

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Probabilistic access forecasting for improved offshore operations

Ciaran Gilbert^{*}, Jethro Browell, David McMillan

University of Strathclyde, Glasgow, UK

ARTICLE INFO

Keywords:

Wind energy
Offshore access
Probabilistic forecasting
Multivariate forecasting
Forecast visualisation
Generalised additive models

ABSTRACT

Improving access is a priority in the offshore wind sector, driven by the opportunity to increase revenues, reduce costs, and improve safety at operational wind farms. This paper describes a novel method for producing probabilistic forecasts of safety-critical access conditions during crew transfers. Methods of generating density forecasts of significant wave height and peak wave period are developed and evaluated. It is found that boosted semi-parametric models outperform those estimated via maximum likelihood, as well as a non-parametric approach. Scenario forecasts of sea-state variables are generated and used as inputs to a data-driven vessel motion model, based on telemetry recorded during 700 crew transfers. This enables the production of probabilistic access forecasts of vessel motion during crew transfer up to 5 days ahead. The above methodology is implemented on a case study at a wind farm off the east coast of the UK.

© 2020 The Authors. Published by Elsevier B.V. on behalf of International Institute of Forecasters. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Offshore wind is now playing a major role in the portfolio of European electricity markets. As existing sites look toward operating without subsidy and new projects compete in auctions for financial support, reducing the cost of energy is essential. It is estimated that 20%–30% of the total cost of energy for an offshore wind farm is due to Operations & Maintenance (O&M) in the UK (Crabtree, Zappalá, & Hogg, 2015). Since O&M savings can be achieved by operators at any stage of the project life cycle and independently of turbine manufacturers there is a great opportunity to reduce this sizeable percentage. Therefore, improving installation, operations, and maintenance practices is a current focus in both industry and academia.

Access forecasting is concerned with predicting conditions for the transfer of technicians to and from offshore

structures at the site. This is clearly highly dependent on the local wave climate and sea-state forecasting plays a major role in the current scheduling practices in offshore wind. These forecasts are typically deterministic forecasts of significant wave height. However, this provides limited information into the state of the weather over the next few days. Probabilistic forecasts, which quantify the uncertainty around future values, provide a route to risk minimisation (Morales, Conejo, Madsen, Pinson, & Zugno, 2013). For instance, scheduling tools could evaluate the spread of possible metocean conditions for the target day. Additionally, probabilistic forecasts give users more actionable information if the underlying information content is communicated effectively (Bessa et al., 2017).

O&M planning at different horizons can be classified into strategic, tactical, and operational decision making (Shafiee, 2015), the latter of which is the horizon of focus of this paper. Relevant studies in this area of decision support are based on optimising the scheduling

^{*} Corresponding author.

E-mail address: ciaran.gilbert@strath.ac.uk (C. Gilbert).

<https://doi.org/10.1016/j.ijforecast.2020.03.007>

0169-2070/© 2020 The Authors. Published by Elsevier B.V. on behalf of International Institute of Forecasters. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

problems in a metaheuristic manner (Stock-Williams & Swamy, 2018), or via exact optimisation methods (Irawan, Ouelhadj, Jones, Stålhane, & Sperstad, 2017; Kovács, Erds, Viharos, & Monostori, 2011), or through a combination of the two (Dawid, McMillan, & Revie, 2018). Uncertainties in short term maintenance scheduling stem from incorrect fault diagnosis, human error, repair times, and of course the weather forecast. The latter has been examined in the literature by Stålhane, Vefsnmo, Halvorsen-Weare, Hvattum, and Nonås (2016) and Besnard et al. (2011), although most published methods are based on deterministic forecasts of the weather (Dawid et al., 2018; Irawan et al., 2017; Kovács et al., 2011; Raknes, Ødeskaug, Stålhane, & Hvattum, 2017), others have considered climatology from past observations (Dalgic, Lazakis, Dinwoodie, McMillan, & Revie, 2015) and ensemble NWP (Gintautas & Sørensen, 2017).

At lead times of greater than approximately 6 h, Numerical Weather Predictions (NWP) are superior to time series forecasting methods for predicting variables such as significant wave height (Reikard, Pinson, & Bidlot, 2011) and should be employed in day-ahead scheduling decisions (and longer lead-times). Probabilistic forecasting of wave height using time series models, driven by recent observations only, is explored by Taylor and Jeon (2018) and the value of these forecasts is demonstrated via a cost-loss model, but are limited to within day applications. The economic case for improved offshore wind maintenance access forecasts is also discussed by Catterson et al. (2016), where different deterministic models and evaluation metrics are tested; the cost a sub-optimal configuration of these aspects is estimated to cost up to hundreds of thousands pounds per year.

The statistical post-processing of raw weather forecasts to a specific location of interest, using measured data from the site, gives significant improvement in forecast performance (Sweeney, Lynch, & Nolan, 2013). In many cases, engineering additional explanatory features from raw NWP data can significantly improve performance, a practice widely adopted in the energy forecasting community (Andrade & Bessa, 2017; Landry, Erlinger, Patschke, & Varrichio, 2016; Silva, 2014). Within these methods the statistical learning tools perform automatic feature selection and regularisation, which improves on out-of-sample performance and have been successful in international forecasting competitions (Hong et al., 2016).

Offshore operations scheduling is a type of time-dependent decision making; consider that accessibility at a single point in time is not sufficient as technicians require to be picked-up and returned to port at the end of the shift. Therefore, turbine accessibility is often framed in terms of weather windows, where forecasts are used to specify access conditions throughout a specified time period (Gintautas & Sørensen, 2017; O'Connor, Lewis, & Dalton, 2013; Walker, Van Nieuwkoop-McCall, Johanning, & Parkinson, 2013). Browell, Dinwoodie, and McMillan (2017) calculate the probability of access from weather forecasts, which are coupled to cost-loss decision model and compared with the deterministic case; this is found to increase the proportion of access windows utilised and reduce operational expenditure.

Specific types of uncertainty forecasts are required to inform time-dependant decision making. These are termed trajectory, scenario or ensemble forecasts (Pinson, Madsen, Papaefthymiou, & Klöckl, 2009), which maintain the dependency structure between variables and over time. These are required because meteorological forecast errors stemming from NWP are highly structured. This type of forecast is comparable to ensemble-NWP forecasts which can be used in this type of application (Bessa et al., 2017), although ensemble calibration would likely be necessary (Pinson & Girard, 2012). Generating statistical scenario forecasts requires the dependency between the marginal distributions of each lead time to be modelled; the most common method to model the dependency is to use copulas (Pinson et al., 2009).

Numerous bivariate copula families exist, which have been tested rigorously for hindcast metocean data by Fazeres-Ferradosa, Taveira-Pinto, Vanem, Reis, and Neves (2018), where the dependency is modelled between variables including asymmetries. Leontaris, Morales-Nápoles, and Wolfert (2016) use copulas to simulate wave height and wind speed time series, with a case study application for cable installation at an offshore wind farm. The most widely used copula for spatial-temporal forecasting of wind power is the Gaussian copula (Bessa, 2016; Pinson et al., 2009; Tastu, Pinson, & Madsen, 2015), because of the ability to effectively model high dimensional distributions and the apparent absence of tail dependencies in the data. Other feasible methods for such applications include copula vines (Bessa, 2016; Wang, Wang, Liu, Wang, & Hou, 2018) which are able to model more complicated dependency structures, at added computational expense.

The contribution of this work is in providing an end-to-end framework for generating access forecasts based on vessel motion during transfer up to 5 days ahead, including quantifying the uncertainty due to weather conditions. To this end, a method is developed to produce probabilistic forecasts of significant wave height and peak wave period using statistical post-processing of NWP. Temporal and cross-variable dependency is modelled in a copula setting to generate scenario forecasts. These are converted to vessel specific forecasts using a data-driven vessel motion model which captures the displacement of the vessel during transfer, based on a related study (Gilbert, Browell, & McMillan, 2019a). An option for transformation of the vessel motion forecast into an 'access score' is also presented. This score enables simple visualisation and communication of uncertainty information for decision-makers. Additional options for access score visualisation are discussed by Gilbert, Browell, and McMillan (2019b). A flowchart summarising the entire modelling chain in both training and operation phases is shown in Fig. 1.

This paper is organised as follows: Section 2 covers some background on maintenance access and the metocean environment, Section 3 details the forecast post-processing method, Section 4 describes the data driven vessel motion model, followed by Section 5 where results are presented and discussed for a UK offshore wind farm. Future work is outlined and conclusions are drawn in Section 6.

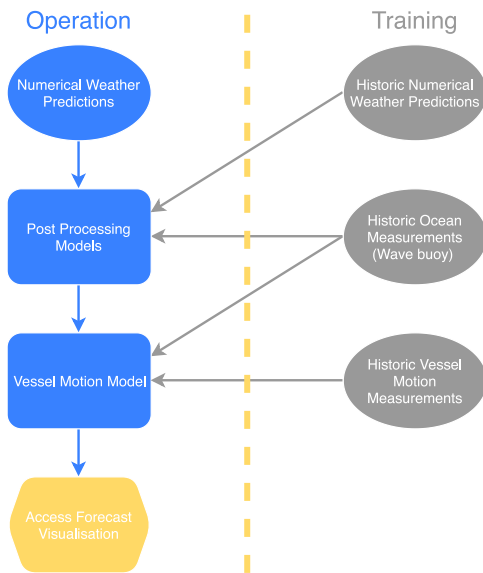


Fig. 1. Flowchart of the modelling chain in operation and training.

2. Access in the offshore environment

Typically, vessel dispatch is managed by a marine coordinator in a control room at the operations base. This work aims to innovate in the decision space where coordinators makes the scheduling/dispatch decision depending on the weather forecast, any available live measurements, accrued experience of the site, and the list of work orders. This schedule is typically made first thing in the morning and then updated at night for the next day accounting for the completed work, any new tasks, and updated weather forecasts. As the day progresses the marine coordinator has to deal with deviations as a result of turbine inaccessibility, technician sea-sickness, or delays; these first two issues are clearly partly due to meteocean forecast errors. Innovations in probabilistic access forecasting are then useful for both marine coordinators and schedulers.

Crew transfer refers to the process of transferring an individual from a vessel to an offshore structure. In the offshore wind industry, it is routine practice for technicians to transfer from dedicated Crew Transfer Vessel (CTV) to a wind turbine via a specially designed ladder which the CTV pushes up against. CTVs are equipped with a rubber fender shaped to fit the ladder and the vessels propulsion system is used to create friction between the fender and the ladder in order to stabilize the vessel. Once stable, the crew member may proceed with the transfer and climb the ladder to the wind turbine. Safety is paramount in this dangerous environment and the individual transferring, the vessel master and marine coordinator all have the power to stop operations if they are deemed unsafe. In contrast, it is common for contractual levers to be in place between asset owners and operators specifying a significant wave height threshold below which CTVs are expected to attempt transfers (Maples, Saur, Hand, Van De Pietermen, & Obdam, 2013).

