Downloaded from orbit.dtu.dk on: Nov 07, 2026

DTU Library

Probabilistic Forecasting of the Wave Energy Flux

Pinson, Pierre; Reikard, G.; Bidlot, J.-R.

Published in:
Applied Energy

Link to article, DOI:
10.1016/j.apenergy.2011.12.040

Publication date:
2012

Document Version
Early version, also known as pre-print

Link back to DTU Orbit

Citation (APA):
Pinson, P., Reikard, G., & Bidlot, J.-R. (2012). Probabilistic Forecasting of the Wave Energy Flux. *Applied Energy*, *93*, 364-370. https://doi.org/10.1016/j.apenergy.2011.12.040

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Probabilistic forecasting of the wave energy flux

P. Pinson[a,*], G. Reikard[b], J.-R. Bidlot[c]

[a] *Technical University of Denmark, Dpt. of Informatics and Mathematical Modelling, Denmark*
[b] *Leap Wireless, Denver, Colorado (USA)*
[c] *European Centre for Medium-range Weather Forecasts, Reading, United Kingdom*

Abstract

Wave energy will certainly have a significant role to play in the deployment of renewable energy generation capacities. As with wind and solar, probabilistic forecasts of wave power over horizons of a few hours to a few days are required for power system operation as well as trading in electricity markets. A methodology for the probabilistic forecasting of the wave energy flux is introduced, based on a log-Normal assumption for the shape of predictive densities. It uses meteorological forecasts (from the European Centre for Medium-range Weather forecasts - ECMWF) and local wave measurements as input. The parameters of the models involved are adaptively and recursively estimated. The methodology is evaluated for 13 locations around North-America over a period of 15 months. The issued probabilistic forecasts substantially outperform the various benchmarks considered, with improvements between 6 and 70% in terms of Continuous Rank Probability Score (CRPS), depending upon the test case and the lead time. It is finally shown that the log-Normal assumption can be seen as acceptable, even though it may be refined in the future.

Keywords: wave energy, forecasting, statistical models, adaptive estimation, forecast skill, probabilistic calibration

1. Introduction

Ocean wave energy is a new, but rapidly expanding field, which offers the promise of generating significant amounts of electricity in coastal areas [1]. While only a small number of wave farms are now in operation, there are plans for significant expansion in several countries. Status and

*Corresponding author
Email address: pp@imm.dtu.dk (P. Pinson)
URL: www2.imm.dtu.dk/~pp/ (P. Pinson)

perspectives for wave energy developments in Europe were reviewed by Clement *et al.* [2] while a more recent global update is given by Esteban & Leary [3]. As with other forms of renewable energy, such as wind and solar power, wave energy can be highly variable and of limited predictability. The problem for Transmission System Operators (TSOs), utilities and trading agents is therefore to forecast ocean wave energy over horizons consistent with the operation of electricity grids and markets. These horizons are typically short, in the area of six hours in the United States or United Kingdom, and somewhat longer, in the order of two days in European electricity markets. Even though we focus here on lead times up to 48 hours, TSOs and utilities will also be interested in forecasting over somewhat longer horizons, in the range of several days, for operational planning, reserve usage and peak load matching. Very-short lead times (between 0 and 30 minutes) could also be considered in the future for the purpose of the control and optimal economic operation of the wave energy farms.

The value of probabilistic forecasts in decision-making related to the optimal integration of renewable energy generation in power systems is increasingly recognized, as for the example case of reserve quantification [4]. This follows from a more general result: for a large class of decision-making problems, optimal decisions directly relate to quantiles of conditional predictive distributions [5]. Consequently, and similar to other renewable energy sources, wave energy forecasts should optimally take the form of predictive densities.

Since very little data exists for real-world projects of operating wave energy converters, emphasis is placed on the wave energy flux, which is the central variable representing the power these wave energy converters may harvest from the sea. The capability to forecast wave energy flux at these horizons currently exists. In prior studies, it was determined that statistical models could predict ocean wave energy more accurately at horizons of 1-6 hours, while for longer horizons, physics-based models were more accurate [6]. However, it was found that combining both methods predicted more accurately than either one individually [7], in line with parallel works that led to similar conclusions for wind [8] and solar [9] energy prediction. The objective of the present study is to explore such a combined approach in a more general setup by producing full probabilistic forecasts of wave energy flux. A parametric approach to probabilistic forecasting based on log-Normal distributions is described, the moments of which are determined through common linear dynamic models using exogenous information (provided by the raw forecasts of the European Center

for Medium-range Weather Forecasts (ECMWF)) and local observations. Model parameters are adaptively and recursively estimated in a Recursive Least Squares (RLS) framework.

The organization of this manuscript is as follows. Section 2 introduces the general probabilistic forecasting methodology using log-Normal predictive densities, as well as the chosen models for their location and scale parameters. The approach to the adaptive and recursive estimation of the model parameters is subsequently described in Section 3. The methodology is applied in Section 4 for the probabilistic forecasting of wave energy flux at 13 locations around North-America, where the corresponding results are discussed. The manuscript ends in Section 5 with conclusions and perspectives regarding further work.

2. Probabilistic forecasting methodology

2.1. Proposal of parametric predictive densities

The wave energy flux is seen as a stochastic process $\{Y_t\}$, the realisations of which define the time-series $\{y_t\}$. t is the time index and for simplicity it is assumed that the observations are sampled at regular intervals. The wave energy flux corresponds to a power content per unit of surface of the crest length. It is usually expressed in kW.m^{-1} and is given by

$$Y_t := \frac{g^2\rho}{64\pi} H_t^2 T_t, \quad Y_t \geq 0 \tag{1}$$

where g is the gravitational acceleration constant and ρ the density of seawater. H_t and T_t denote the significant wave height and mean wave period, respectively. In addition, the multiplicative nature of the relationship between wave height and period means that the uncertainty in forecasts of the wave energy flux is greater than for the height and period individually.

