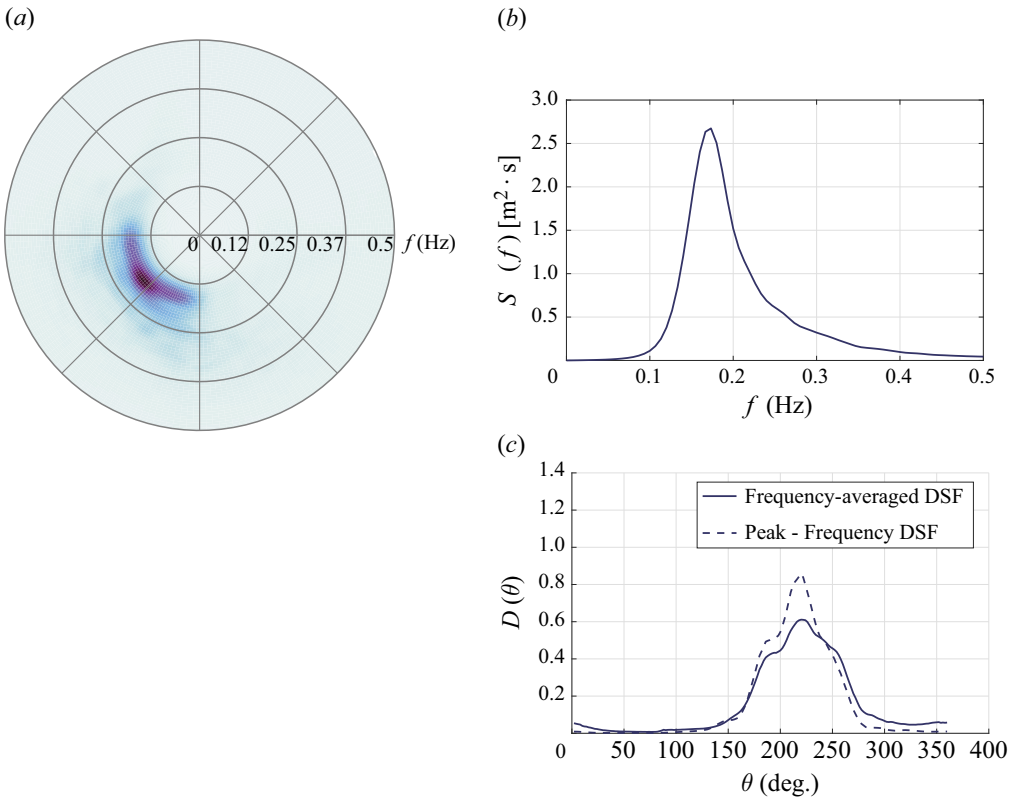


Figure 15. Intrinsic directional spectrum estimated from the Yellow Sea dataset. (*a*) Directional spectrum (*b*) Frequency spectrum (*c*) Directional spreading function (DSF), frequency-averaged and at peak frequency.

prediction horizon τ in $[\underline{\tau}_p, \overline{\tau}_p]$, as follows:

$$\tilde{\epsilon}_{emp}(x', y', \tau) = \frac{1}{m_0[i_{max} - i_{min} + 1]} \sum_{i=i_{min}}^{i_{max}} [\eta_{p|o(t_i)}(x', y', t_i + \tau) - \eta(x', y', t_i + \tau)]^2, \tag{6.1}$$

where

$$m_0 = \frac{1}{N_t N_x N_y} \sum_{i=1}^{N_t} \sum_{j=1}^{N_x} \sum_{k=1}^{N_y} \eta(x_j, y_k, t_i)^2 \tag{6.2}$$

is the variance of η, and i_{min} and i_{max} are the first and last time indices where the prediction can be carried out, depending on the values of $\underline{\tau}_o$, $\overline{\tau}_o$, $\underline{\tau}_p$ and $\overline{\tau}_p$. Allowing the prediction location (x', y') to vary across the whole experimental domain, empirical NMSE maps are constructed, which can be directly compared with the theoretical NMSE.

In the procedure outlined above, a significant implementation detail must be considered. Indeed, the actual wave time series may be subject to measurement noise, or unmodelled nonlinear dynamics. Furthermore, the intrinsic spectrum and the current may not have been estimated perfectly. Those effects all imply that actual wave statistics can depart slightly from those predicted in the Gaussian approximation. As a consequence, some robustness must be introduced in the prediction matrix calculation, to ensure that empirical results are not too sensitive to even minor deviations from the expected statistics. To

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

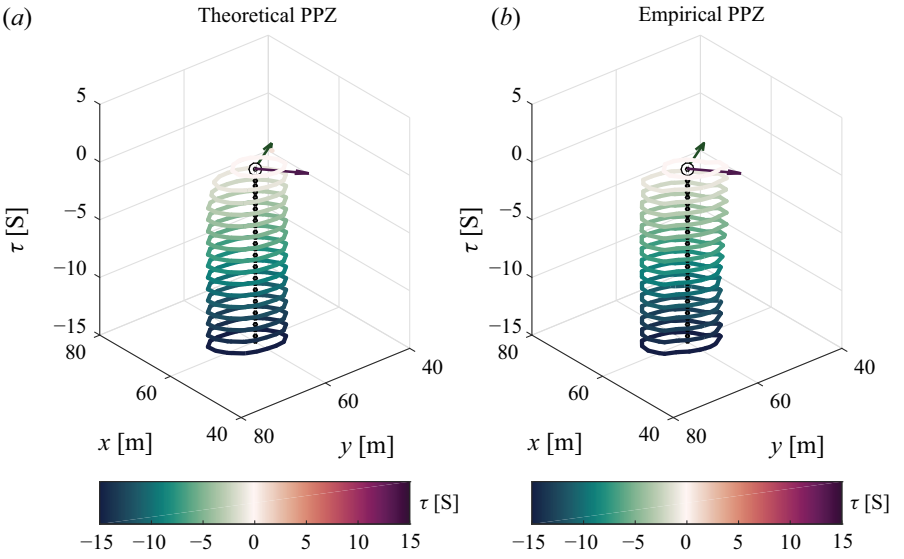


Figure 16. Theoretical and empirical PPZs for a 20 % NMSE threshold. Waves are observed at a single location, looking for 30 s into the past. The peak wave direction and current direction are represented by purple and green arrows, respectively. (*a*) Theoretical PPZ computed from the wave spectrum. (*b*) Empirical PPZ obtained by applying the forecasting model to the experimental data. Vertical spacing between (x, y) contour lines: 1 s.

that aim, it is assumed that some level of uncorrelated, Gaussian noise is also present in the observations, which amounts to adding a positive, constant term σ_{n}^2 to the diagonal elements of Σ_{oo}, where σ_{n}^2 represents the white noise variance. Without noise modelling ($\sigma_n^2 = 0$), the empirical NMSE is far below the theoretical one. By gradually increasing σ_n^2, the empirical NMSE increases significantly, while the theoretical NMSE remains almost unchanged. Here σ_n^2 can thus be tuned so that theoretical and empirical results closely match, as done by Mérigaud *et al.* (2018). A value of $\sigma_{\mathrm{n}}^2 = 0.025 m_0$ (in other words, 2.5 % of the wave signal variance) is adopted throughout this experimental study, and found sufficient to ensure robustness of results.