In order to plan and execute maintenance operations, including crew transfer, forecasts of the sea state are then required. In practice, these typically comprise of significant wave height at 1- or 3-hour intervals for a single location in space, representative of the wind farm, for the next 48 h. In this work, three crucial environmental factors for access quality are forecast: significant wave height, peak wave period, and peak wave direction at 1 h intervals up to 5 days ahead. The spatial resolution is for a single location in space representative of the wind farm. Other factors, such as lightning risk and visibility, are reserved for future work, as well as incorporating information from wave spectra.

3. Sea state forecasting methodology

Here, the method for generating scenarios of significant wave height and peak wave period are detailed. The method for post-processing wave direction for the vessel motion model is also described. The NWP outputs used here include significant wave height, peak wave period, and mean wave direction. It is common practice in contemporary regression models to engineer additional features and use cross-validation for algorithm-specific parameter tuning. Feature engineering is the practice of combining or transforming raw explanatory variables to produce new features with greater explanatory power. For example, the NWP error characteristics suggest that including leading and lagged lead-times will be beneficial. We also consider rolling averages, rolling variances, diurnal, and seasonal effects. However, the significantly increased dimensional of the input space necessitates feature selection techniques and/or regularisation, which is discussed in the proceeding sections. All engineered temporal features are calculated per issue time. A full list of features used in each model can be found in the attached supplementary material.

3.1. Parametric & non-parametric regression

Three methods are considered for producing density forecasts of significant wave height or peak wave period. All are based on post-processing NWP, i.e. learning the relationship between historical observations and concurrent weather predictions. One non-parametric and two parametric density forecasting methods are compared. Parametric techniques assume that the predictive distribution follows a parametric distribution and the forecasting task is to predict the parameters of that distribution, whereas non-parametric techniques make no assumptions of this sort. Typically, kernel density estimation or multiple quantiles of the predictive distribution are used to construct the non-parametric predictive density. Here, gradient boosting machines are used for quantile regression, and two variations of generalised additive models for location, scale, and shape are used to produce parametric density forecasts.

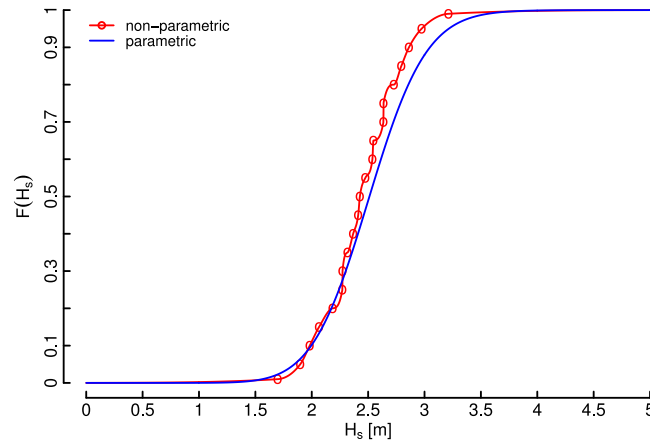


Fig. 2. Example predictive CDF for significant wave height at a single lead time, based on either multiple quantile regression or distributional regression. In the latter, the conditional distribution family is the Generalised Beta Prime.

3.1.1. Gradient boosting machines

The use of gradient boosting machines as a statistical learning technique is motivated by the success of this algorithm in the energy forecasting arena (Andrade & Bessa, 2017; Landry et al., 2016). A predictive model is generated from an ensemble of *weak learners* where, in this case, each learner is a regression tree. The ensemble of regression trees is constructed sequentially by estimating a new tree according to some user-specified differentiable loss function, here quantile loss (Morales et al., 2013). Importantly, the optimisation is solved by steepest descent (Friedman, 2001).

The user must specify the number of trees to fit, the number of splits for each tree, and a shrinkage parameter to reduce the impact of individual trees in the ensemble. For more information on this algorithm the reader is referred to Friedman (2001). The two key parameters tuned to minimise out-of-sample error via k -fold cross-validation are the interaction depth and shrinkage. These levers allow the user to regularise the model and to automatically perform feature selection on high dimensional input data.

The predictive distribution is constructed for each variable and lead-time from multiple quantile forecasts (probability levels: 0.01, 0.05, 0.1, ..., 0.95, 0.99) using cubic spline interpolation with knots at each predicted quantile and fixed boundaries at zero and the maximum value observed in the training data. Monotone cubic spline interpolation is used to guarantee a valid Cumulative Distribution Function (CDF) at each lead time (Fritsch & Carlson, 1980). An example predictive CDF of significant wave height for a single lead time is shown in Fig. 2 using the multiple quantiles and spline interpolation method, as well as a parametric predictive CDF which is discussed in the following section.

Modelling extreme quantiles, those in the tails of the predictive distribution, is challenging due to high estimation error of quantile regression at the extremes and motivates the use of extreme value theory (e.g. Pareto-type tails) and parametric methods. Here we found negligible difference between the described approach and use of Generalised Pareto tails, which require an additional

parameter to be estimated and does not remove the need to impose boundaries.

3.1.2. Generalised additive models for location, scale, and shape

Since the main objective of the sea-state forecasting stage is to produce scenario forecasts, parametric regression models are considered because the tails of the distribution are well defined compared to quantile regression, where tails require special treatment. The tails of the distribution have a large impact on dependency structure estimation and scenario forecast production. Additionally, the full distribution is described by fewer parameters. To this end, two variations of generalised additive models for location, scale, and shape are used. One uses maximum likelihood to optimise the model fit and the other uses boosting; these are termed the `gamLSS` (Rigby, Stasinopoulos, & Lane, 2005) and `gamboostLSS` (Hofner, Mayr & Schmid, 2016; Mayr, Fenske, Hofner, Kneib, & Schmid, 2012) models respectively. The use of boosting in this case is employed to allow for the use of feature engineering as discussed above.

`gamLSS` models are termed ‘semi-parametric’ models, because a parametric distribution is assumed for the target variable and the parameters of that distribution may include non-parametric smoothing functions of explanatory variables — this should not be confused with non-parametric probabilistic forecasts in the form of quantiles. This framework is an extension of the more familiar generalised additive models (Hastie, Tibshirani, & Friedman, 2009) in that any parameter of the distribution can be a function of the explanatory variables, not just the conditional mean.

If we have observations y , in this case significant wave height or peak wave period, the conditional density typically $f_d(y|\theta)$ depends on up to four parameters; these are the location (θ_1), scale (θ_2), and shape parameters (θ_3, θ_4). Distributions with less than four parameters are supported. Note that the time index of the observation from the above is dropped to avoid notational clutter. So, an additive regression predictor η_{θ_i} is generated for each distribution parameter θ_i for $i = 1, \dots, 4$. Let \mathbf{x}_i be the

pool of N_i explanatory variables in the sub-model for θ_i , and $g_i(\cdot)$ the link function, then the model formulation of `gam1ss` is

$$g_i(\theta_i) = \eta_{\theta_i} = \beta_{0\theta_i} + \sum_{n=1}^{N_i} f_{n\theta_i}(x_{i,n}), \quad i = 1, \dots, 4 \quad (1)$$

where the function $f_{n\theta_i}$ is the effect of explanatory variable n on the distribution parameter θ_i , which can be linear or non-linear effects such as penalised splines; $\beta_{0\theta_i}$ are the intercepts of each sub-model. Typically, these models are fitted iteratively using a combination of maximum likelihood, transformation of distribution parameters θ using the inverse link function, and successive back-fitting of the predictor functions in each sub-model η_{θ_i} (Rigby et al., 2005).

However, when \mathbf{x}_i becomes large, feature selection procedures should be carried out to avoid over-fitting and the computational expense of repeated model estimation for feature selection can increase significantly. Model fitting based on component-wise gradient boosting is an attractive solution to this problem (Mayr et al., 2012). Formally, given y_t observations and $\boldsymbol{\eta}_t$ additive predictors of the four sub-models, the `gamboostLSS` (Hofner, Mayr & Schmid, 2016; Mayr et al., 2012) algorithm minimises the loss function $L(\cdot)$ i.e the negative log likelihood

$$\frac{1}{T} \sum_{t=1}^T L(y_t, \eta_{\theta_1 t}, \eta_{\theta_2 t}, \eta_{\theta_3 t}, \eta_{\theta_4 t}) = \frac{1}{T} \sum_{t=1}^T L(y_t, \boldsymbol{\eta}_t) \quad (2)$$

Similarly to the `gam1ss` model, for each distribution parameter a set of base learners $h_{i,n}(\cdot)$ (penalised splines, cyclical splines, linear effects etc.) are specified for each explanatory variable, and the model formula can be different for each predictor. Where $L(\cdot)$ is differentiable, the vector of the negative gradient \mathbf{r}_i is defined as

$$\mathbf{r}_i^{[m-1]} = - \left[\frac{\partial L(y_t, \boldsymbol{\eta}_t)}{\partial \eta_{\theta_i}} \right]_{t=1, \dots, T} \quad (3)$$

where the boosting iteration is m and $\boldsymbol{\eta}_t = \hat{\boldsymbol{\eta}}_t^{[m-1]}$ are the current estimates of the additive predictors. To begin the algorithm, additive predictors $\hat{\eta}_{\theta_i}^{[0]}$ are initialised with offset values. The base learners $h_{i,n}(\cdot)$ are fit to this negative gradient and only the best base learner (n^*), according to the least squares error, is used to update the additive predictor

$$\hat{\eta}_{\theta_i}^{[m-1]} = \hat{\eta}_{\theta_i}^{[m-1]} + \lambda h_{i,n^*}(\cdot) \quad (4)$$

where λ is a shrinkage parameter, or step-length, which is included for regularisation. The additive predictor $\hat{\eta}_{\theta_i}^{[m]}$ is then set equal to $\hat{\eta}_{\theta_i}^{[m-1]}$ and the process is repeated for the remaining θ parameters in this boosting iteration. Following this, the boosting process is repeated until the user specified $m = m_{stop}$ is reached. Therefore, k -fold cross-validation is used in this case to tune the total number of boosting iterations and the value of the shrinkage term. This process is known as component-wise gradient boosting, enabling an intrinsic feature selection capability (as some features are never the best learner n^* and therefore do not form part of the model), which performs well with high dimensional input data.