The approach here is probabilistic forecasting in a parametric setup, hence based on assumption for the shape of predictive densities. For any given time t it is aimed at issuing forecasts of the characteristics of the random variable Y_{t+k} conditional to the models and information available up to time t. k is referred to as the lead time.

Based on an empirical analysis of the wave energy flux in logs, it appears that the formulation of a log-Normal assumption about predictive densities would be appropriate. This assumption will be verified through the case-study applications and the discussion of our forecasting results. For a given lead time k it is therefore assumed that

$$Y_{t+k} \sim \ln\mathcal{N}(\hat{\mu}_{t+k|t}, \hat{\sigma}^2_{t+k|t}) \tag{2}$$

where $\hat{\mu}_{t+k|t}$ and $\hat{\sigma}_{t+k|t}$ are forecasts of the location and scale parameters of the log-Normal distribution. The density for the log-Normal distribution reads

$$f(y) = \frac{1}{y\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\ln(y)-\mu)^2}{2\sigma^2}\right\}, \quad y > 0 \tag{3}$$

Assuming a variable Y_t to be distributed log-Normal is then equivalent to considering $X_t = \ln(Y_t)$ to be distributed Gaussian,

$$X_t = \ln(Y_t), \quad X_t \sim \mathcal{N}(\mu_t, \sigma_t^2) \tag{4}$$

The introduction of the various models and estimation approaches introduced in the following will extensively rely on this idea. Note that one could extend this proposal by considering generalized logit-Normal distributions instead, as introduced by Pinson [10] for the case of wind power generation. This extension could be especially useful in the future when accounting for the wave energy converters and their nonlinear effect on the processes involved. Here by using the logarithmic transformation of (4), the resulting stochastic process $\{X_t\}$ can be modelled in a Gaussian framework.

If aiming at extracting point forecasts from the predictive densities that will be produced, the properties of log-Normal distribution tells us that the conditional expectation of wave energy flux is

$$\hat{y}_{t+k|t} = \exp\left\{\hat{\mu}_{t+k|t} + \frac{\hat{\sigma}^2_{t+k|t}}{2}\right\} \tag{5}$$

while more intuitively, one could think that it would simply be given by $\exp(\hat{\mu}_{t+k|t})$. $\exp(\hat{\mu}_{t+k|t})$ actually corresponds to the median of the conditional density for that random variable.

2.2. Dynamic models for the location and scale parameters

There are in principle two ways to combine physical and statistical approaches to forecasting. One may include the output of the physical model (ECMWF here) in a regression model of the flux. Alternatively, one may fit a statistical model to the residuals from the physics model in order to correct for its dynamic deficiencies. Such a post-processing of physical model forecasts is commonly referred to as Model Output Statistics (MOS). Both ideas mentioned above are here merged, since regressing on the physical model output while also correcting for dynamical deficiencies in a time-series analysis framework.

To begin with,

$$\tilde{x}_{t+k|t} := \ln\left(\frac{g^2\rho}{64\pi}\tilde{H}^2_{t+k|t}\tilde{T}_{t+k|t}\right) \tag{6}$$

denotes the log-transformed value of the wave energy flux predicted by the ECMWF model at time t for time $t+k$. It should be understood that for both location and scale parameters separate models are built for every individual lead time k. Similarly x_t is the log-transformed value of the measured wave energy flux at time t. The log-transformation is there to allow for a Gaussian assumption for the residuals of the statistical models to be proposed, which would not be acceptable for the non-transformed variable. Even though modeling is to be based on the values of $\tilde{x}$, the final forecasts are issued in the space of the original quantity, then respecting its original dimension.

For lead times between 1 and 48 hours, a relevant model for the location parameter $\hat{\mu}_{t+k|t}$ in its vectorial form is

$$\hat{\mu}_{t+k|t} := \boldsymbol{\theta}_t^\top \mathbf{z}_t = \theta_0 + \theta_1 \tilde{x}_{t+k|t} + \theta_2 x_t + \ldots + \theta_{l+2} x_{t-l} \tag{7}$$

with

$$\boldsymbol{\theta}_t := [\theta_0,\ \theta_1,\ \theta_2, \ldots,\ \theta_{l+2}]^\top \tag{8}$$

$$\mathbf{z}_t := \left[1,\ \tilde{x}_{t+k|t},\ x_t, \ldots,\ x_{t-l}\right]^\top \tag{9}$$

where l is then the order of the autoregressive part of the model, which should be selected depending upon the lead time of interest. The autoregressive order is typically larger for short lead times (say, up to 6-9 hours) and then null for further lead times. The maximum lag l can be determined after examining the autocorrelation and partial autocorrelation functions of the $\{x_t\}$ time-series, by minimisation of either Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC), or finally through a cross-validation exercise. In the above, θ_0 has the same dimension as the x and $\tilde{x}$ variables involved in this regression. This is while the θ_is ($i > 0$) are dimensionless.

In (7) the model parameters are time varying, i.e., they evolve slowly over time. This is also the case with wind model dynamics, as proposed by Gneiting *et al.* [11], Hering and Genton [12] and Pinson [10] among others. In essence, the model includes a physics component based on the ECMWF model forecasts ($\theta_1 \tilde{x}_{t+k|t}$), and an autoregressive component ($\theta_0 + \theta_2 x_t + \ldots + \theta_{l+2} x_{t-l}$) that allows recalibration based on the recent history of the wave energy flux time-series.

In parallel, since it was not found that the scale parameter should be made an explicit function of anything else than time (and lead time, obviously), it then takes the following form,

$$\hat{\sigma}_{t+k|t} := \sqrt{\beta}_t \tag{10}$$

i.e. the scale parameter also evolves over time, with β_t the actual model parameter which will be tracked adaptively. A different model parameter β_t is used for each individual lead time k.