6.3. *Single observation point*

Starting with a single measurement point (chosen in the centre of the observation frame), and setting $\underline{\tau}_o = -30$ s, $\overline{\tau}_o = 0$ s, with a time step of 0.5 s, theoretical and empirical NMSE maps in the (x, y) plane are constructed. The resulting theoretical and empirical PPZs are compared in figure 16, for a 20 % threshold value, showing a very good agreement between the theoretically predicted PPZ and the empirical one. Like in § 5.1, predictable zones at specific time horizons resemble ellipses, with their principal axis almost aligned with the peak wave direction. However, because of surface current which is almost perpendicular to the peak wave direction, the PPZ orientation is slightly deviated. Regardless of the precise shape of PPZs, empirical results tend to confirm that, using a single observation location in a directional wave field, no forecasts of practical interest can be obtained, unless the prediction requirements are extremely low, in terms of time horizon and prediction quality. Most importantly, neither the theoretical, nor the empirical results, support the existence of fan-shaped predictable zones, as constructed in the deterministic theory.

figure 16 is only one example, since only one NMSE threshold is considered, but a more complete set of results, including a comparison of theoretical and empirical NMSE

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

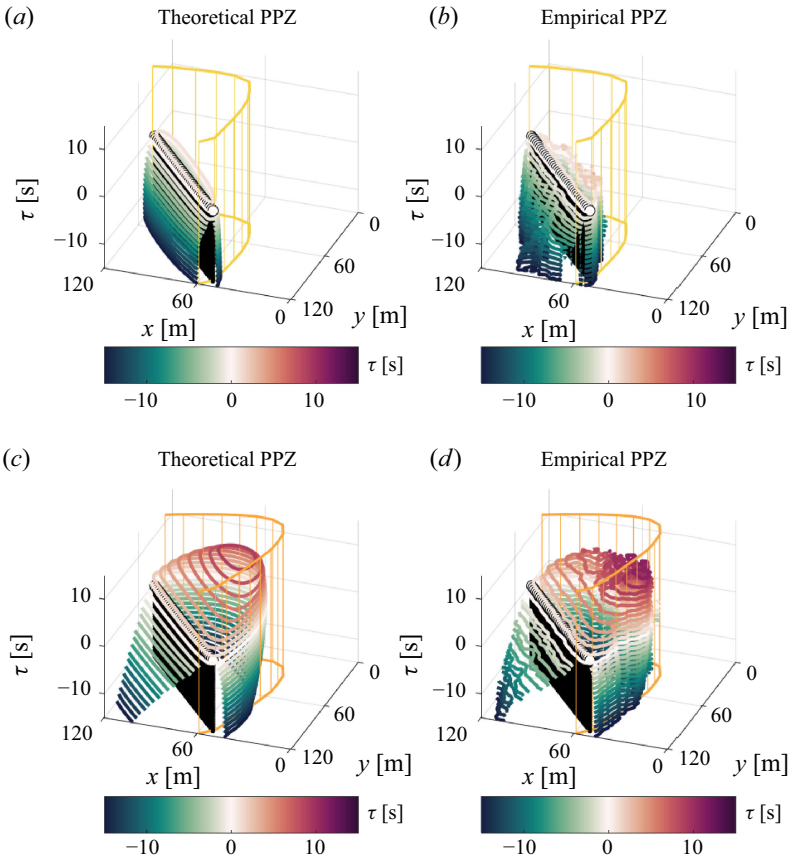


Figure 17. Theoretical and empirical PPZs for 10 % and 20 % NMSE thresholds. Waves are observed along a line perpendicular to the main intrinsic wave direction, looking for 30 s into the past. (*a*,*c*) Theoretical PPZ computed from the wave spectrum. (*b*,*d*) Empirical PPZ obtained by applying the forecasting model to the experimental data. Vertical spacing between (x, y) contour lines: 1 s. The effect of directional spreading is approximated by cylindrical envelopes (vertically stretched 10 % and 20 % GOP contour lines). (*a*,*b*) 10% PPZ. (*c*,*d*) 20% PPZ.

maps at selected prediction horizons, is available as Supplementary materials. Those additional results all confirm the agreement between theoretical and empirical prediction performance metrics.

6.4. *A line of observation points*

A line of measurement points is also investigated, oriented perpendicular to the peak intrinsic wave direction, similarly to § 5.3. The line length is two peak wavelengths like in § 5.3, with a point density of 20 points per wavelength. Like in the single-point case, observations are made between $\underline{\tau}_o = -30$ s and $\overline{\tau}_o = 0$ s, with a time step of 0.5 s. Theoretical and empirical results are shown in figure 17, in the form of the 10 % and 20 % PPZs. Considering the figure, the line of observations does not allow significant prediction horizons if a 10 % NMSE criterion is retained, which can be attributed to the large directional spreading of the wave energy content. Loosening the criterion to 20 % (figure 17*b*), the prediction horizon is extended to almost 10 s. Like in § 5.3, cylindrical domains, obtained from GOP contour lines, are also represented, as an approximation to the effect of directional spreading. In this case, the effect of current is taken into account

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

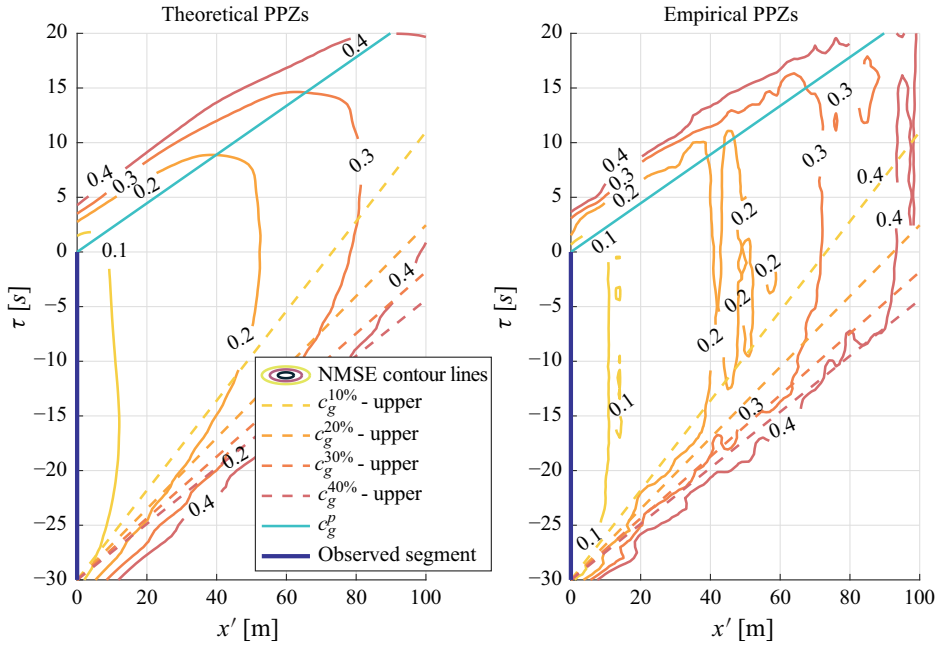


Figure 18. NMSE contour lines defining PPZs for selected threshold values, along with limiting slopes determined in § 4. Waves are observed like in figure 17. Results are taken along a segment aligned with the intrinsic peak wave direction, and crossing the middle observation point. x' denotes the longitudinal distance along the segment (i.e. the distance from the observation line).

to calculate the GOP, as documented in appendix B. As can be appreciated in figure 17(a), the GOP 10 % contour line is not a satisfactory approximation to the spatial extent of the 10 % PPZs, both theoretical and empirical. The agreement is significantly more convincing for the 20 % PPZs in figure 17(b). Such discrepancies may be due to the fact that GOP is calculated from the directional energy content at the peak frequency, while the actual directional energy content varies significantly across frequencies, as can be appreciated in figure 15.