For both parametric additive models explored, the selection of the base learner is important, and types of learners available are very similar. Taking the `gam1boostLSS` model notation, the typical base learner $h_{i,n}(\cdot)$ specification in this case is a penalised B -spline (i.e. the p -spline), with cyclical splines used for direction variables, and a bivariate p -spline for the seasonal terms – time-of-day and day-of-year – to include the smooth interaction of these variables. Taking the most commonly used base learner as an example, the p -spline is defined as

$$h_{i,n}^{p-s}(x_{i,n}) = \sum_{k=1}^K a_k B_k(x_{i,n}) \quad (5)$$

where the k th B -spline basis function is $B_k(x_{i,n})$, and a_k are the associated spline coefficients. The coefficients are estimated with penalisation to enforce a degree of smoothness to the fit (Eilers & Marx, 1996). The exact method of penalised coefficient estimation varies across the two implementations tested here; the `gamboostLSS` implementation is expanded by Hofner, Kneib and Hothorn (2016), with information on the constrained cases, such as circular variables. Various penalised spline implementations are explored and compared by Perperoglou, Sauerbrei, Abrahamowicz, and Schmid (2019), along with details on the `gam1ss` procedure. For a full description of each model formulation used in the entire analysis, the readers are referred to the supplementary material.

A key component of distributional regression is choosing an appropriate conditional distribution family for the target variable. In the case of wave height and period regression the distribution should support values on the positive real line. A number of candidate families meeting this criteria are tested and the best performing candidate identified by evaluating resulting forecasts in a cross-validation exercise. An example predictive CDF of significant wave height using a candidate distribution family is shown in Fig. 2 for a single lead time, and an example density forecast is shown in Fig. 3(a).

3.1.3. Benchmark models

Two benchmark methods are included as both a ‘naive’ and ‘smart’ comparison. In both cases the target variable is related to a corresponding single input from the NWP source, e.g. significant wave height to significant wave height, and the target variable is assumed to follow a Gamma distribution, as this was found to be a competitive model during exploratory analysis and in related work (Dowell, Zitrou, Walls, Bedford, & Infield, 2014). These models are implemented in `gam1ss`, in which the variance of the distribution is also influenced by the mean due to the parameterization. Therefore, the shape of the distribution is not constant and both models are very competitive. The naive benchmark is a generalised linear model with a single linear base learner, this is termed the benchmark – *glm*. The second benchmark is a simple generalised additive model with a single penalised spline as the base learner, termed benchmark – *gam*.

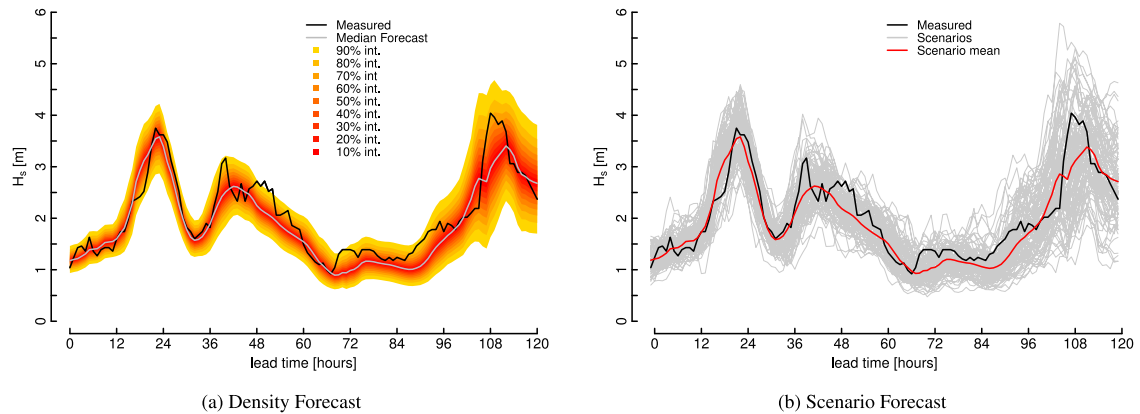


Fig. 3. Example probabilistic forecasts of significant wave height using gamboostLSS parametric regression. In Fig. 3(a) the intervals plotted cover specified probability level ranges, e.g. the 90%int. is the 5% – 95% quantile range.

3.2. Scenario forecasting

Where probabilistic forecasts are used for time-dependant decision making, scenario forecasts are required (Bessa et al., 2017). Here, the Gaussian copula is used (Pinson et al., 2009). The marginals of the copula are the density forecasts for each variable and lead-time. If the random variable Y_h denotes the target variable at lead time h , and y_h is the corresponding realisation, the predictive CDF of the h th lead time is

$$F_h(y_h) = P(Y_h \leq y_h) \tag{6}$$

for $h = 1, 2, \dots, H$ lead times. With continuous marginals $F_i(\cdot)$ there is a unique copula function that describes the H -dimensional cumulative distribution (Nelsen, 2006)

$$F(y_1, y_2, \dots, y_H) = C(F_1(y_1), F_2(y_2), \dots, F_H(y_H)) \tag{7}$$

Therefore, the joint predictive distribution can be estimated using the marginals for each lead time, and the dependence structure via a copula function; this is a useful decoupling. Note the copula function links uniformly distributed marginals $u_h = F_h(y_h)$ and consequently the calibration of the density forecasts has a direct impact on the quality of the dependency estimation. The Gaussian copula is described by

$$C(F_1(y_1), F_2(y_2), \dots, F_H(y_H)) = \Phi_{\Sigma}(\Phi^{-1}(F_1(y_1)), \Phi^{-1}(F_2(y_2)), \dots, \Phi^{-1}(F_H(y_H))) \tag{8}$$

where $\Phi^{-1}(\cdot)$ is the inverse standard normal distribution function and $\Phi_{\Sigma}(\cdot)$ the H -dimensional normal distribution function with covariance matrix Σ and zero mean. In this context a single covariance matrix encodes the temporal dependence structure between all the lead times. Note that

$$v_h = \Phi^{-1}(F_h(y_h)) \tag{9}$$

is a transformation of the uniformly distributed variable into the Gaussian domain where $v_h \sim \mathcal{N}(0, 1)$. Here, the covariance is simply estimated by the sample covariance matrix of the transformed normally distributed variables. Therefore, by sampling from the multivariate distribution H -temporal scenarios of \hat{v}_h are back-transformed to the

uniform domain, and then finally transformed again into the original domain of the target variable using the inverse predictive CDF for the h th lead time. An example scenario forecast of significant wave height based on this technique is shown in Fig. 3(b).

Here three configurations of the dependency are tested: (1) *Independence* – the benchmark where no temporal correlation is embedded in the high dimensional dependence, (2) *Linked* – the full temporal inter-dependency between significant wave height and peak wave period is modelled across the lead times, and (3) *Temporal* – the dependency is modelled for each variable separately across the forecast lead times. The *linked* case is motivated by the idea that the uncertainty is linked between the variables, because they are summary statistics from the same source wave spectrum forecast; again the cross-variable dependency matrix is simply estimated using the sample covariance matrix of the transformed normally distributed variables.

3.3. Wave direction regimes: Clustering & logistic regression

Wave direction can have a significant impact on vessel motion and on the characteristics of NWP errors, particularly if wave direction is associated with different physical processes. This section describes the wave direction post-processing strategy. For the purpose of the vessel motion model, a small number of distinct directional regimes are considered rather than incorporating direction as a continuous variable. Peak wave direction is incorporated by first clustering the wave buoy measurements into two distinct regimes, motivated by the fact that the wave climate at the case study location is dominated by locally driven wind waves or waves from the swell, though this technique could easily be extended to wave climates with more than two distinct regimes. Logistic regression is then used with NWP to predict the cluster membership at any forecast lead time. Therefore, wave direction prediction is simplified into a straightforward classification problem.

To cluster the measured variables, we define the input space $z_t = (\omega_{p_t}, T_{p_t}, H_{s_t})$, where the three environmental factors are peak wave direction, peak wave period, and

significant wave height respectively. The k -means clustering algorithm is used to define the two regimes (Hastie et al., 2009). This algorithm generates disjoint regions R_k that collectively cover the input space spanned by z_t . Note that all input variables are scaled and the wave direction variable is linearised. Since, in this case there are only two regimes defined, logistic regression is used to determine the probability of regime membership. The gradient boosting machine described above is used as the logistic regression tool with inputs features engineered similar to the continuous target variable regression case. Again, a full list of input features for the wave direction regime forecasting can be found in the supplementary material. For more information on the regression technique, please refer to Friedman (2001).

4. Vessel motion during transfer

Accessibility is constrained by vessel motion during push-on and transfer. Therefore, in order to provide forecasts of accessibility to both wind farm and vessel operators it is necessary to forecast the sea conditions and understand how individual vessels will respond in those conditions. Here a data-driven approach to vessel motion modelling is undertaken, and for more information on this part of the modelling the reader is referred to Gilbert et al. (2019a).

Distributional models using generalised additive models for location, scale and shape, described in Section 3, are used to learn the relationship between met-ocean observations and the vessel motion data (heave peak-to-peak) during push-on instances. The main mode of movement during push-ons which impede transfers is vertical displacement of the vessel fender and the turbine transition piece due to the oncoming wave field, this can result in a 'slip' event which can have serious safety implications. This motivated the use of the heave motion of the vessel as a key transfer quality and safety indicator. The other degrees-of-freedom can clearly have an impact on transfer quality, but this is reserved for future work.

In operation, forecasts from Section 3, i.e. the scenarios of wave height and period, as well as the forecast regime membership, are used as inputs to drive the vessel motion model; this process generates vessel-specific scenarios of motion during transfer. A visualisation stage, discussed briefly in Section 5 and more extensively by Gilbert et al. (2019b), completes the forecasting process, as shown in Fig. 1.

5. Case study

The methodology is tested at an east coast offshore wind farm in the UK. Ocean measurements are collected from a Centre of Environment Fisheries and Aquaculture Science wave-buoy within the site boundary. NWP's of the wave climate from the European Centre for Medium-Range Weather Forecasts are extracted at the closest grid point to the site from 0-120 h ahead at hourly intervals, with 2 issue times per day.

Vessel telemetry data is from two purpose built offshore wind service vessels with the same specification:

length 19.2 m, width 8.2 m, maximum draft 2 m, passengers 12, aluminium catamaran; this data is collected during the construction phase of the wind farm, which contains around 700 push-on instances alongside concurrent wave buoy measurements. Transfer events are identified using the measured push-on force as well as time-stamped *swipes* from technicians' ID cards when transferring. The vessel's average peak-to-peak heave is determined during each push-on attempt and used in the following analysis. The time resolution is reconciled by matching transfer events to the closest buoy measurement in time.