3. Adaptive and recursive estimation of model parameters

The parameters $\boldsymbol{\theta}_t$ and β_t are estimated within a Recursive Least Squares (RLS) method [13, 14]. Recursive estimation offers the advantage that only the last available measurements are used at each time step for updating the model parameters. This contrasts with the more computationally-intensive idea of employing a sliding window, as done by Gneiting *et al.* [11] and Hering and Genton [12], for which estimation is performed on a bulk of data at each time step.

Focusing on the model related to lead time k, the estimate of the parameters $\boldsymbol{\theta}_t$ at time t is defined as

$$\hat{\boldsymbol{\theta}}_t := \arg\min_{\boldsymbol{\theta}} \sum_{i=1}^{t} \lambda^{t-i} \left(x_i - \boldsymbol{\theta}^\top \tilde{\mathbf{z}}_{i-k} \right)^2 \tag{11}$$

where λ is the forgetting factor ($\lambda \in (0,1)$), allowing for the exponential forgetting of past observations, for which the corresponding effective number of observations is defined as $n_\lambda = (1-\lambda)^{-1}$. The value of the forgetting factor is typically slightly below 1.

In a recursive framework the formulae for updating the LS-estimate of $\boldsymbol{\theta}_t$ based on the newly available information at time t summarises to

$$\hat{\mathbf{R}}_t = \lambda \hat{\mathbf{R}}_{t-1} + \tilde{\mathbf{z}}_{t-k} \tilde{\mathbf{z}}_{t-k}^\top \tag{12}$$

$$\hat{\boldsymbol{\theta}}_t = \hat{\boldsymbol{\theta}}_{t-1} + \hat{\mathbf{R}}_t^{-1} \tilde{\mathbf{z}}_{t-k} \varepsilon_t \tag{13}$$

where ε_t is defined as $\varepsilon_t := x_t - \hat{\boldsymbol{\theta}}_{t-1}^\top \tilde{\mathbf{z}}_{t-k}$. R_t is covariance matrix for the variables entering the regression model of Eq. (7) as estimated at time t. It summarizes the variances and covariances of these variables, similar to the case of the LS estimates in a regression model. The derivations necessary for obtaining this updating scheme can be found in Madsen [14] among others.

Since the scale parameter is given by a constant only, its adaptive estimation at time t and for a given led time k can be performed thanks to an exponential smoothing scheme, being consistent with the update of the dynamic model for the location parameter. This writes

$$\hat{\beta}_t = \lambda_t \hat{\beta}_{t-1} + (1-\lambda_t) {\varepsilon_t}^2 \tag{14}$$

The coefficients $\hat{\boldsymbol{\theta}}_t$ and $\hat{\beta}_t$ are initialized as

$$\hat{\boldsymbol{\theta}}_0 := [0\ 1\ 0\ \ldots\ 0]^\top \tag{15}$$

$$\hat{\beta}_0 := \epsilon \tag{16}$$

with ϵ a small value. Consequently at $t = 0$ the output location parameter correspond to the original ECMWF forecast. $\hat{\beta}_0$ could alternatively be set to an expert guess. The initial inverse covariance matrix $\hat{\mathbf{R}}_0$ takes the form of an identity matrix times a small constant, yielding a small loading on its diagonal. Only the updating formula (12) is used in the recursive estimation procedure as long as $\hat{\mathbf{R}}_t$ is not invertible. When that point is reached, both (13) and (14) are also employed at each time step for updating the models for the location and scale parameters.

4. Application and results

4.1. Data characteristics

The original measurement dataset consists of hourly values for the significant wave height and wave period. These were obtained from 11 buoys from the U.S. National Data Buoy Centre (NDBC) and two from the Canadian Integrated Science Data Management (ISDM). The NDBC wave spectral data were downloaded from the National Oceanographic Data Centre website (www.nodc.noaa.gov) whereas the ISDM data were downloaded from the web portal (http://www.meds-sdmm.dfo-mpo.gc.ca/isdm-gdsi/waves-vagues/index-eng.htm). Following a basic quality control procedure applied to the observed spectra [18], hourly time series for significant wave height ($\{H_t\}$) and the mean energy wave period ($\{T_t\}$) were derived. This procedure consists in discarding values that are *(i)* out of the acceptable physical range, *(ii)* remaining constant for 1 day or more, and *(iii)* seen as outliers owing to their substantial deviation from the mean behaviour, in terms of process magnitude values or in terms of the amplitude of the variations from one hour to the next.The wave energy flux is computed from these values using (1). Measurements were collected for a period ranging from the 1$^{\text{st}}$ December 2007 until the 31$^{\text{st}}$ March 2010. For 2 of the buoys (41048 and 51001) the datasets were not complete, with measurements missing for a few months at the beginning or at the end of the period. A complete overview of the buoys locations and collected datasets are given in Figure 1 and Table 1.

The physics model-based forecasts are those from the Wave Model (WAM) of ECMWF. The basic principle is to model waves as a function of wind inputs, wave-wave interactions, and dissipa-