In addition to directional spreading, it is interesting to verify the two other ingredients governing wave predictability, highlighted in § 5, namely the two limiting group velocities. To that end, the NMSE contour lines are visualised along a segment, aligned with the peak wave direction and crossing the middle observation point. This results in NMSE contour lines in the $x' - \tau$ plane (because the current is oriented almost perpendicularly to the waves, it is not expected to have a significant effect in this cross-section of the results), where x' is the distance from the middle observation point. Those contour lines are represented in figure 18 for specific NMSE threshold values. In the same figure, straight lines are also represented, with slopes corresponding to the limiting group velocities identified in § 4. At small distances, each z % PPZ is bounded from above by the group velocity at peak frequency, and from below by $c_g^{z\,\%,\mathrm{upper}}$, determined from the energy criterion of § 4. However, instead of remaining guided between the two limiting slopes, the PPZs are spatially limited under the effect of directional spreading, long before the two slopes intersect. Compared with the two-dimensional results of § 4, it is more difficult, from figure 18 alone, to assess exactly the accuracy of the proposed limiting slopes $c_g^{z\,\%,\mathrm{upper}}$ and c_g^p, due to the fact that the PPZs collapse rapidly under the directional

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

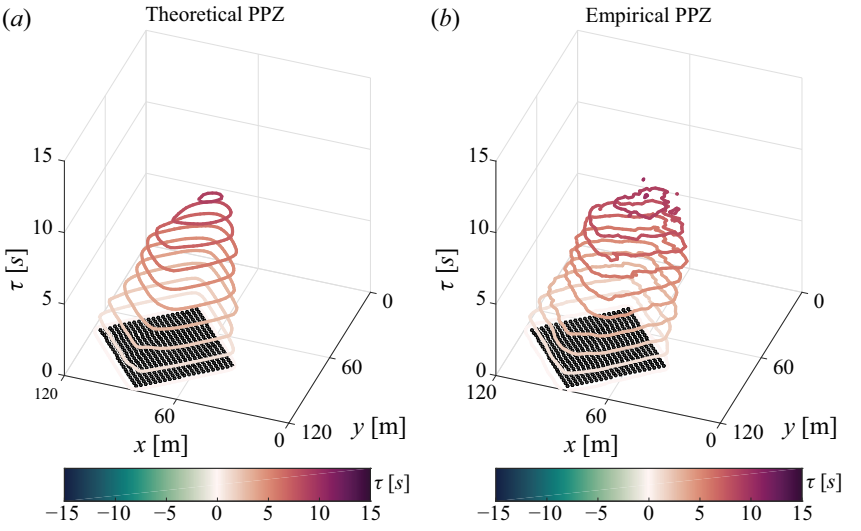


Figure 19. Theoretical and empirical PPZs for a 20 % NMSE threshold. Waves are observed in $\tau = 0$ in a 21 × 21-point square, aligned with the main intrinsic wave direction, with the sides equal to one peak wavelength. (*a*) Theoretical PPZ computed from the wave spectrum. (b) Empirical PPZ obtained by applying the forecasting model to the experimental data. Vertical spacing between (x, y) contour lines: 1 s.

spreading effect. Nevertheless, those two group velocities remain, from this set of results, very plausible approximations.

Overall, there is a good agreement between theoretical and empirical results, which is further supported by additional figures, available as Supplementary materials, including theoretical and empirical NMSE maps in the (x, y) plane for selected prediction horizons.

6.5. *A square snapshot of observation points*

The square snapshot configuration of § 5.4 is now considered, where the square is aligned with the peak intrinsic wave direction, the square side length is one peak wavelength, and the point density in both longitudinal and transverse directions is 20 points per peak wavelength. The NMSE results are presented as a PPZ for the 20 % NMSE criterion, in figure 19. Overall, empirical results (figure 19*b*) follow those, calculated from the wave spectrum (figure 19*a*). The empirical prediction performance is, however, slightly better than expected, without calling into question the overall shape and behaviour of the PPZ. Like in the line configuration, the area of better prediction accuracy is transported along the main wave direction at the peak group velocity, and in the current direction at the current speed (see appendix B for more detail). Overall, retaining a 20 % NMSE criterion, this square configuration allows forecasts up to approximately 10 s ($\approx 1.5T_p$) into the future.

Finally, figure 20 shows an example of the time evolution of the wave field, predicted in the square snapshot configuration. For selected prediction horizons, figure20(*a,c*) represent the actual FSE, in and around the initial observation domain, while figure 20(*b,d*) shows the FSE, predicted in the same area, using the square snapshot (bounded by dotted lines) taken at $\tau = 0$. The theoretical 20 % PPZ (i.e. figure 19*a*) is also materialised by means of dashed, green lines, for each sample prediction horizon. Note that, at time $\tau = 0$, the 20 % PPZ includes both the square observed snapshot, as well as the surrounding area where the FSE can be inferred from the snapshot with NMSE smaller than 20 %.

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

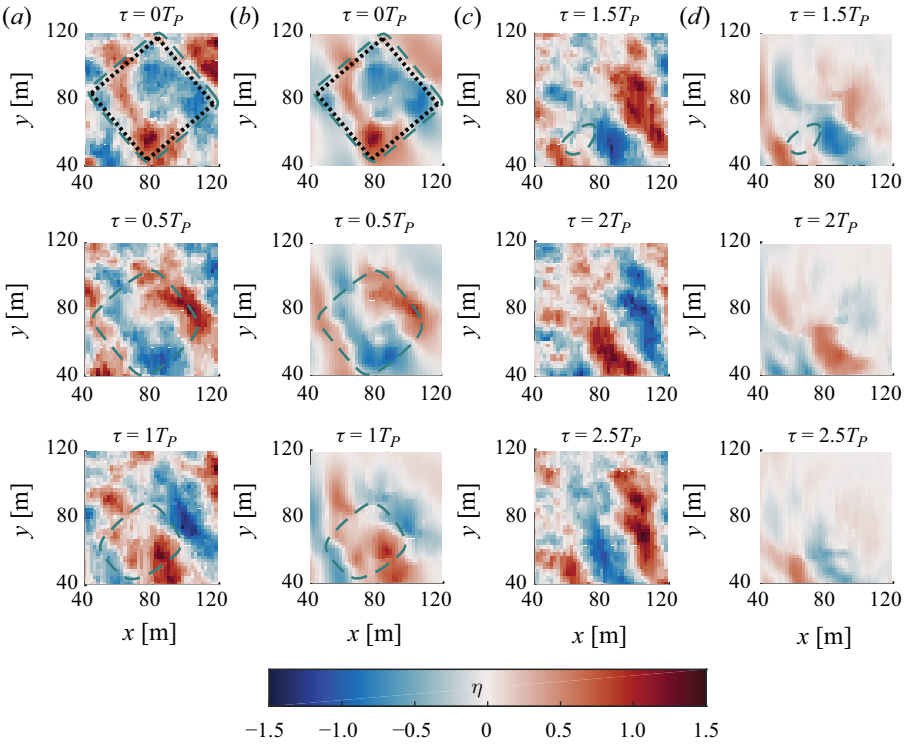


Figure 20. An example of wave field prediction from a square snapshot of observations, at different prediction horizons τ. The square snapshot is represented by dotted lines at $\tau = 0$. (*a,c*) True wave field, (*b,d*) predicted wave field. Green, dashed lines: theoretical 20 % PPZ.