For the regression problems the data is partitioned into 4.5 (January 2013 - June 2017) and 1 (July 2017 - June 2018) year(s) for training and testing respectively. This allows for sufficient data in the modelling of the copula and an entire year to evaluate the subsequent forecasts in out-of-sample tests. To make decisions on the best configuration for each forecasting task and to tune the algorithm specific hyperparameters, 4-fold cross-validation is used on the training data only. To refresh, the four main forecasting tasks are: (1) significant wave height regression, (2) peak wave period regression, (3) wave regime clustering & logistic regression, and (4) copula dependency modelling. The method is implemented in R (R Core Team, 2016) using the package ProbCast, which is in development, although a 'beta version' is available with accompanying scripts for this methodology (Gilbert, Browell, & McMillan, 2020). ProbCast is developed for the modelling, evaluation, and plotting of probabilistic forecasts, using `gbm`, `gamLSS`, and `gamboostLSS` for the regression models (Greenwell, Boehmke, Cunningham, & GBM Developers, 2019; Hofner, Mayr & Schmid, 2016; Rigby et al., 2005).

Probabilistic density forecasts of significant wave height and peak wave period are evaluated according to the principle that the forecast should be optimally sharp subject to calibration (Gneiting, Balabdaoui, & Raftery, 2007); sharpness evaluates the spread of the distribution and calibration (i.e. reliability) demands that the empirical spread of forecast matches the observations. Implementing quantile and parametric regression for density forecasting introduces a compromise when using evaluation metrics for comparison. Here the Probability Integral Transform (PIT) is used to measure the calibration of the full distribution, motivated by the direct impact of this variable on the dependency structure, and the pinball loss score is used to measure the sharpness and calibration at discrete quantiles of the distribution. The PIT measure is

$$u_t = F_t(y_t) \quad (10)$$

and if the forecasts are well calibrated and the sample is sufficiently large then u should be uniformly distributed, which is visually inspected via a histogram. For target quantile α and corresponding forecast $\hat{y}_t^{(\alpha)}$ the pinball loss score is

$$PB_t^{(\alpha)} = \begin{cases} (1 - \alpha)(\hat{y}_t^{(\alpha)} - y_t) & y_t < \hat{y}_t^{(\alpha)} \\ \alpha(y_t - \hat{y}_t^{(\alpha)}) & y_t \geq \hat{y}_t^{(\alpha)} \end{cases} \quad (11)$$

and the score is averaged over all lead times. Note that the quantile regression models will be directly optimised to minimise this score. It is also important to understand the growth in the uncertainty of the density forecasts as a function of lead time. Here, sharpness is tested in terms of average interval width plots (Pinson, Nielsen, Møller, Madsen, & Kariniotakis, 2007). For an interval with a nominal coverage rate of $1-\beta$ the interval size is

$$\delta_t^{(\beta)} = \hat{y}_t^{(1-\beta/2)} - \hat{y}_t^{(\beta/2)} \quad (12)$$

and this measure is averaged over all cases, grouped by each lead time. It is important to note that this measure of sharpness is a final illustrative layer to the forecast which is proven to be sharp and reliable via the pinball loss and PIT histogram respectively.

Scenario forecasts are evaluated via multivariate probabilistic forecast verification methods. Two metrics capable of evaluating the trajectories are the Energy Score (ES) developed for meteorological applications (Pinson & Girard, 2012), and the p -Variogram Score (VS- p) (Scheuerer & Hamill, 2015), which has greater discrimination ability when focusing on the underlying dependence structure generating the multivariate probabilistic forecasts (Bessa, 2016; Tastu, 2013); both are evaluated per issue time of the forecast. The ES is given by

$$ES = \frac{1}{J} \sum_{j=1}^J \|\mathbf{y} - \hat{\mathbf{y}}^{(j)}\|_2 - \frac{1}{2J^2} \sum_{i=1}^J \sum_{j=1}^J \|\hat{\mathbf{y}}^{(i)} - \hat{\mathbf{y}}^{(j)}\|_2 \quad (13)$$

where $\|\cdot\|_2$ represents the ℓ_2 norm, \mathbf{y} is the observed time trajectory of the measured variable, $\hat{\mathbf{y}}^{(j)}$ is the j th scenario forecast, and J is the total number of samples taken from the underlying multivariate distribution. The p -variogram score is

$$VS-p = \sum_{i,j=1}^H w_{ij} \left(|y_i - y_j|^p - \frac{1}{J} \sum_{z=1}^J |\hat{y}_i^{(z)} - \hat{y}_j^{(z)}|^p \right)^2 \quad (14)$$

where p is the order of the variogram, H is the length of the trajectory vectors, and $\hat{y}^{(z)}$ is the z th forecast scenario. Here, the weights, w_{ij} , are set to the inverse square distance between the i th and j th components. It is clear that this score is based on pairwise differences. Results in Scheuerer and Hamill (2015) show that a VS- p with an order of less than one has the best discriminative ability.

The regime classification forecasts are evaluated in terms of the Area Under the Receiver Operator Curve (AUROC) (Fawcett, 2006). In evaluating logistic regression, the ROC curve plots the true positive rate versus the false positive rate for different threshold values at which the predicted probability is cut to define the two prediction classes. The maximum area AUROC is equal to one, and optimal forecasts are as close to this value as possible.

5.1. Sea state forecasting

Here we detail the results of the wave height, period, and direction forecasting tasks including density, scenario, and regime membership forecasting.

5.1.1. Parametric & non-parametric regression

To make decisions on the best regression technique for significant wave height and peak wave period, the calibration and sharpness are compared by 4-fold cross-validation on the training data. Cross-validation is also used to tune the model hyperparameters of the boosting methods and to refine the model formula for the `gamLSS` model. For the parametric regression techniques, the conditional distribution of the measured variable is also defined using cross-validation. A large range of distributions with $(0, \infty)$ support were tested. For both significant wave height and peak wave period the Generalised Beta Prime distribution (Rigby et al., 2005; Stasinopoulos & Rigby, 2007) produced the best forecasts in terms of sharpness, subject to calibration. It is a flexible four parameter distribution, which nests other common distributions. For regression model selection in the case study, the calibration of the density forecasts is crucial due to the direct impact on the dependency structure and therefore scenario generation quality.

For the significant wave height density forecasting task, Figs. 4(a) and 4(b) show the pinball loss in both cross-validation and testing respectively, which reveals that the two boosting models reduce the pinball loss across the quantiles evaluated, compared to the `gamLSS` model, although to a lesser extent at the tails of the distribution. Comparing the `gbm` and `gamboostLSS` techniques the former gives lower error scores in cross-validation and the latter in testing, although the differences are generally minor and only evident in the $p30 - p70$ range. The *benchmark - gam* model is very competitive, and provides significant improvement over the naive *benchmark - glm*; the only difference in these models is a simple change of base learner from a linear effect to a penalised spline.

In Figs. 4(c) and 4(d) the PIT histograms are presented; comparatively, the advanced parametric regression techniques result in better calibrated forecasts, especially in the tails of the distribution because of the difficulty in estimating quantile regression models in this region. Both benchmark models here show poor calibration, specifically under-confidence. Comparing the *benchmark-gam* and the `gamLSS` model, the reliability is much improved for latter even though the models are somewhat similar; this validates the choice of the Generalised Beta Prime distribution. However, the two also have different input features, which can be found in the supplementary material.

Based on the cross-validation results, the density forecasts of significant wave height obtained via `gamboostLSS` regression are selected for use in the later stages of the access forecasting process, based on the principle that density forecasts should be sharp subject to calibration. The `gbm` approach has a lower pinball loss, but the calibration is poor, notably in both tails of the predictive distribution, and is excluded as a result. The `gamLSS` approach is well calibrated but has a higher pinball loss than `gamboostLSS`.

For peak wave period regression, the pinball loss scores are shown in Figs. 5(a) and 5(b) for cross-validation and testing. In cross-validation, the lowest pinball loss is more clearly defined using the gradient boosting machine regression technique. However, in testing the two boosting

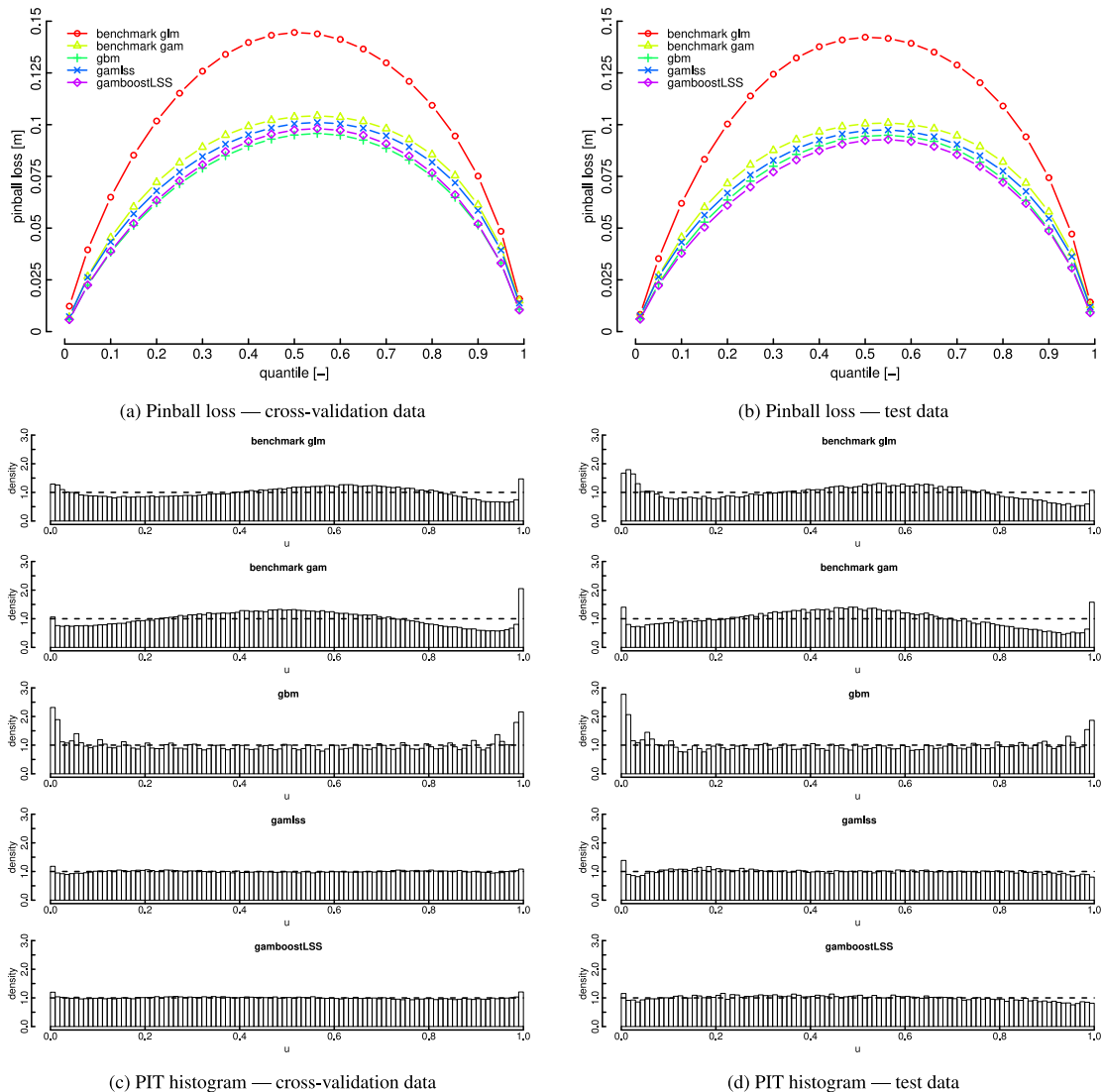


Fig. 4. Results for all lead times – significant wave height density forecasting by regression model.

models return very similar pinball loss scores across all the tested quantiles. It should be emphasised that the gbm quantile models are directly optimised to minimise this score. A similar behaviour to the wave height case is found for two benchmark models.