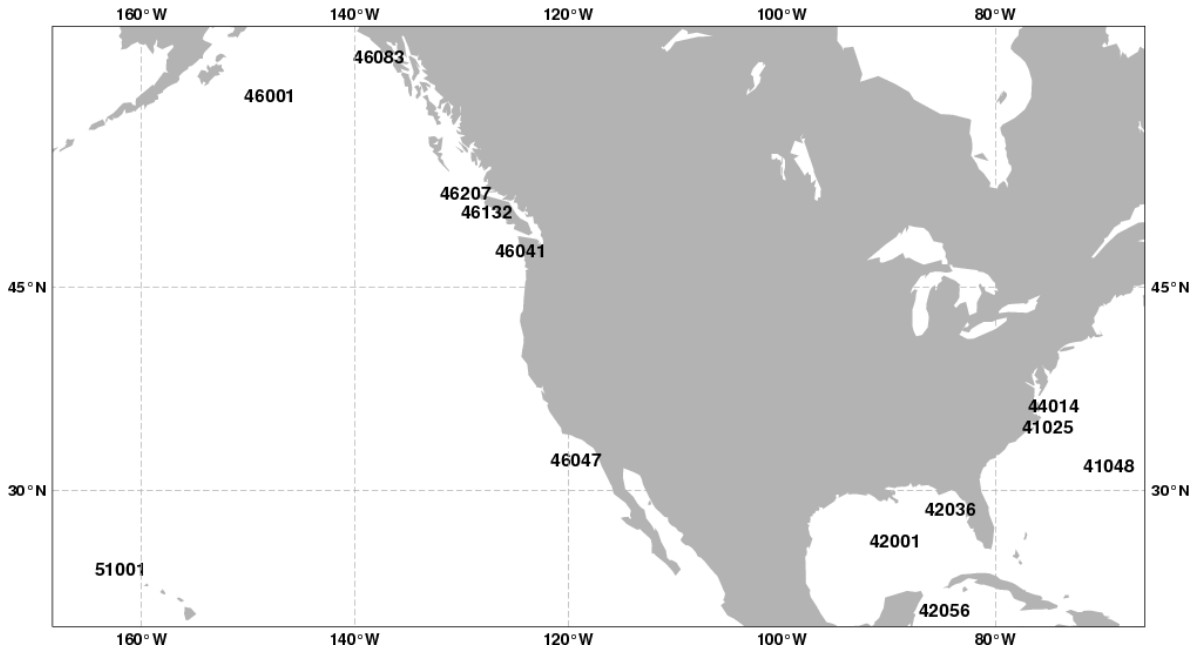

Figure 1: Map with the locations of the 13 buoys considered in the empirical work.

Table 1: Characteristics of the buoys and of the corresponding datasets.

Id.	Dates	Depth [m]	Location	(Latitude, Longitude)	Operator
41025	1.12.2007-31.3.2010	68	Diamond Shoals, NC	(35.001 N, 75.40 W)	NDBC
41048	1.12.2007-31.12.2009	5261	West Bermuda	(31.98 N, 69.65 W)	NDBC
42001	1.12.2007-31.03.2010	3246	Mid-Gulf, Louisiana	(25.90 N, 89.67 W)	NDBC
42036	1.12.2007-31.03.2010	54	West Tampa, FL	(28.50 N, 85.52 W)	NDBC
42056	1.12.2007-31.03.2010	4446	Yucatan Basin	(19.874 N, 85.06 W)	NDBC
44014	1.12.2007-31.03.2010	48	Virginia Beach, VA	(36.61 N, 74.84 W)	NDBC
46001	1.12.2007-31.03.2010	4206	Gulf of Alaska, AK	(56.30 N, 148.02 W)	NDBC
46041	1.12.2007-31.03.2010	132	Cape Elizabeth, WA	(47.35 N, 124.73 W)	NDBC
46047	1.12.2007-31.03.2010	1393	Tanner Banks, CA	(34.43 N, 119.533 W)	NDBC
46083	1.12.2007-31.03.2010	136	Yakutak, AK	(52.24 N, 137.99 W)	NDBC
46132	1.12.2007-31.03.2010	2040	South Brooks, BC	(49.74 N, 127.93 W)	ISDM
46207	1.12.2007-31.03.2010	2125	East Dellwood, BC	(50.88 N, 129.92 W)	ISDM
51001	21.02.2008-24.12.2009	3430	Kauai, HI	(23.45 N, 162.28 W)	NDBC

tion. The WAM is a third-generation system which has been used primarily to model deep water domains, although it can also be applied in shallow water domains [15, 16]. The ECMWF WAM model is further described in Janssen [17] and Bidlot *et al.* [18].

The model forecasts are derived from the global high resolution version of WAM. It had a latitude-longitude grid with an average resolution of 40 km until the end of January 2010. Since then, the resolution has been increased to 28 km. Because the wave model is actively coupled to

the ECMWF global atmospheric model, it provides feedback to the atmosphere in the form of the change to the surface roughness, which in turn influences the varying momentum and energy flux from the atmosphere to waves. Shallow water effects are also included. ECMWF WAM forecasts are issued twice daily (0 and 12 UTC starting time) with outputs having a 3-hour resolution. A maximum horizon of 48 hours is considered here. The model forecasts include the significant wave height and period. The forecasts for the flux are calculated from these values using (1), being consistent with treatment of measurements. Forecasts are available for the whole period over which the measurements were collected. The temporal resolution of the measurements is matched with that of the forecasts, i.e. 3-hourly.

4.2. Model set-up and evaluation framework

From the available data, two periods are defined, the first one being used for identification (and initial training) of the statistical models, and the second for evaluating model performance. The learning set covers the period from the 1st December 2007 until the 31st December 2008. The remainder of the dataset, that is, from the 1st January 2009 until the 31st March 2010, is used for out-of-sample evaluation of forecast performance.

Over the 13-month learning period, a part of the data is used for one-fold cross validation (the last 3 months) in order to select an optimal autoregressive order for the dynamic models, as well as the optimal value for the forgetting factor. The selection of optimal structures and forgetting factor is done in a trial-and-error manner, by evaluating the results obtained over the validation set with different model setups. The underlying idea is that while minimizing forecast evaluation scores on the training data only could lead to over-fitting, the evaluation of forecast performance on an independent cross-validation dataset permits to inform of how the model would perform when facing new data (the so-called generalization ability).

The maximum autoregressive lag l is expected at maximum 3, while the range of envisaged value for the optimal forgetting factor is $[0.98, 0.999]$. For more information on cross validation, the reader is referred to Stone [19]. The criterion to be minimised over the cross-validation set is the negative logarithmic score, also known as Ignorance after the work of Roulston and Smith [20]. For a given predictive distribution $\hat{f}_{t+k|t}(y)$ and corresponding wave energy flux measurement y_{t+k}, the Ignorance score value $\mathrm{Is}_{t,k}$ is given by

$$\mathrm{Is}_{t,k} := -\ln(\hat{f}_{t+k|t}(y_{t+k})) \tag{17}$$

so that Ignorance can be averaged over the validation and evaluation sets, calculated separately for each horizon,

$$\mathrm{Is}_k \;:=\; -\frac{1}{N}\sum_{t=1}^{N}\ln(\hat{f}_{t+k|t}(y_{t+k})) \tag{18}$$

where N stands for the number of forecast series. Only one forgetting factor value is selected for each buoy, being the same for all lead times. Further refinement may consider optimizing the forgetting factor for each lead time individually.