figure 20 illustrates how the achievable prediction quality degrades with the time horizon, increasingly resembling a low pass-filtered version of the true wave field, whereby only larger wavelengths are preserved – an observation already suggested by Mérigaud & Tona (2022) by analysing the frequency content of the prediction error signal. In addition, incoming waves mostly propagate from the top-right corner of the domain. As a result, for larger time horizons (beyond T_p), the top-right part of the wave field essentially consists of newly entered, independent wave groups, which could by no means be inferred from the observation square at initial time $\tau = 0$. Since there is virtually no information about the FSE in that area of the domain for larger time horizons, the optimal forecast approaches zero, i.e. the unconditional mean of the FSE field.

7. Conclusion

This work studies the real-time predictability of ocean waves, using the framework of Gaussian stochastic fields, which allows for a quantitative definition of predictable zones, with respect to a chosen NMSE threshold. In both two-dimensional (unidirectional) and three-dimensional (omnidirectional) cases, relatively straightforward physical interpretations can be given for the geometry of PPZs.

In two-dimensional wave fields, the triangular predictable zones, theorised in the deterministic theory, are recovered, although with two novel specifications. First, the lower limiting group velocity is determined quantitatively from the wave spectrum, using an energy integral criterion, depending on the prescribed NMSE threshold. Second, the upper

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

limiting group velocity is that of the peak frequency, regardless of the NMSE threshold (at least thresholds of practical interest), due to the sharp decrease of the spectrum energy below the peak frequency. Hence, the spectrum peak could be interpreted as the centre frequency of the lowest-frequency wave packets present in the wave field. The latter finding corroborates simulation and laboratory results of Law *et al.* (2020); Hlophe *et al.* (2021); Huchet *et al.* (2023), as well as the real-sea experimental results of the present study.

In three-dimensional wave fields, the group velocities and wave dispersion do not suffice to account for wave predictability. When the waves are observed in only one location, the different – in fact, infinitely many – directions which constitute the wave field cannot be distinguished within the observed wave signal, which results in the prediction quality deteriorating very rapidly away from the observation location, regardless of the observation duration. The latter result is in contradiction to the fan-shaped predictable zones constructed from the deterministic theory, which would allow arbitrarily long prediction horizons by expanding the wave observation duration. According to the present probabilistic model, the time horizons, achievable in realistic wave fields, become drastically more pessimistic, including when the wave spectrum is directionally truncated.

The wave field directional components can be resolved, however, by observing arrays of field points (§§ 5.3 and 5.4). In such a case, accurate predictions can be achieved for larger horizons and distances – a fact already well known. The lower and upper group velocities, highlighted in the two-dimensional case, continue to play a role in bounding the predictable zone geometry in the main wave direction. However, the predictable zones also shrink in the direction transverse to the main wave direction, which is the effect of new, independent wave groups, incoming from the sides, without travelling through the observation array. By considering simple, densely populated array layouts, with dimensions larger than a peak wavelength, the directional spreading effect is reasonably well approximated by the GOP map of the observation array, i.e. the fraction of the wave energy, which travels through the array prior to reaching any given point in space and time. In the simple layouts considered here, the GOP is straightforward to calculate, from a representative directional spreading function and the array geometrical parameters. This model, however, should be improved, for it to be of any practical use to accommodate more complex wave spectra and measurement layouts.

The results of § 6 provide empirical confirmation of the key parameters which govern the predictability of random ocean waves. They also support the fact that the stochastic approach, presented in § 3, is suitable both to comprehend wave predictability, and as a candidate practical tool for real-time wave prediction, since empirical results are close to the theoretical maximum achievable prediction performance – setting aside computational aspects which are beyond the scope of this work. More generally, the present empirical results suggest that, in this dataset, actual ocean waves behave as expected from the Gaussian stochastic wave model – at least to a good approximation. Indeed, empirical results are obtained by applying a linear predictor, which is itself calculated from the wave spectrum based on linearity and Gaussianity assumptions. If strong nonlinear effects were present in this dataset, empirical results would be affected by additional prediction error due to those unmodelled nonlinear effects, hence yielding prediction performance significantly worse than the Gaussian theory predicts – a deviation not observed in this work.

Finally, the generality of the results presented here deserves to be discussed. First, it must be clear that the present stochastic approach only applies to ordinary, wind-generated ocean wave fields. Those wave fields are, in nature, made of so many, infinitesimal contributions, that the Gaussian random field model is an accurate description, with

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

the spatial correlation structure given by the dispersion relation and surface current. In contrast, with punctual phenomena such as tsunamis, or with artificial waves, such as those in a ship's wake, where incoming wave groups are not independent, the stochastic approach would not work, while a deterministic model, based on physical wave propagation equations, might.

Second, the physical interpretations of probabilistic prediction zones, provided throughout this study in §§ 4, 5 and 6, are made possible because relatively simple wave spectra and observation configurations are considered. Sea states composed of two or more wave systems, or observation configurations with more complex geometries, would probably not lend themselves to such simplified physical approximations for the shape of predictable zones. Nevertheless, it is believed by the authors that adopting the present stochastic approach, for the study of any sea state and observation configuration, remains valuable.

Third and last, the empirical forecasting results in § 6 are obtained using a linear prediction function, of which the parameters are calculated from the wave spectrum, based on the theory of § 3. It may thus be argued that, by using a different forecasting method (for example, a deterministic wave propagation model), better empirical results may be obtained in this dataset. In such a case, the empirical results provided in the present work would not constitute evidence of wave predictability in general, but merely of how well the specific prediction approach presented here performs. Yet, under the Gaussian wave field assumptions of § 3, the calculated linear prediction function is optimal, and consequently outperforms any other prediction model employing the same wave observations. Additionally, as discussed above, the agreement between empirical and theoretical prediction performance suggests that the Gaussian wave field assumption is approximately valid in this dataset. As a consequence – and this is a claim crucial to this work – no wave prediction model, including those based on deterministic wave propagation equations, can provide any substantial improvement to the experimental results presented in § 6, except maybe for modest nonlinear corrections. In that sense, the results of this study are indeed a true assessment of the predictability of random ocean waves.

Supplementary material. A supplementary material is available at https://doi.org/10.1017/jfm.2025.84.

Declaration of interests. The authors report no conflict of interest.