Figs. 5(c) and 5(d) detail the PIT histograms for the peak wave period regression models under testing and cross-validation. Clearly again in this case, the calibration of the gbm model in the tail region is comparatively poor against the two advanced parametric regression techniques. However, in testing all of the models here present deviations from uniformity. The calibration of two benchmark models is again found to be poor. Based on the cross-validation results, the forecasts based on the gamboostLSS regression are selected for implementation due to producing sharp forecasts subject to calibration.

In Fig. 6 the sharpness, or average interval width, is plotted against forecast lead-time for both significant wave height and peak wave period during testing. The

forecast models evaluated are the final gamboostLSS densities chosen for further implementation. They show that the uncertainty grows with lead time, which is to be expected. This is particularly pronounced for significant wave height, shown in Fig. 6(a), where the average interval size grows considerably; the 90% interval more than doubles in average width from 0.5 m at issue time to over 1.2 m at 120 hours-ahead. For this reason, and because we are motivated by day-ahead decision-making, we only consider lead-times of up to 120 h despite NWP with lead-times of 120 h–240 h being available. For wave period in Fig. 6(b), the growth of the interval size is not as pronounced, however the interval widths have a greater spread at the earliest lead times.

An important aspect of using the gamboostLSS method, is that it is very memory intensive; it was not possible to model the density forecasts using a conventional desktop computer (8 virtual cores, 3.6 GHz CPU, 16 GB RAM), a cloud instance was used instead. However,

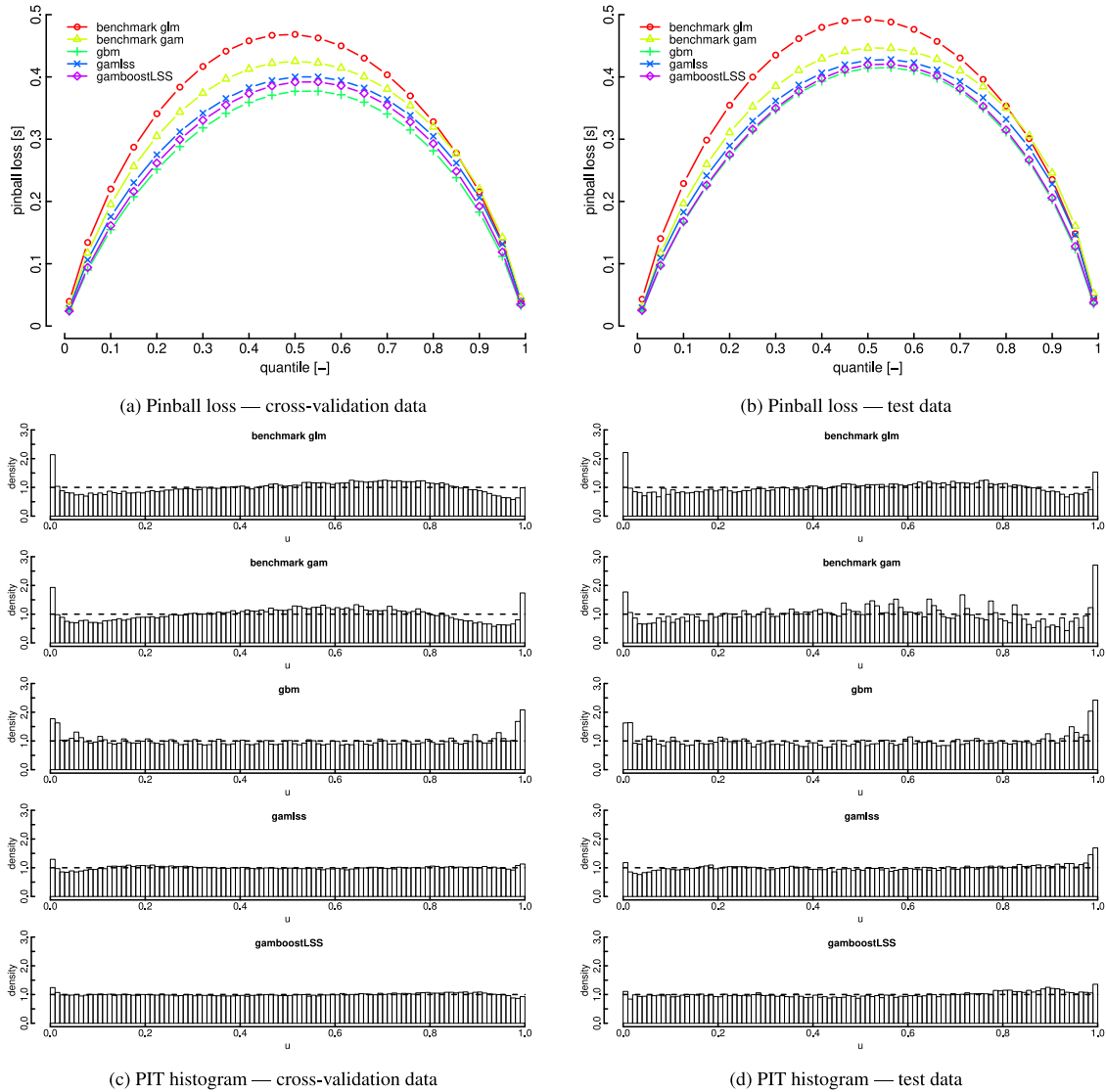


Fig. 5. Results for all lead times – peak wave period density forecasting by regression model.

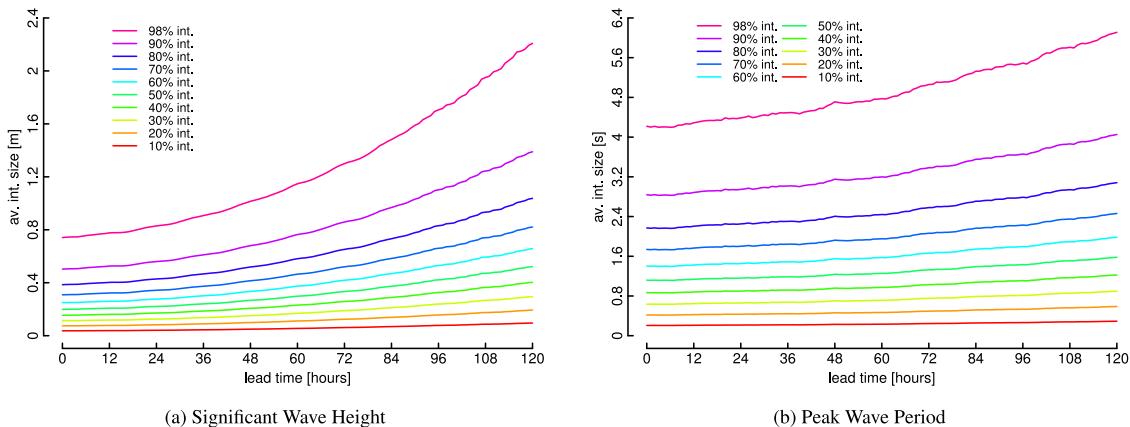


Fig. 6. Evaluation of the sharpness for gamboostLSS density forecasts vs lead-time during testing. The width of the prediction intervals increases with lead-time as the forecasts becomes less confident further into the future.

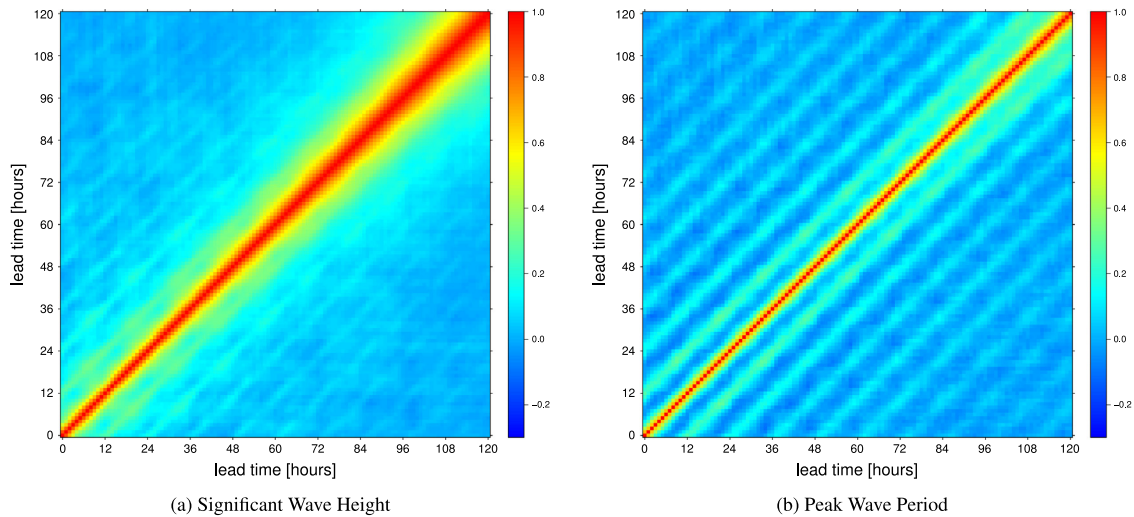


Fig. 7. Temporal correlation matrices during testing.

Table 1

Multivariate forecast evaluation results. Linked means the temporal dependency between significant wave height and peak period is modelled. Best results during testing are in bold.