The intercept term θ_0 (see (8)) was found necessary for all buoys and all lead times as well as the linear regression on the ECMWF WAM model output (related to the coefficient θ_1). Regarding the autoregressive part of model (7), it was found that in a generic manner for all buoys the maximum lag l should be 2 for 3-hour ahead forecasts, 1 for 6-hour ahead forecasts, and 0 for further lead times. Note that this may not be the case for other buoys depending upon the local characteristics. Optimal model structures should always be decided upon based on a thorough analysis of available data and forecast experiments. The optimal forgetting factor values obtained through this cross-validation exercise are collated in Table 2. For all buoys the changes in dynamics are fairly slow, with equivalent training window sizes (defined as $1/(1-\lambda)$) between 100 and 1000 time steps. No relationship can be found between the value of the optimal forgetting factor and the location and depth of the buoys considered.

Table 2: Optimal forgetting factor values for the 13 buoys as selected based on minimizing the average Ignorance score over the validation set.

Id.	41025	41048	42001	42036	42056	44014
λ	0.999	0.999	0.999	0.993	0.999	0.996
46001	46041	46047	46083	46132	46207	51001
0.991	0.999	0.999	0.991	0.999	0.998	0.996

4.3. Evaluating the skill of the forecasts

Before evaluating the performance of the probabilistic forecasting methodology, let us first illustrate the outcome of that methodology based on the example forecasts of Figure 2 issued for West Tampa (buoy id. 42036) on the 1$^{\mathrm{st}}$ January 2010 at 0 UTC. Predictive distributions are represented by a set of prediction intervals with increasing nominal coverage rates (from 10 to 90%) in a fashion similar to Pinson and Kariniotakis [21]. The point forecasts that may be extracted

from these predictive densities (following (5)) are also represented, along with corresponding observations. This plot nicely represents how the formulated log-Normal assumption, combined with the dynamic models for every lead time, makes the shape and width of predictive densities evolve with the lead time and the level of wave energy flux.

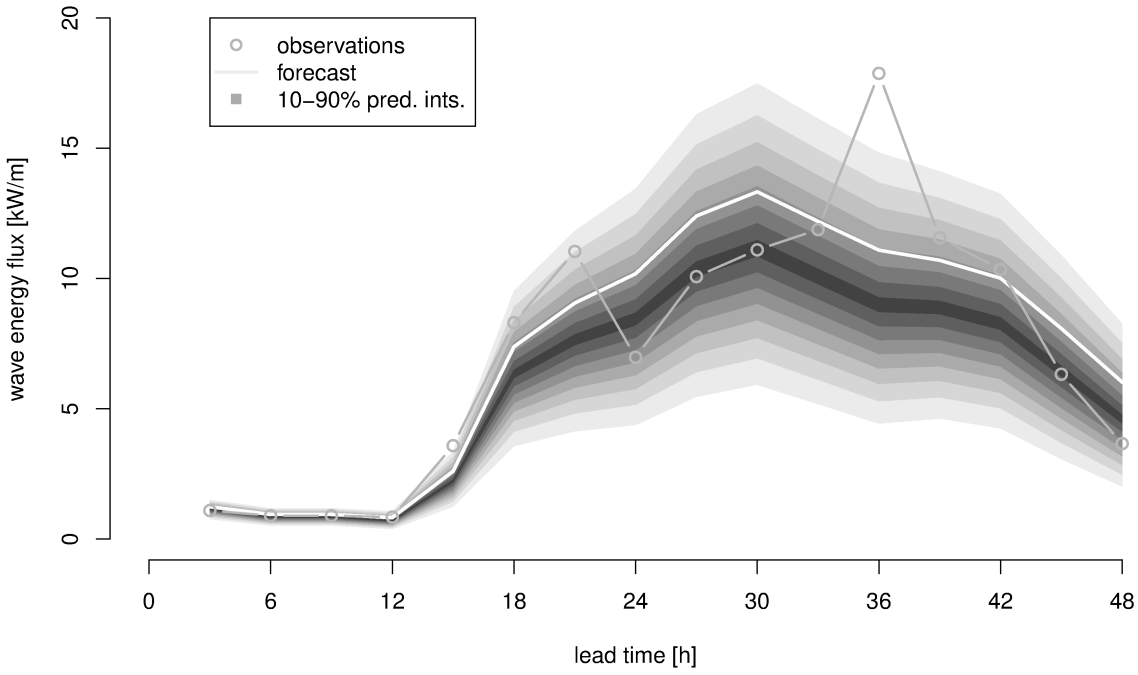

Figure 2: Probabilistic forecast of the wave energy flux issued for West Tampa (buoy id. 42036) on the 1st January 2010 at 00:00 UTC. The point forecast (i.e. the conditional expectation) is also represented.

Probabilistic forecasts of the wave energy flux are assessed in terms of their overall skill. Such a skill is quantified with proper skill scores like Ignorance used in the above. Another proper skill score which is widely used in the literature is the Continuous Rank Probability Score (CRPS) [22]. The value of the CRPS for a given forecast density $\hat{f}_{t+k|t}(y)$, with related cumulative distribution function $\hat{F}_{t+k|t}(y)$, and corresponding measurement y_{t+k}, is calculated as

$$\text{CRPS}_{t,k} := \int_y \left(\hat{F}_{t+k|t}(y) - H(y - y_{t+k})\right)^2 dy \tag{19}$$

where $H(.)$ is the Heaviside step function, taking the value 1 for $y \geq y_{t+k}$ and 0 otherwise. The CRPS for a given lead time over the whole evaluation set is calculated as the average over t of

$\text{CRPS}_{t,k}$ values as in (18). As for the case of the Mean Absolute Error (MAE) criterion used for assessing single-valued forecasts, both Ignorance and CRPS are negatively-rewarding score: the lower the better. The Ignorance score has no unit while the CRPS has the same unit as the wave energy flux (kW.m^{-1}).