Appendix A. Effect of the low-frequency spectrum decrease rate onto two-dimensional predictable zones

This brief section examines why the largest group velocity, governing the upper limit of two-dimensional predictable zones, should be taken as the group velocity at peak frequency. As hypothesised in § 4.1, the reason may be in the fast decrease rate of the spectrum energy below the peak frequency, which would make the spectrum peak behave like one wave packet. In order to verify that possibility, the fictitious spectrum of figure 21(*b*) is built, with a linear decrease rate below the peak frequency, and the resulting PPZs are represented in figure 21(*a*). It is found that, in contrast to realistic wave spectra, both the lower and upper limits of PPZs can be retrieved from an energy integral criterion, in the upper and lower frequency ranges. As a consequence, it is indeed the specific, sharply decreasing energy content of real ocean wave spectra, which makes the peak group velocity the adequate largest velocity shaping predictable zones.

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

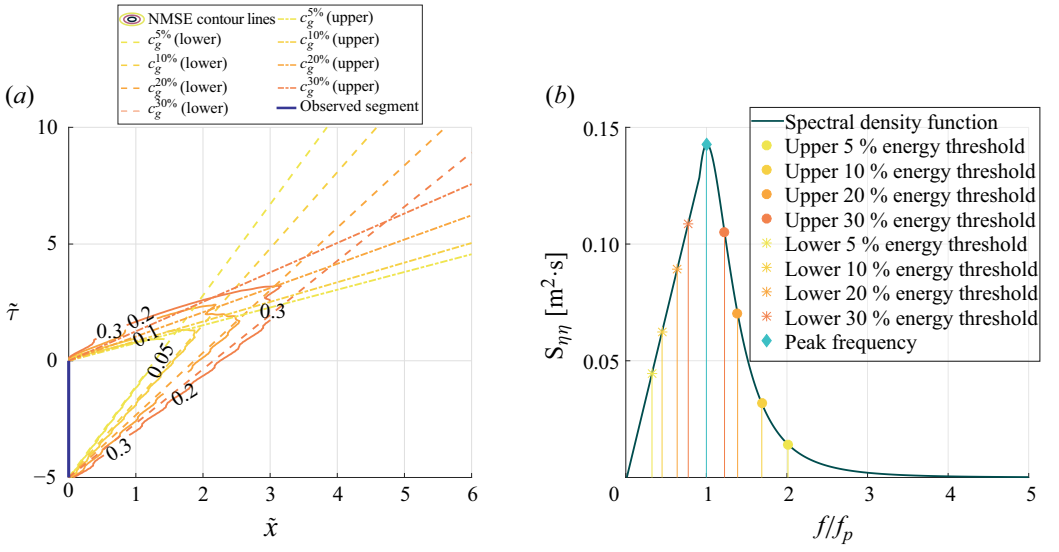


Figure 21. Two-dimensional PPZs (*a*) for a fictitious spectrum with a linear decrease rate below the spectrum peak (*b*). Both lower and upper PPZ limiting slopes can be retrieved accurately from an energy criterion: let $f_{z\,\%}^{\mathrm{upper}}$ (respectively, $f_{z\,\%}^{\mathrm{lower}}$) be the frequency above (respectively, below) which $z\,\%$ of the spectrum energy are confined, and $c_g^{z\,\%,\mathrm{upper}}$ (respectively, $c_g^{z\,\%,\mathrm{lower}}$) the corresponding group velocity. Then the $z\,\%$ PPZ is bounded from below by a slope $\approx 1/c_g^{z\,\%,\mathrm{upper}}$ and from above by a slope $\approx 1/c_g^{z\,\%,\mathrm{lower}}$.

Appendix B. Surface current effect on wave predictability

Surface current is readily taken into account in the present probabilistic framework, through the modified Wiener–Khinchine relation (3.6). This section gives a brief account of how the geometry of PPZs is modified by current, represented by means of a current velocity vector $\mathbf{U} = (U_x, U_y)$.

Two-dimensional wave predictability in the presence of current

In a two-dimensional scenario, only the x-wise current vector component matters. figure 22 shows the PPZs for selected NMSE threshold values, with a surface current going against the waves (figure 22*a*), without surface current (figure 22*b*) and with a surface current going with the waves (figure 22*c*). The physical interpretation in terms of limiting group velocities, highlighted in § 4, can be retained if every group velocity is replaced with its effective value, that is $c_g + U_x$.

With current going against the waves, the effective peak group velocity is slowed down, which should be favourable to longer-horizon forecasts. However, the lowest group velocities are slowed down even more, relatively to their intrinsic value. Therefore, current reduces the PPZ extent (figure 22*a*). In contrast, when current goes with the waves, lowest group velocities are, relatively, more accelerated than the largest (peak) group velocity, which overall expands PPZs.

Three-dimensional wave predictability in the presence of current

The three-dimensional case is now considered, beginning with a line of observation points, perpendicular to the peak wave direction. In the example of figure 23, the 20 % PPZ without current (figure 23*a*) is compared with that obtained with a current of 2 m s^{-1},

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

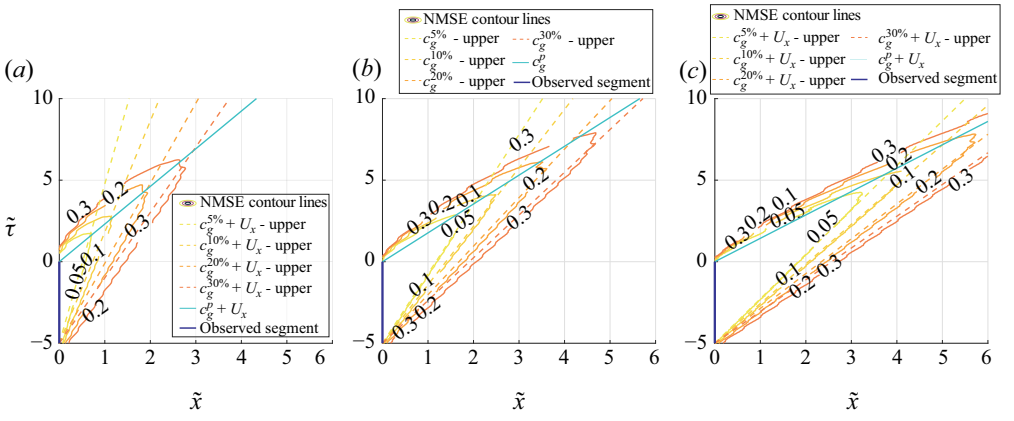


Figure 22. Effect of surface current on two-dimensional PPZs. (*a*) $U_x = -2$ m/s (*b*) $U_x = 0$ m/s (*c*) $U_x = 2$ m/s.