Variable	Dependency	Data	ES	VS-1	VS-0.5
Wave height	Temporal	CV	2.06	3.11	5.76
		Test	1.97	3.32	5.48
	Linked	CV	2.06	3.10	5.76
		Test	1.98	3.32	5.48
	Independent	CV	2.09	12.71	15.93
		Test	2.01	13.62	15.49
Wave Period	Temporal	CV	7.84	105.13	34.52
		Test	8.38	130.68	40.21
	Linked	CV	7.84	105.13	34.51
		Test	8.38	130.67	40.21
	Independent	CV	7.87	169.47	51.65
		Test	8.40	190.5	56.93

this problem can be significantly reduced by reducing the number of input features, reducing the number of cross validation folds, reducing the number of boosting iterations, and by using a distribution defined by fewer parameters. The gbm models were fit using a desktop, although 21 quantile models are required for each fold. Finally the gamlss model is computationally cheap, due to the reduced number input variables. Operationally, the time required to issue a forecast is negligible and re-training models would be required infrequently.

5.1.2. Scenario forecasting

As discussed in Section 3.2, for scenario forecasting three configurations are tested: *independence* is a benchmark, *linked* indicates that the full temporal inter-dependency between the significant wave height and peak wave period is modelled, and *temporal* is where the dependency is modelled for each variable separately across the lead times. Again, 4-fold cross-validation is used to determine the most appropriate dependency structure.

In Table 1 the scenario forecast scores are presented. For significant wave height it is evident in training and

cross-validation there are improvements in modelling the temporal dependency structure across all scores; however, there is no real improvement in the linked case compared to the temporal dependency. Comparing the scores, as expected the improvement of modelling the temporal dependency is more significant in the case of both the variogram scores than the energy score because of the better discriminative ability of covariance structures of the former. For evaluation of these forecasts 1000 scenarios are used because, empirically, the mean and standard deviation of the energy score were found to stabilise around this point. In the case of temporal scenarios for significant wave height it takes approximately 0.6 s per issue time to generate 1000 samples and transform them into the original domain using the desktop described previously.

Table 1 also details the results of the peak wave period scenario forecasting; these follow a similar profile to the significant wave height case. Again, the linked dependency structure doesn't prove more valuable than the temporal case across the scores detailed; the temporal dependency also provides improvements against the independent case, especially when measuring improvement via the variogram scores.

The temporal correlation matrices for both wave height and period scenario forecasting cases are plotted in Fig. 7; these are the matrices used to generate scenarios over the test dataset. Here, the dependency characteristics of the uncertainty across the lead time differs between the variables, in that the correlation between consecutive lead times is clearly more prominent in the wave height case. However, they also share some subtle characteristics; the correlation strength persists to a greater extent further ahead in the forecast lead times, and there is clearly a strong diurnal pattern, especially in the peak wave period case. Note that time-of-day, seasonal, and interaction effects are included in the marginal distributions regression formulation.

The energy score improvements for the wave height scenario forecasting task are plotted in Fig. 8, compared

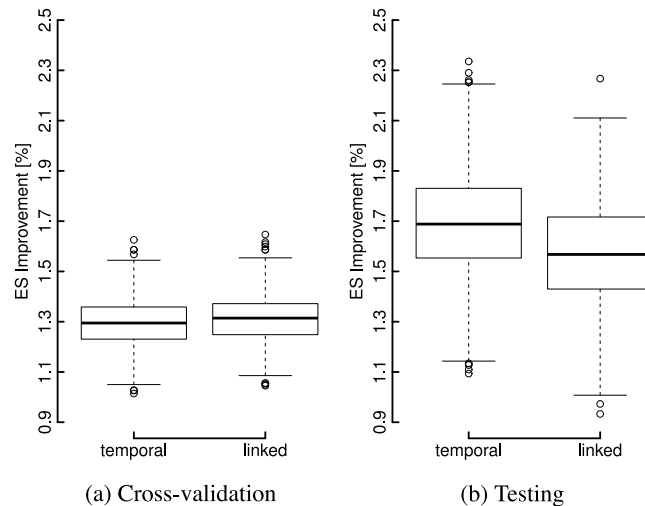


Fig. 8. Boxplots showing the block-bootstrap sample distributions of Energy Score improvement for significant wave height scenario forecasting. The benchmark is the *independence* case, with no temporal correlation. Linked means the temporal inter-dependency between significant wave height and peak period is modelled.

against the independence benchmark. Here, a simple block-bootstrapping approach is used to estimate the significance in the forecast improvement (Efron & Tibshirani, 1994). The scores are split into non-overlapping blocks of 7-days length, to account for correlation in the governing weather patterns. These blocks are re-sampled with replacement and then forecast improvement is determined. This process is repeated 1000 times to estimate the sampling variation of the score improvement in Table 1. The results are presented via boxplots in Fig. 8 and clearly illustrate that modelling the temporal dependency is valuable; the sampling variation is greater during testing because of the smaller size of the dataset. The linked dependency structure shows no significant improvement to the temporal dependency case in cross-validation.

5.1.3. Clustering & logistic regression

Wave direction forecasting for input to the vessel motion model is reduced into two stages, clustering the measured wave buoy data and then applying logistic regression to predict regime membership based on NWP data. Please refer back to Section 3.3 for more details. When using *k*-means clustering, the random assignment of the cluster centres at the start of the algorithm can lead to different results if the data-set is small or not amenable to a clustering algorithm. Therefore, in this case the algorithm was set with different random seeds over 4 tests and 98.8% of data points in a large data-set (>140,000 rows) are assigned to the same cluster. The defined regimes are plotted in Fig. 9(a) via a parallel coordinate plot which allows for visualisation of circular variables (Will, 2016). Clearly the two regimes are separated mostly by peak wave direction; waves from the north east which are driven by the swell and waves driven by the prevailing south westerly winds. The swell driven waves also on average have longer periods and both regimes have similar average wave heights. The repeatability of the clustering using the measured data and

the physical explanation give confidence in the defined regimes at the site.

Post-processing results for the logistic regression are shown in Fig. 9(b) via the ROC curve for the testing and cross-validation phases; the AUROC is very close to one for both cases, being 0.97 and 0.95 in cross-validation and testing respectively. This means that for an optimally defined threshold probability, the classification provides a high rate of positive predictions when the measured value is positive, and a low rate of positive predictions when the measured value is negative. The optimal threshold probability which is used to split the predicted probability space into the predicted regimes is defined by the one which maximises accuracy, so assuming there is an equal weighting to false positive predictions and the false negative predictions.

5.2. Vessel motion

Here, we explore the mapping of the relationship between the measured significant wave height at the buoy and vessel motion measurements during transfer. These results are described in more detail in Gilbert et al. (2019a). Several vessel motion models are tested, starting from basic linear regression models and leading to truncated regression models with penalised smooth base learners in the *gamlss* environment. Truncated regression is used to respect the reality that vessel heave peak to peak measurements are always positive. The Akaike Information Criterion (AIC) is used to measure goodness of fit

$$\text{AIC} = 2k - \ln(\hat{L}), \quad (15)$$

which rewards the model with the highest likelihood function \hat{L} , penalised by the number of parameters k used to estimate the model; overfitting is then less likely for the model with the minimum AIC. As shown in Table 2, the minimum AIC of the models tested is found using a

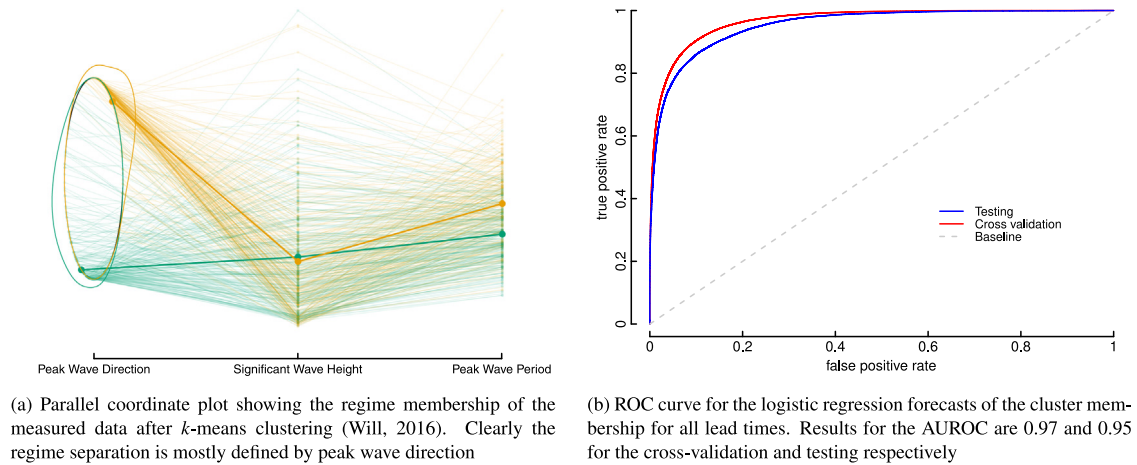


Fig. 9. Sea-state regime classification plots – regime forecasts are used in the vessel motion model.

Table 2

Selected formulations for the vessel motion model during transfer dependent on observed sea-state. The symbols *ps* indicates a penalised beta spline, † is a varying coefficient model, *poly* is a fractional polynomial model, H_s is significant wave height, T_p is peak wave period, ω_p is peak wave direction, and *reg.* is the regime membership.

ID	Distribution	Formula & Explanatory variables			AIC
		Location	Scale	Shape	
N-1	Normal	$\sim H_s$	~ 1	n/a	-1653
N-2	Normal	$\sim H_s + T_p$	~ 1	n/a	-1687
N-3	Normal	$\sim H_s + T_p * H_s$	~ 1	n/a	-1695
N-4	Normal	$\sim H_s + T_p * H_s$	$\sim H_s$	n/a	-1787
tr-N-1	Tr. Normal [0, ∞)	$\sim H_s + \text{poly}(T_p + T_p^2)$	$\sim H_s + \mathbf{1}(\omega_p < 120)$	n/a	-1823
tr-T-1	Tr. <i>t</i> -family [0, ∞)	$\sim H_s + \text{poly}(T_p + T_p^2)$	$\sim H_s + \mathbf{1}(\omega_p < 120)$	$\sim H_s$	-1827
tr-T-2	Tr. <i>t</i> -family [0, ∞)	$\sim H_s + \text{ps}(T_p)$	$\sim H_s + \mathbf{1}(\omega_p < 120)$	$\sim H_s$	-1829
tr-T-3	Tr. <i>t</i> -family [0, ∞)	$\sim \text{ps}^\dagger(H_s, \text{reg.}) + \text{ps}(T_p)$	$\sim H_s + \text{reg.}$	~ 1	-1830
tr-T-4	Tr. <i>t</i> -family [0, ∞)	$\sim \text{ps}^\dagger(H_s, \text{reg.}) + \text{ps}(T_p) + T_p * H_s$	$\sim H_s + \text{reg.}$	~ 1	-1838
tr-T-5	Tr. <i>t</i> -family [0, ∞)	$\sim \text{ps}^\dagger(H_s, \text{reg.}) + \text{ps}(T_p) + T_p * H_s$	$\sim \text{ps}^\dagger(H_s, \text{reg.})$	~ 1	-1846

student-*t* distribution, truncated at 0. Note that not all of the models tested are detailed here for brevity, only a selection to understand the model development.