Over the evaluation set, the skill of the forecasts is evaluated against that of two common benchmarks, climatology and persistence. Climatology relies on all available past measurements gathered to form the so-called climatological distribution G_c. This distribution is then used as an unconditional predictive density, i.e. at any time t and lead time k by assuming that

$$Y_{t+k} \sim G_c \tag{20}$$

Since it is generally easy to outperform for short lead times (say, less than a day), another relevant benchmark is the probabilistic generalization of persistence. Persistence is based on a random walk model. In the present case, this translates to defining predictive densities based on the assumption such that

$$Y_{t+k} \sim \ln\mathcal{N}\left(x_t, \hat{\sigma}_t^{p2}\right) \tag{21}$$

This means that the location parameter of log-Normal predictive densities is directly given by the last (log-transformed) measurement. The scale parameter $\hat{\sigma}_t^p$ may be obtained through an exponential smoothing scheme similar to that of (14). This involves defining ε_t as $\varepsilon_t = (x_t - x_{t-1})$ and using a fairly low forgetting factor instead to accommodate rapid changes in the dynamics of ε_t. It is chosen as 0.9 here.

As an illustrative example, Figure 3 depicts the skill evaluation results at Diamond Shoals (buoy id. 41025), which is the site where the skill of the probabilistic forecasts is the poorest. Whatever the skill score considered, persistence is the most serious benchmark for short lead times, while climatology takes over for further lead times. The skill of persistence appears to degrade more rapidly when using the Ignorance score. This is owing to the higher sensitivity of this score which harshly penalizes large forecast errors. With both scores, the proposed probabilistic approach significantly outperforms both benchmark, even though the improvements over persistence are fairly small for the very short-term (say, between 0 and 6 hours ahead).

In addition, a skill score based on the CRPS is introduced: it expresses the improvement over

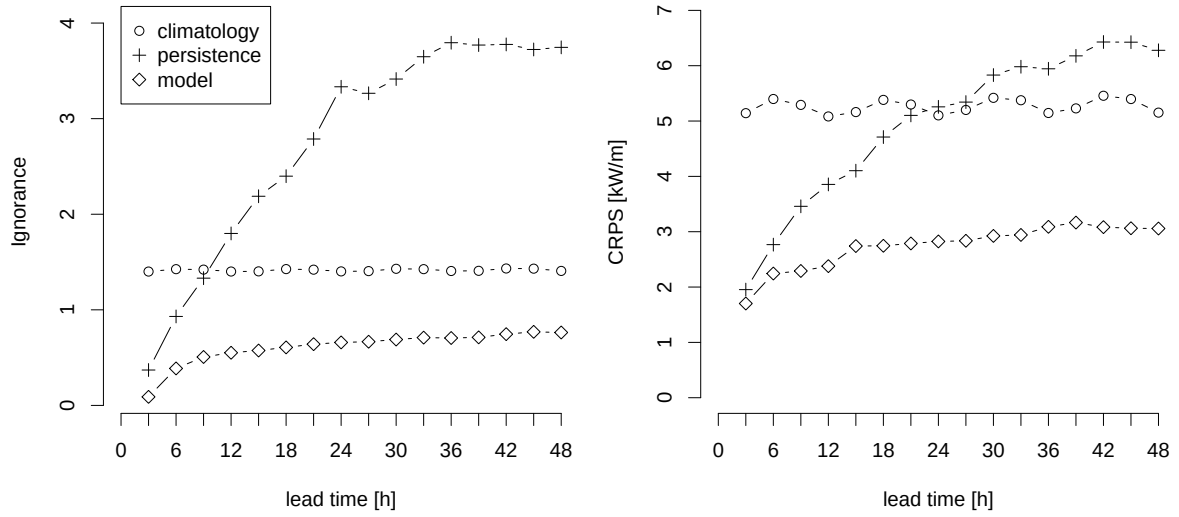

Figure 3: Probabilistic forecast skill as a function of the lead time at Diamond Shoals, as evaluated with both the CRPS and the Ignorance score. The two benchmarks considered are persistence and climatology.

the best of the two benchmarks considered. This skill score for a given lead time k is defined as

$$\mathrm{SCRPS}_k := 100\ \frac{\min(\mathrm{CRPS}_k^p, \mathrm{CRPS}_k^c) - \mathrm{CRPS}_k}{\min(\mathrm{CRPS}_k^p, \mathrm{CRPS}_k^c)} \tag{22}$$

where CRPS^c and CRPS^p denote the CRPS value for the climatology and persistence benchmarks, respectively. SCRPS is expressed in percent with positive values indicating improvements over the benchmarks. Note that such type of skill score cannot be readily defined based on the Ignorance score.

Figure 4 gathers the results for all 13 locations in terms of the evolution of the SCRPS with the lead time. All the skill score values are positive, indicating that the probabilistic forecasting approach is always better than the best of the benchmarks. SCRPS values increase with the lead time and reach a maximum, after 24 hours, between 40 and 70% depending on the buoy location. The lowest skill overall is at Diamond Shoals (buoy id. 41025) while the best results are obtained in the Gulf of Alaska (buoy id. 46001).