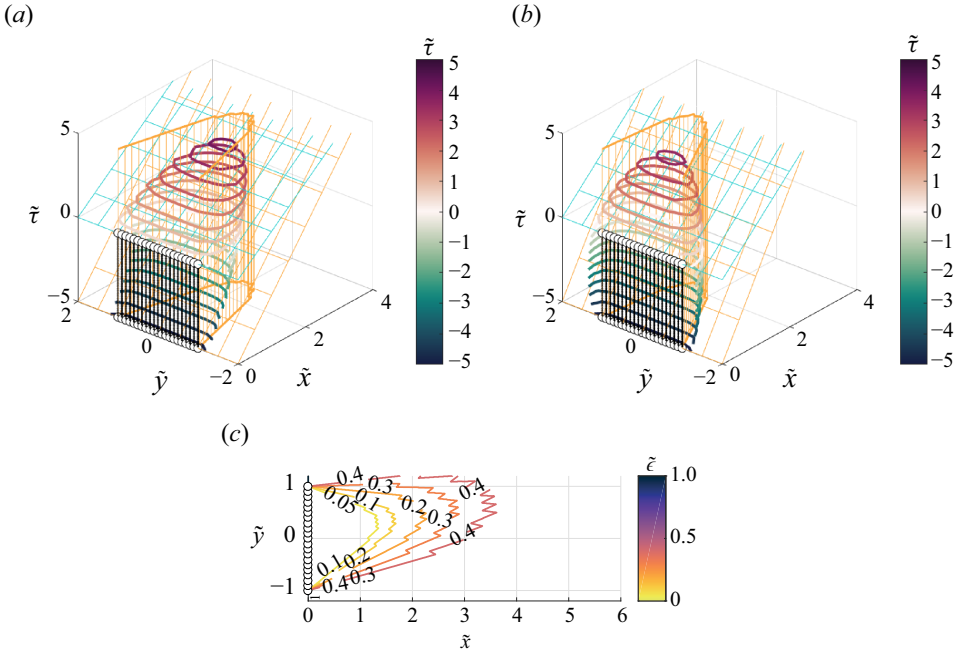


Figure 23. Three-dimensional PPZs, without current (*a*) and in the presence of surface current (*b*), with a line of observation points. figure 23(*c*) shows contour lines of the GOP of § 5.3, modified by current as in (B1). (*a*) $U = 0$ m/s (*b*) $U_y = -U_x = \sqrt{2}$ m/s (*c*) GOP contour lines, $U_y = -U_x = \sqrt{2}$ m/s.

forming a 135° angle with the peak wave direction (aligned with the x axis), that is $U_x = -\sqrt{2}\,\mathrm{m\,s^{-1}}$ and $U_y = \sqrt{2}\,\mathrm{m\,s^{-1}}$ (figure 23*b*).

Accordingly to the two-dimensional results of § B.1, the two planes which bound the PPZ from below and above are modified, by adding the x-wise velocity component U_x to the relevant group velocities. In addition to the latter deformation, which takes place in the peak wave direction, the PPZ is also deformed laterally with respect to the incoming wave direction, as can be appreciated in figure 23(*b*). The PPZ envelope can be approximated in a way similar to that of § 5.3, by modifying slightly the GOP calculation to take current

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

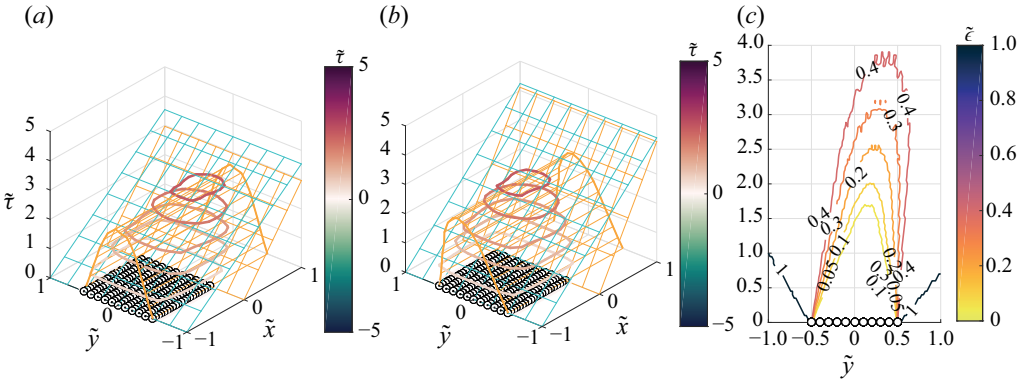


Figure 24. Three-dimensional PPZs, without current (*a*) and in the presence of surface current (*b*), with a square snapshot of observation points. Panel (*c*) shows contour lines of the GOP of § 5.4, modified by current as in (B3). (*a*) $U = 0$ m/s (*b*) $U_y = -U_x = \sqrt{2}$ m/s (*c*) GOP contour lines, $U_y = -U_x = \sqrt{2}$ m/s.

into account. Assuming, again, that the wave field is made of wave groups centred around the peak frequency, it is possible to define an effective propagation angle: consider a wave group propagating at speed c_g^p with intrinsic propagation angle θ. During a time duration T, the wave group travels along its intrinsic direction $\mathbf{u}(\theta) = (\cos(\theta), \sin(\theta))$, while also being transported by current. Therefore, the total wave group displacement is $\mathbf{X} = T[c_g^p \mathbf{u}(\theta) + \mathbf{U}]$, and the effective travelling angle is $\theta_{\text{eff}}(\theta) = \angle\{c_g^p \mathbf{u}(\theta) + \mathbf{U}\}$.

The integration domain, in the GOP calculation of (5.3), is thus modified as follows:

$$\text{GOP}'(x, y) = \int_{\mathcal{S}'(x,y)} D_p(\theta) \mathrm{d}\theta, \tag{B1}$$

where

$$\mathcal{S}'(x, y) = \{\theta : \theta_{\text{eff}}(\theta) \in \mathcal{S}(x, y)\} \tag{B2}$$

and $\mathcal{S}(x, y)$ is the angular sector geometrically defined as in figure 10. The resulting modified GOP contour lines are drawn in figure 23(*c*). The 20 % contour line is extended vertically to form a cylindrical domain, also shown in figure 23(*b*), thus providing a satisfactory approximation to the deformed PPZ envelope, along with the two limiting planes.

Current-induced modifications of the PPZ are illustrated in figure 24, in the square-snapshot measurement configuration, with the same current $U_x = -\sqrt{2}$ m s^{-1} and $U_y = \sqrt{2}$ m s^{-1}. Here as well, the two limiting planes are modified through the addition of U_x to the limiting group velocities. Furthermore, the GOP calculation of § 5.4 is straightforwardly modified to take into account the fact that, during time lag τ, the wave group is transported laterally by a distance τU_y. Thus, the modified GOP is

$$\text{GOP}'(y, \tau) = \text{GOP}(y - \tau U_y, \tau) \tag{B3}$$

where $\text{GOP}(y, \tau)$ is given by (5.4) and (5.5). The resulting modified GOP contour lines are drawn in figure 24(*c*). The 20 % contour line is extended horizontally to form a cylindrical domain, also shown in figure 24(*b*), thus providing a satisfactory approximation to the deformed PPZ envelope, along with the two limiting planes.