The marginal effect plot of significant wave height is shown in Fig. 10 for the best model (tr-T-5) with peak period and regime membership held constant; the motivation behind using a model with conditional heteroscedasticity is clear, as the uncertainty grows with significant wave height. Importantly, this model uses penalised varying coefficient splines for the location and shape parameters, where the coefficients vary by the regime membership. This allows for the regime membership to influence the model fit more flexibly than varying the intercept terms and the resulting shape of the uncertainty to change depending on the regime membership.

The estimated model structure can be further examined in the supplementary material; here there are two interactive 3d plots which show the model fit against both wave height and peak wave period, as well as in the different regime classes. For a more in-depth discussion of these results the reader is referred to Gilbert et al. (2019a).

5.3. Forecasting vessel motion during transfers

Here, the outputs of the access forecasting methodology are described. From Fig. 1 the last stage of the modelling is a visualisation stage, which is described. This involves transforming the raw vessel forecast scenarios to make the uncertainty information more interpretable.

5.3.1. Raw output

To generate the vessel motion scenarios, the forecast sea-state scenarios and forecast regime membership are used to drive the vessel motion model. This allows for a forecast estimation of the heave motion of the vessel during transfer. In the presented case, the mean output of the vessel motion model is taken for each scenario input by sampling each generated distribution due to the asymmetric nature of the truncated distribution. A scenario forecast, generated via the mean output of the vessel motion model, is presented in Fig. 11 and illustrates the motivation behind the visualisation stage; this uncertainty information is difficult to interpret by any decision maker. However, the raw output could be useful for driving scheduling optimisation tools for instance.

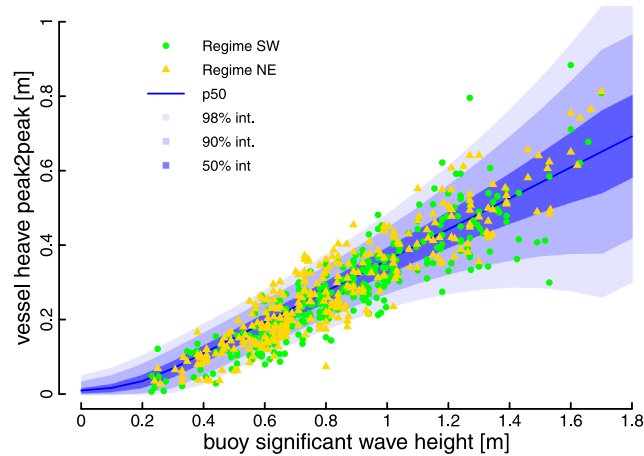


Fig. 10. Plot of vessel motion during push-on and concurrent ocean measurements. The distributional model fit is for the tr -T-5 model in the south west regime with a fixed peak period. Note that push-ons only occur in a small range of the possible significant wave height forecast values and the variance is capped at the highest observed significant wave height. See Sections 5.3.1 and 5.3.2 for more detail on dealing with this effect.

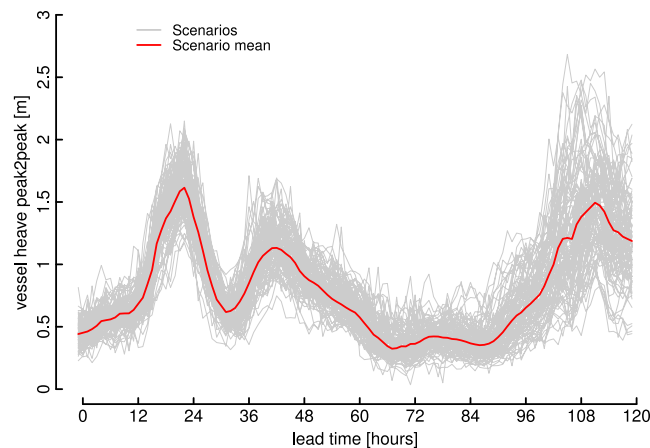


Fig. 11. Example scenario forecast of mean vessel motion during transfer.

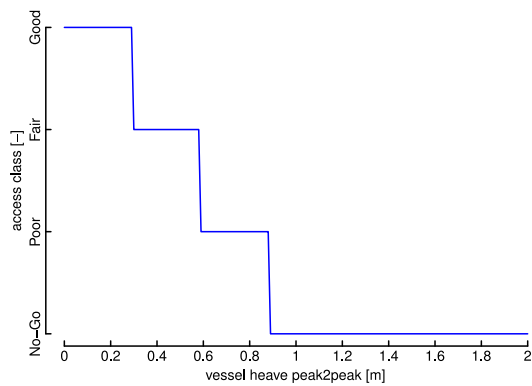
An important subtlety of the vessel motion model presented in Fig. 10 is that data is only collected in a sub-range of the significant wave height marginal distribution at the site; this is a result of push-ons only being attempted in conditions conducive to safe transfer. For forecast values outside this range a threshold significant wave height feature (equal to the maximum observed significant wave height during push-on) is used for the shape parameter in the vessel motion model. This means that stable predictions are obtained across the full distribution of forecast significant wave height values at the site. However, the vessel forecast scenarios above this threshold must be then processed to represent a zero chance of transfer.

Uncertainty information during transfers is explored here, although the conditional expectation of displacement is used for the operational forecast. Future work should consider the incorporation of uncertainty in the vessel motion model, for instance by sampling the vessel motion distribution.

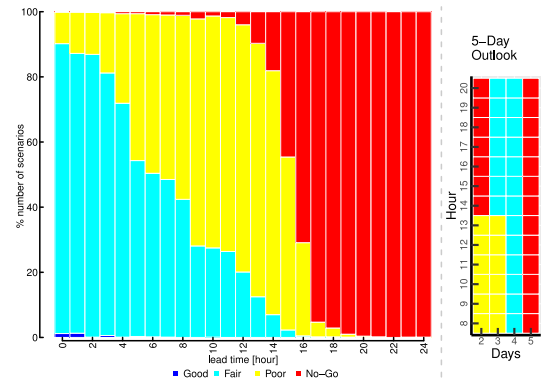
5.3.2. Visualisation

Numerous visualisation options are possible based on the forecast from Fig. 11. For a more in depth discussion of this stage based on continuous, classification, and threshold transfer quality forecasts the reader is referred to Gilbert et al. (2019b). The motivation and end-user requirements must be considered before visualisation options are explored; end-users require the forecast and associated uncertainty to be communicated as simply as possible, and the motivation is to make the forecast interpretable, as well as processing the scenarios so that those which have a zero chance of success, according to the historical data, are conveyed as such.

Here, a user-defined function for transforming the vessel scenarios into classes is shown in Fig. 12(a). This transformation is flexible and can be based on the vessel capabilities, specific mission, experience of the site, and appetite for risk. Additionally, to distil the information content the detailed forecast visualisation focuses on the



(a) User-defined function for transforming mean vessel motion scenarios into a transfer quality forecast



(b) Example visualisation option of Figure 11 as a transfer quality forecast with a focus on the upcoming day on the left panel and further horizons summarised in the right panel

Fig. 12. Visualisation stage plots.

upcoming day with extended lead times summarised to the right of the main plot.

An example of the resulting transfer quality forecast is shown in Fig. 12(b). Here the bars quantify the percentage of scenarios belonging to each class at every time step. Details on conditions further into the horizon are shown via a panel plot where the colour of each panel indicates the majority class over the next 4 work-days. The advantage of the classification method is that every scenario is accounted for and therefore the end user views a complete picture of the spread of possibilities at every time step.

6. Conclusions & future work

This work describes a novel forecasting solution for predicting safety-critical conditions during transfer for offshore operations with a case study at an east coast wind farm in the UK. The proposed access forecasts predict vessel motion during transfer, accounting for weather uncertainty, up to 5-days ahead. Sharp and calibrated density forecasts of peak wave period and significant wave height are generated by post-processing Numerical Weather Predictions, with boosted generalised additive models for location, scale and shape outperforming non-parametric methods. Scenario forecasts of these variables have then been produced using the Gaussian copula to model temporal dependence and used as inputs to a data-driven vessel motion. Modelling cross-variable dependency added no value in terms of the multivariate skill scores. A method of visualisation of these forecasts is also suggested to best communicate the information content for end users.

Future work on the methodology should consider the feasibility of a turbine (or region) specific forecast for large wind farms where access is constrained by local bathymetry. Alternatives to the vessel motion model where data is not available should be considered, as well as investigating the value in transforming the motion to the point of contact with the fender and ladder. Embedding the forecasts into a schedule optimisation tool or

a cost/loss model with a corresponding power forecast could further support offshore wind farm operations and be used off-line to demonstrate the value in accounting for the uncertainty in access conditions.

Regarding sea-state forecasting, the dependency structure of the scenario forecasts could also be made conditional on the dominant forecast direction regime at the site, which could improve the quality of the scenarios. The high-dimensional nature of the dependency structures mean that alternative copulas are somewhat limited. Options include: copula vines, though this would increase computation cost significantly for both fitting and sampling; the empirical copula, in which training data are re-sampled to produce scenarios; or using the Gaussian copula with parametric covariance matrices, though the diurnal patterns observed in the empirical covariance (Fig. 7) suggest that this would not be trivial.

Extending the forecast horizon beyond 5 days is of interest, though it may be necessary to consider lower temporal resolution forecasts, e.g. daily accessibility, due to the reduced skill in NWP at these lead-times.

Some important meteorological factors that restrict offshore access have not been considered here, such as lightning, visibility, and surface wind speeds; forecasts of these should be provided to decision-makers. The forecasting in this paper is based on summary statistics of the wave field; an interesting extension would be to use the forecast wave spectrum from the NWP, which would enable a more complete picture of expected conditions.

Acknowledgments

This work is supported by the EPSRC Supergen Wind Hub project ORACLES, EP/L014106/1. The authors would like to thank the editor, associate editor, and anonymous reviewers who helped improve the quality of this paper with insightful comments and suggestions. The authors would also like to thank Amine Hadjer, John Best, and Pete Leach at James Fisher Marine Services, Lars Johanning and Giovanni Rinaldi at University of Exeter, Sally Shenton at Generating Better, and Christopher Briggs at

SiemensGamesa for their contribution to discussions and feedback throughout this work.

Data Statement: The R code used in this work is available in Gilbert et al. (2020). Numerical weather predictions from ECMWF are available at www.ecmwf.int for qualifying research institutions or commercial customers. Wave buoy measurements from the Centre of Environment Fisheries and Aquaculture Science (CEFAS) are available at <http://wavenet.cefas.co.uk/>. Due to confidentiality agreements with research collaborators, access to vessel motion data is restricted.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.ijforecast.2020.03.007>.