4.4. Reliability of the probabilistic forecasts

A crucial aspect of probabilistic forecasts is reliability, i.e., the correspondence of predicted and observed probabilities. For continuous variables like the wave energy flux, reliability can be visually assessed with Probability Integral Transform (PIT) diagrams. For an extensive description of how

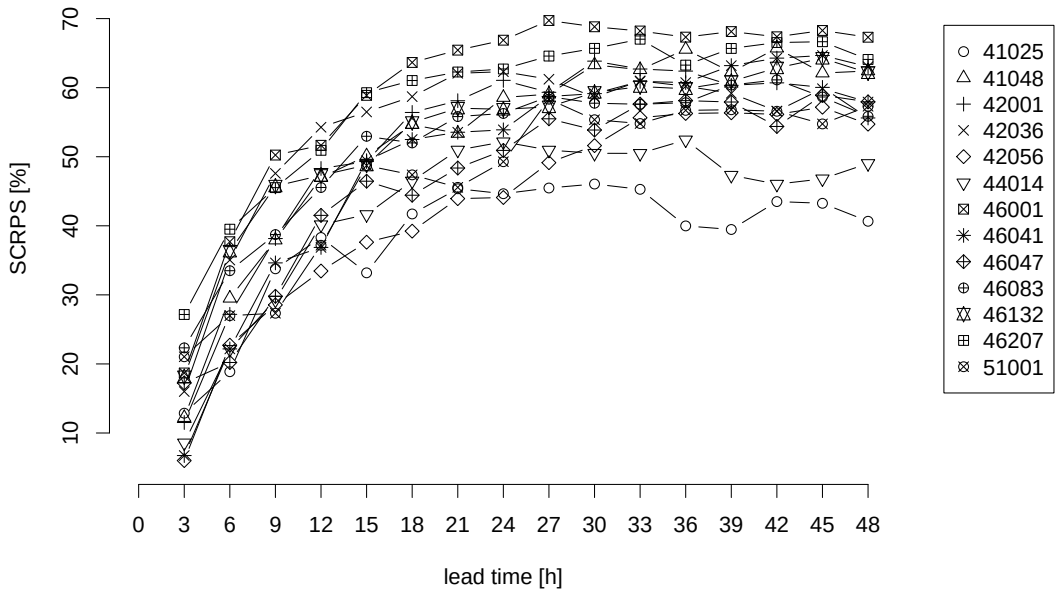

Figure 4: Comparison of the skill of the probabilistic forecast approach with respect to the climatology and persistence benchmarks based on the SCRPS as a function of the lead time.

these diagrams are built, see Pinson *et al.* [23]. To summarize, a number of quantiles forecasts $\hat{q}_{t+k|t}^{(\alpha_j)}$ with various nominal proportions (denoted α_j) are extracted from the predictive densities. By definition, a quantile forecast $\hat{q}_{t+k|t}^{(\alpha_j)}$ with nominal proportion α_j is defined by

$$\mathrm{P}_{\hat{F}_{t+k|t}}\left[y_{t+k} < \hat{q}_{t+k|t}^{(\alpha_j)}\right] = \alpha_j, \quad \alpha_j \in [0,1] \tag{23}$$

This means that having a frequentist approach, a quantile forecast with nominal proportion α_j should cover the observation $\alpha_j\%$ of the times. PIT diagrams then allow for the direct comparison of the nominal and observed proportions. For the calculation of these observed proportions, let us introduce the indicator variable $\xi_{t,k}^{(\alpha_j)}$. Given a quantile forecast $\hat{q}_{t+k|t}^{(\alpha_j)}$ issued at time t for lead time $t+k$, and the verification y_{t+k}, $\xi_{t,k}^{(\alpha_j)}$ is given by

$$\xi_{t,k}^{(\alpha_j)} := \mathbf{1}\{y_{t+k} < \hat{q}_{t+k|t}^{(\alpha_j)}\} = \begin{cases} 1, & \text{if} \quad y_{t+k} < \hat{q}_{t+k|t}^{(\alpha_j)} \\ 0, & \text{otherwise} \end{cases} \tag{24}$$

Finally the observed proportion $\hat{\alpha}_{k,j}$ corresponding to the nominal one α_j for forecasts with lead time k is obtained as

$$\hat{\alpha}_{k,j} := \frac{1}{N} \sum_{t=1}^{N} \xi_{t,k}^{(\alpha_j)} \tag{25}$$

where N is the number of time indices in the evaluation set. PIT diagrams plot $\hat{\alpha}_{k,j}$ against α_j for a chosen lead time k and with α_j varying between 0 and 1. Here the nominal proportions α_j are chosen in $\{0.05, 0.1, \ldots, 0.9, 0.95\}$.

Results from the evaluation of the reliability of the probabilistic forecasts are gathered in Figure 5 for two example lead times, 12 hours ahead (top row) and 48 hours ahead (bottom row). In the ideal case of perfectly calibrated forecasts and of an evaluation set of infinite size, the curves in these plots should lie along the diagonal. For evaluation sets of limited sizes, however, deviation from the diagonal are to be expected owing to sampling (and potentially correlation) effects [21, 24]. The magnitude of potential deviations are illustrated by 95% intervals determined from making a Binomial assumption on the actual observation falling above or below each of the quantiles considered (that is, for $\alpha_j = 0.05, 0.1$, etc.). It was verified that no correlation effect should be accounted for.

The evaluation results are split into two groups: the buoys for which probabilistic forecasts may be considered as having acceptable reliability (left column), and those for which their reliability is not acceptable (right column). These groups were formed after visual inspection of the curves for the observed proportions of the quantiles of the predictive densities lying or not within the range of 95%-intervals around the diagonal. They were noticed to be the same for most lead times, with the probabilistic forecasts at buoys 42001 (Mid-Gulf, Louisiana), 46047 (Tanner Banks, CA), 46132 (South Brooks, BC) and 46207 (East Dellwood, BC) being consistently less reliable than for the other locations. At all these sites, predictive densities tend to significantly overestimate the observed wave energy flux. Interestingly, these sites do not appear to have a common characteristics e.g. water depth, location or distance to shore, which would account for this finding. It shows however that the log-Normal assumption used in the present work may be further evaluated and refined in the future so as to improve its general applicability to probabilistic forecasting of the wave energy flux.