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

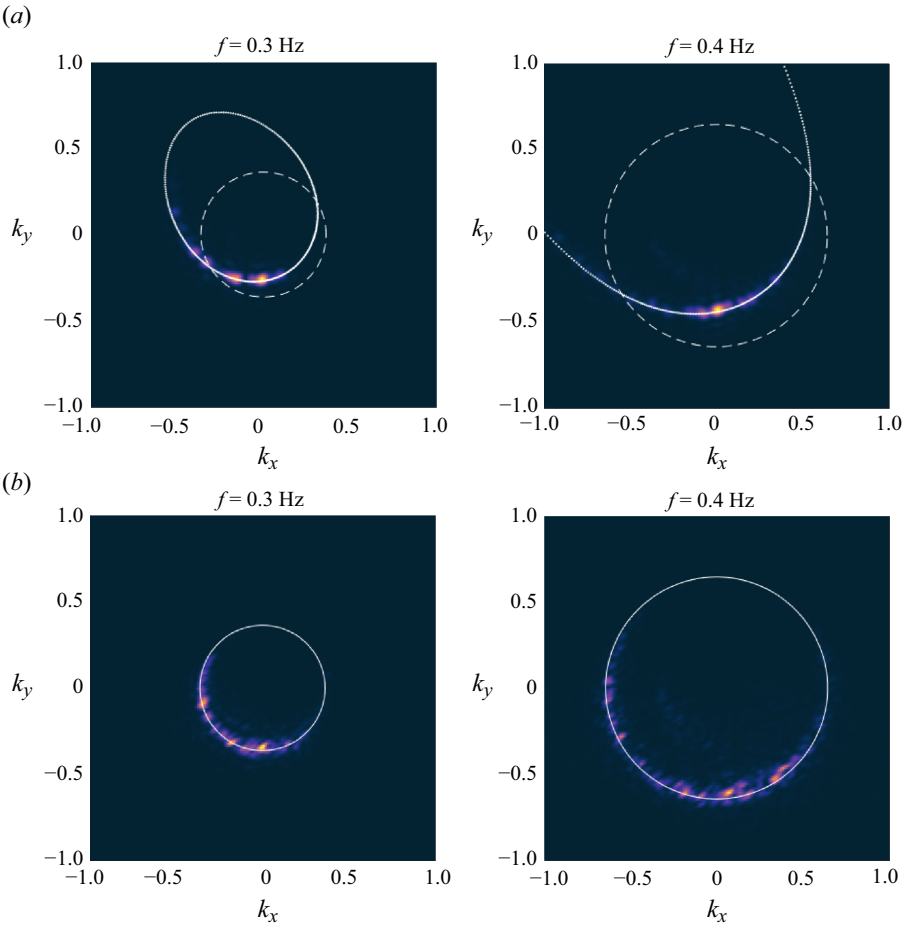


Figure 25. Apparent spectrum and intrinsic spectrum in the wavevector domain at selected frequencies. Linear dispersion shells are shown, without (dashed lines) and with (solid lines) the fitted current. (*a*) Apparent spectrum (*b*) Intrinsic spectrum.

Appendix C. Experimental wave spectral analysis

This section documents how the intrinsic wave spectrum, necessary for the calculation of the Wiener–Khinchine relation (3.6), is estimated from the stereo dataset.

In a first step, the three-dimensional spectrum (wavevector and frequency) of the FSE field is estimated. To that end, a two-dimensional fast Fourier transform is carried out in the spatial domain, while a windowed spectrum estimation (periodogram) is calculated in the time dimension. Because of the relatively small size of the observed domain (approximately two peak wavelengths), spatial fast Fourier transform is calculated after zero-padding, three times the size of the initial domain, is implemented. This provides a smoother spectrum in the wavenumber domain, and more accurate peak localisation. Those steps yield the three-dimensional apparent spectrum, that is, for each frequency, a two-dimensional spectral density function in the wavenumber domain. The apparent spectrum is illustrated at selected frequencies, in figure 25(*a*).

In theory, the spectrum energy, at a given frequency ω, should be concentrated along a circular shell, with radius $k(\omega)$, corresponding to the dispersion relation. However, as can be appreciated in figure 25(*a*), the dispersion shell is significantly distorted, which is

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

accounted for by the presence of current. Such a distortion is more clearly visible at larger frequencies.

For a representation of the intrinsic spectrum in both frequency and direction, the apparent spectrum must be corrected for the influence of current. To do so, the current magnitude and direction are fitted so that the corresponding deformed dispersion shell (solid lines in figure 25*a*) matches the observations reasonably well. Once **U** is estimated, it is possible to reconstruct the intrinsic spectrum, as detailed by Benetazzo *et al.* (2018). The intrinsic spectrum obtained is represented, for each of the selected frequencies, in figure 25(*b*), where it can be seen that the spectrum energy is now distributed along the currentless linear dispersion relation.

The wavenumber spectrum $S_{\eta\eta}(\mathbf{k})$ is then obtained by integrating the three-dimensional intrinsic spectrum over the frequency range, and $S_{\eta\eta}(\mathbf{k})$ is converted to a frequency-direction representation $S_{\eta\eta}(\omega, \theta)$, for ease of computation in the Wiener–Khinchine relation (3.6). The result is represented in figure 15.

Finally, the authors of the present work have also assessed possible spatial variations in surface current within the observation domain, using three-dimensional Fourier analysis in smaller subdomains (not reported here), and determined that no appreciable variation in surface current can be observed.

References

Abusedra, L. & Belmont, M. 2011 Prediction diagrams for deterministic sea wave prediction and the introduction of the data extension prediction method. *Intl. Shipbuilding Prog.* **58** (1), 59–81.

Al-Ani, M., Belmont, M. & Christmas, J. 2020 Sea trial on deterministic sea waves prediction using wave-profiling radar. *Ocean Engng* **207**, 107297.

Al-Ani, M., Christmas, J., Belmont, M., Duncan, J.M., Duncan, J. & Ferrier, B. 2019 Deterministic sea waves prediction using mixed space–time wave radar data. *J. Atmos. Ocean. Technol.* **36** (5), 833–842.

Ardhuin, F., Gille, S.T., Menemenlis, D., Rocha, C.B., Rascle, N., Chapron, B., Gula, J. & Molemaker, J. 2017 Small-scale open ocean currents have large effects on wind wave heights. *J. Geophys. Res. Oceans* **122** (6), 4500–4517.

Belmont, M., Al-Ani, M., Challenor, P., Christmas, J. & Wilson, P. 2019 Obtaining the distribution of quiescent periods directly from the power spectral densities of sea waves. *Appl. Ocean Res.* **85**, 65–72.

Belmont, M., Christmas, J., Dannenberg, J., Hilmer, T., Duncan, J., Duncan, J.M. & Ferrier, B. 2014 An examination of the feasibility of linear deterministic sea wave prediction in multidirectional seas using wave profiling radar: theory, simulation, and sea trials. *J. Atmos. Ocean. Techol.* **31** (7), 1601–1614.

Benetazzo, A., Bergamasco, F., Yoo, J., Cavaleri, L., Kim, S.-S., Bertotti, L., Barbariol, F. & Shim, J.-S. 2018 Characterizing the signature of a spatio-temporal wind wave field. *Ocean Model.* **129**, 104–123.

Desmars, N., Hartmann, M., Behrendt, J., Hoffmann, N. & Klein, M. 2023 Nonlinear deterministic reconstruction and prediction of remotely measured ocean surface waves. *J. Fluid Mech.* **975**, A8.

From, P.J., Gravdahl, J.T., Lillehagen, T. & Abbeel, P. 2011 Motion planning and control of robotic manipulators on seaborne platforms. *Control Engng Practice* **19** (8), 809–819.

Fucile, F., Bulian, G. & Lugni, C. 2018 Prediction error statistics in deterministic linear ship motion forecasting. In *ASME 2018 37th International Conference on Ocean, Offshore and Arctic Engineering*, pp. V07AT06A003. ASME.

Goda, Y. 1975 Computation of refraction and diffraction of sea waves with mitsuyasu's directional spectrum. *Tech. Note of Port and Harbour Res. Inst.* **230**, 45.

Hasselmann, K. et al. 1973 Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP). *Deuterium Hydrogen Z* **8**, 1–95.