References

- Andrade, J. R., & Bessa, R. J. (2017). Improving renewable energy forecasting with a grid of numerical weather predictions. *IEEE Transactions on Sustainable Energy*, 8(4), 1571–1580.
- Besnard, F., Patriksson, M., Stromberg, A., Wojciechowski, A., Fischer, K., & Bertling, L. (2011). A stochastic model for opportunistic maintenance planning of offshore wind farms. In *IEEE PowerTech* (pp. 1–8). Trondheim.
- Bessa, R. J. (2016). On the quality of the Gaussian copula for multi-temporal decision-making problems. In *19th Power systems computation conference, PSCC 2016*.
- Bessa, R., Möhrlein, C., Fundel, V., Siefert, M., Browell, J., Haglund El Gaidi, S., et al. (2017). Towards improved understanding of the applicability of uncertainty forecasts in the electric power industry. *Energies*, 10(9), 1402.
- Browell, J., Dinwoodie, I., & McMillan, D. Forecasting for day-ahead offshore maintenance scheduling under uncertainty. In *European safety and reliability (ESREL) conference* (pp. 1137–1144).
- Catterson, V. M., McMillan, D., Dinwoodie, I., Revie, M., Dowell, J., Quigley, J., et al. (2016). An economic impact metric for evaluating wave height forecasters for offshore wind maintenance access. *Wind Energy*, 19(2), 199–212.
- Crabtree, C. J., Zappalá, D., & Hogg, S. I. (2015). Wind energy: UK experiences and offshore operational challenges. *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, 229(7), 727–746.
- Dalgic, Y., Lazakis, I., Dinwoodie, I., McMillan, D., & Revie, M. (2015). Advanced logistics planning for offshore wind farm operation and maintenance activities. *Ocean Engineering*, 101, 211–226.
- Dawid, R., McMillan, D., & Revie, M. (2018). Decision support tool for offshorewind farm vessel routing under uncertainty. *Energies*, 11(9), 1–17.
- Dowell, J., Zitrou, A., Walls, L., Bedford, T., & Infield, D. (2014). Analysis of wind and wave data to assess maintenance access to offshore wind farms. In *Safety, reliability and risk analysis: beyond the horizon - Proceedings of the European safety and reliability conference, ESREL 2013*.
- Efron, B., & Tibshirani, R. (1994). *An introduction to the bootstrap*. Chapman & Hall.
- Eilers, P. H. C., & Marx, B. D. (1996). Flexible smoothing with b-splines and penalties. *Statistical Science*, 11(2), 89–121.
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874.
- Fazeres-Ferradosa, T., Taveira-Pinto, F., Vanem, E., Reis, M. T., & Neves, L. d. (2018). Asymmetric copula-based distribution models for met-ocean data in offshore wind engineering applications. *Wind Engineering*, 42(4), 304–334.
- Friedman, J. (2001). Greedy function approximation: a gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232.
- Fritsch, F. N., & Carlson, R. E. (1980). Monotone piecewise cubic interpolation. *SIAM Journal on Numerical Analysis*, 17(2), 238–246.
- Gilbert, C., Browell, J., & McMillan, D. (2019a). A data-driven vessel motion model for offshore access forecasting. *OCEANS 2019 - Marseille*, 1–6.
- Gilbert, C., Browell, J., & McMillan, D. (2019b). Visualisation of probabilistic access forecasts for offshore operations. *Journal of Physics: Conference Series*, 1222(1), 012040.
- Gilbert, C., Browell, J., & McMillan, D. (2020). Code for: "Probabilistic access forecasting for improved offshore operations". University of Strathclyde, <https://doi.org/10.15129/d4055a23-f2f7-482d-9cb2-4e15960b17f6>.
- Gintautas, T., & Sørensen, J. D. (2017). Improved methodology of weather window prediction for offshore operations based on probabilities of operation failure. *Journal of Marine Science and Engineering*, 5(2), 20.
- Gneiting, T., Balabdaoui, F., & Raftery, A. E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, 69, 243–268.
- Greenwell, B., Boehmke, B., Cunningham, J., & GBM Developers (2019). *gbm: Generalized boosted regression models: R package version 2.1.5*. <https://cran.r-project.org/package=gbm>.
- Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.
- Hofner, B., Kneib, T., & Hothorn, T. (2016). A unified framework of constrained regression. *Statistics and Computing*, 26(1–2), 1–14.
- Hofner, B. H., Mayr, A., & Schmid, M. (2016). gamboostLSS: An R package for model building and variable selection in the GAMLSS framework. *Journal of Statistical Software*, 74(1), 1–31.
- Hong, T., Pinson, P., Fan, S., Zareipour, H., Troccoli, A., & Hyndman, R. J. (2016). Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond. *International Journal of Forecasting*.
- Irawan, C. A., Ouelhadj, D., Jones, D., Stålhane, M., & Sperstad, I. B. (2017). Optimisation of maintenance routing and scheduling for offshore wind farms. *European Journal of Operational Research*, 256(1), 76–89.
- Kovács, A., Erds, G., Viharos, Z. J., & Monostori, L. (2011). A system for the detailed scheduling of wind farm maintenance. *CIRP Annals - Manufacturing Technology*, 60(1), 497–501.
- Landry, M., Erlinger, T. P., Patschke, D., & Varrichio, C. (2016). Probabilistic gradient boosting machines for {GEFCom2014} wind forecasting. *International Journal of Forecasting*, 32(3), 1061–1066.
- Leontaris, G., Morales-Nápoles, O., & Wolfert, A. R. (2016). Probabilistic scheduling of offshore operations using copula based environmental time series – An application for cable installation management for offshore wind farms. *Ocean Engineering*.
- Maples, B., Saur, G., Hand, M., Van De Pietermen, R., & Obdam, T. (2013). *Installation, operation, and maintenance strategies to reduce the cost of offshore wind energy: Technical Report*, National Renewable Energy Laboratory (NREL).
- Mayr, A., Fenske, N., Hofner, B., Kneib, T., & Schmid, M. (2012). Generalized additive models for location, scale and shape for high dimensional data—a flexible approach based on boosting. *Journal of the Royal Statistical Society. Series C. Applied Statistics*, 61(3), 403–427.
- Morales, J. M., Conejo, A. J., Madsen, H., Pinson, P., & Zugno, M. (2013). *Integrating renewables in electricity markets: Operational problems*, (Vol. 205). Springer Science & Business Media.
- Nelsen, R. B. (2006). *Springer series in statistics, An introduction to copulas*. New York, NY: Springer New York.
- O'Connor, M., Lewis, T., & Dalton, G. (2013). Weather window analysis of Irish west coast wave data with relevance to operations & maintenance of marine renewables. *Renewable Energy*, 52, 57–66.
- Perperoglou, A., Sauerbrei, W., Abrahamowicz, M., & Schmid, M. (2019). A review of spline function procedures in R. *BMC Medical Research Methodology*, 19(1), 46.
- Pinson, P., & Girard, R. (2012). Evaluating the quality of scenarios of short-term wind power generation. *Applied Energy*, 96, 12–20.
- Pinson, P., Madsen, H., Papaefthymiou, G., & Klöckl, B. (2009). From probabilistic forecasts to statistical scenarios of short-term wind power production. *Wind Energy*, 12(1), 51–62.
- Pinson, P., Nielsen, H. A., Møller, J. K., Madsen, H., & Kariniotakis, G. N. (2007). Non-parametric probabilistic forecasts of wind power: required properties and evaluation. *Wind Energy*, 10(6), 497–516.

- R Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, <https://www.r-project.org/>.
- Raknes, N., Ødeskaug, K., Stålhane, M., & Hvattum, L. (2017). Scheduling of maintenance tasks and routing of a joint vessel fleet for multiple offshore wind farms. *Journal of Marine Science and Engineering*, 5(1), 11.
- Reikard, G., Pinson, P., & Bidlot, J. R. (2011). Forecasting ocean wave energy: The ECMWF wave model and time series methods. *Ocean Engineering*, 38(10), 1089–1099.
- Rigby, R. A., Stasinopoulos, D. M., & Lane, P. W. (2005). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society. Series C. Applied Statistics*, 54(3), 507–554.
- Scheuerer, M., & Hamill, T. M. (2015). Variogram-based proper scoring rules for probabilistic forecasts of multivariate quantities*. *Monthly Weather Review*, 143(4), 1321–1334.
- Shafiee, M. (2015). Maintenance logistics organization for offshore wind energy: Current progress and future perspectives. *Renewable Energy*, 77(1), 182–193.
- Silva, L. (2014). A feature engineering approach to wind power forecasting: {GEFCom} 2012. *International Journal of Forecasting*, 30(2), 395–401.
- Stålhane, M., Vefsnmo, H., Halvorsen-Weare, E. E., Hvattum, L. M., & Nonås, L. M. (2016). Vessel fleet optimization for maintenance operations at offshore wind farms under uncertainty. *Energy Procedia*, 94(1876), 357–366.
- Stasinopoulos, D. M., & Rigby, R. A. (2007). Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software*, 23(7), 1–46.
- Stock-Williams, C., & Swamy, S. K. (2018). Automated daily maintenance planning for offshore wind farms. *Renewable Energy*, 1–11.
- Sweeney, C. P., Lynch, P., & Nolan, P. (2013). Reducing errors of wind speed forecasts by an optimal combination of post-processing methods. *Meteorological Applications*, 20(1), 32–40.
- Tastu, J. (2013). *Short-term wind power forecasting: probabilistic and space-time aspects* (Ph.D. thesis), Technical University of Denmark.
- Tastu, J., Pinson, P., & Madsen, H. (2015). Space-time trajectories of wind power generation: Parametrized precision matrices under a Gaussian copula approach. *Lecture Notes in Statistics*, 217, 267–296.
- Taylor, J. W., & Jeon, J. (2018). Probabilistic forecasting of wave height for offshore wind turbine maintenance. *European Journal of Operational Research*, 267(3), 877–890.
- Walker, R. T., Van Nieuwkoop-Mccall, J., Johanning, L., & Parkinson, R. J. (2013). Calculating weather windows: Application to transit, installation and the implications on deployment success. *Ocean Engineering*, 68, 88–101.
- Wang, Z., Wang, W., Liu, C., Wang, Z., & Hou, Y. (2018). Probabilistic forecast for multiple wind farms based on regular vine copulas. *IEEE Transactions on Power Systems*, 33(1), 578–589.
- Will, G. (2016). *Visualizing and clustering data that includes circular variables* Technical report, (Master's thesis) Montana State University, <http://www.math.montana.edu/graduate/writing-projects/2016/16will.pdf>.