Hlophe, T., Taylor, P.H., Kurniawan, A., Orszaghova, J. & Wolgamot, H. 2023 Phase-resolved wave prediction in highly spread seas using optimised arrays of buoys. *Appl. Ocean Res.* **130**, 103435.

Hlophe, T., Wolgamot, H., Kurniawan, A., Taylor, P.H., Orszaghova, J. & Draper, S. 2021 Fast wave-by-wave prediction of weakly nonlinear unidirectional wave fields. *Appl. Ocean Res.* **112**, 102695.

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

HLOPHE, T., WOLGAMOT, H., TAYLOR, P.H., KURNIAWAN, A., ORSZAGHOVA, J. & DRAPER, S. 2022 Wave-by-wave prediction in weakly nonlinear and narrowly spread seas using fixed-point surface-elevation time histories. *Appl. Ocean Res.* **122**, 103112.

HUCHET, M., BABARIT, A., DUCROZET, G., FERRANT, P., GILLOTEAUX, J.-C. & DRONIOU, E. 2023 Experimental assessment of a nonlinear, deterministic sea wave prediction method using instantaneous velocity profiles. *Ocean Engng* **281**, 114739.

KIM, I.-C., DUCROZET, G., BONNEFOY, F., LEROY, V. & PERIGNON, Y. 2023 Real-time phase-resolved ocean wave prediction in directional wave fields: enhanced algorithm and experimental validation. *Ocean Engng* **276**, 114212.

KLEIN, M., DUDEK, M., CLAUSS, G.F., EHLERS, S., BEHRENDT, J., HOFFMANN, N. & ONORATO, M. 2020 On the deterministic prediction of water waves. *Fluids* **5** (1), 9.

KÖLLISCH, N., BEHRENDT, J., KLEIN, M. & HOFFMANN, N. 2018 Nonlinear real time prediction of ocean surface waves. *Ocean Engng* **157**, 387–400.

LAW, Y.Z., SANTO, H., LIM, K.Y. & CHAN, E.S. 2020 Deterministic wave prediction for unidirectional sea-states in real-time using artificial neural network. *Ocean Engng* **195**, 106722.

LI, R., ZHANG, J., ZHAO, X., WANG, D., HANN, M. & GREAVES, D. 2023 Phase-resolved real-time forecasting of three-dimensional ocean waves via machine learning and wave tank experiments. *Appl. Energy* **348**, 121529.

MA, Y., SCLAVOUNOS, P.D., CROSS-WHITER, J. & ARORA, D. 2018 Wave forecast and its application to the optimal control of offshore floating wind turbine for load mitigation. *Renew. Energy* **128**, 163–176.

MEISNER, E., GALVAGNO, M., ANDRADE, D., LIBERZON, D. & STUHLMEIER, R. 2023 Wave-by-wave forecasts in directional seas using nonlinear dispersion corrections. *Phys. Fluids* **35** (6), 062104.

MÉRIGAUD, A., HERTERICH, J., FLANAGAN, J., RINGWOOD, J. & DIAS, F. 2018 Incorporating wave spectrum information in real-time free-surface elevation forecasting: Real-sea experiments. In *Proceedings of the 11th IFAC Conference on Control Applications in Marine Systems, Robotics, and Vehicles, Opatija, Croatia, 2018 (CAMS 2018)*, vol. **51**, pp. 232–237. Elsevier.

MÉRIGAUD, A. & RINGWOOD, J.V. 2018*a* Incorporating ocean wave spectrum information in short-term free-surface elevation forecasting. *IEEE J. Oceanic Engng* **99**, 1–14.

MÉRIGAUD, A. & RINGWOOD, J.V. 2018*b* Towards realistic non-linear receding-horizon spectral control of wave energy converters. *Control Engng Practice* **81**, 145–161.

MÉRIGAUD, A. & TONA, P. 2022 A stochastic approach to short-term ocean wave forecasting: Preliminary results using data from a remote sensing imaging system. In *ASME 2022 41st International Conference on Ocean, Offshore and Arctic Engineering*,

NAAIJEN, P., TRULSEN, K. & BLONDEL-COUPRIE, E. 2014 Limits to the extent of the spatio-temporal domain for deterministic wave prediction. *Intl Shipbuilding Prog.* **61** (3-4), 203–223.

OCHI, M.K. 2005 *Ocean Waves: the Stochastic Approach*. Vol. **6**. Cambridge University Press.

PAPOULIS, A. & PILLAI, S.U. 2002 *Probability, Random Variables, and Stochastic Processes*. Tata McGraw-Hill Education.

PIERSON, W.J. 1952 A unified mathematical theory for the analysis, propagation and refraction of storm-generated ocean surface waves, part 1 and 2. NYU, Coll. of Eng., Res. Div. 461. NYU.

QI, Y., WU, G., LIU, Y., KIM, M.-H. & YUE, D.K.P. 2018*a* Nonlinear phase-resolved reconstruction of irregular water waves. *J. Fluid Mech.* **838**, 544–572.

QI, Y., WU, G., LIU, Y. & YUE, D.K.P. 2018*b* Predictable zone for phase-resolved reconstruction and forecast of irregular waves. *Wave Motion* **77**, 195–213.

RAACH, S., SCHLIPF, D., SANDNER, F., MATHA, D. & CHENG, P.W. 2014 Nonlinear model predictive control of floating wind turbines with individual pitch control. In *2014 American Control Conference*, pp. 4434–4439. IEEE.

RINGWOOD, J.V., BACELLI, G. & FUSCO, F. 2014 Energy-maximizing control of wave-energy converters: the development of control system technology to optimize their operation. *IEEE Control Syst. Mag.* **34** (5), 30–55.

SIMANESEW, A., TRULSEN, K., KROGSTAD, H.E. & NIETO BORGE, J.C. 2017 Surface wave predictions in weakly nonlinear directional seas. *Appl. Ocean Res.* **65**, 79–89.

SKAARE, B., NIELSEN, F.G., HANSON, T.D., YTTERVIK, R., HAVMØLLER, O. & REKDAL, A. 2015 Analysis of measurements and simulations from the Hywind Demo floating wind turbine. *Wind Energy* **18** (6), 1105–1122.

SØRENSEN, A.J. 2011 A survey of dynamic positioning control systems. *Annu. Rev. Control* **35** (1), 123–136.

STUHLMEIER, R. & STIASSNIE, M. 2021 Deterministic wave forecasting with the Zakharov equation. *J. Fluid Mech.* **913**, A50.

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press

VERAS GUIMĀRAES, P., ARDHUIN, F., BERGAMASCO, F., LECKLER, F., FILIPOT, J.F.C, SHIM, J.-S., DULOV, V. & BENETAZZO, A. 2020 A data set of sea surface stereo images to resolve space-time wave fields. *Scientific Data* **7** (1), 145.

WOODACRE, J.K., BAUER, R.J. & IRANI, R.A. 2015 A review of vertical motion heave compensation systems. *Ocean Engng* **104**, 140–154.

ZHANG, J., ZHAO, X., JIN, S. & GREAVES, D. 2022 Phase-resolved real-time ocean wave prediction with quantified uncertainty based on variational Bayesian machine learning. *Appl. Energy* **324**, 119711.

https://doi.org/10.1017/jfm.2025.84 Published online by Cambridge University Press