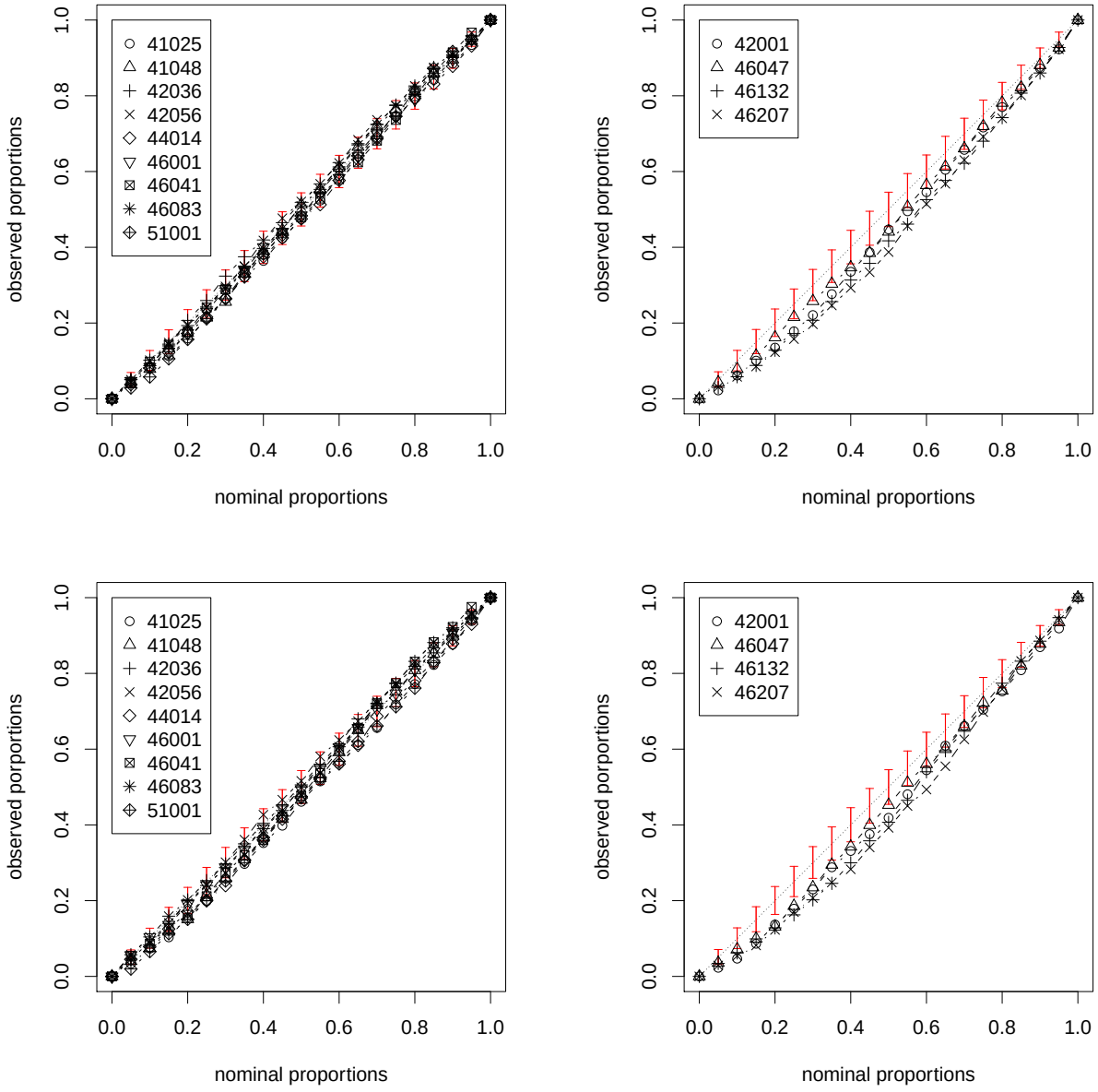

Figure 5: PIT diagrams for the evaluation of the reliability of probabilistic forecasts. These results are for lead times of 12 hours ahead (top) and 48 hours ahead (bottom). The buoys are split into two groups: those where forecasts may be seen as reliable (left), and those where the forecasts can be deemed as non-reliable (right). The bars indicate the potential range of observed proportions for perfectly reliable forecasts.

5. Conclusions and perspectives

The possibility of issuing fully probabilistic forecasts of the wave energy flux by making optimal use of meteorological predictions (from ECMWF) and local wave measurements was investigated.

The forecasts obtained were shown to substantially outperform the usual benchmarks i.e. climatology and persistence. Improvements obtained in terms of CRPS values are between 6 and 70% depending upon the test case and the lead time.

It was proposed to base the forecasting methodology on a log-Normal assumption for the shape of predictive densities. An obvious advantage of such an assumption is that it allows working in a classic Gaussian framework after transformation of the original wave energy flux data. This assumption appeared to be appropriate for most of the sites considered in the case-study application, in view of the acceptable reliability of the probabilistic forecasts issued. It may be refined in the future since the calibration of the forecasts was observed to significantly differ from perfect reliability at a number of sites. Other parametric assumptions e.g. log-logistic or generalized logit-Normal, may be considered. Alternatively, non-parametric approaches such as those employed in wind and solar power forecasting [9, 21] could be envisaged.

Future developments will focus on the possibility of using ensemble forecasts of wave parameters as an input to this type of probabilistic forecasting methodologies. Also, more advanced models may be considered for the post-processing of the raw information provided by meteorological model outputs. Finally, it will be of particular interest to look at the joint forecasting of wind-wave energy output since it is most likely that a number of offshore wind parks and wave energy farms will be jointly operated in the future.

Acknowledgements

The first author of this work was partly supported by the Danish Council for Strategic Research, Technology and Production through the Ensymora project (10-093904/DSF), which is hereby acknowledged. The authors are grateful to the providers of the data used (NDBC, ISDM, ECMWF). The Editor and anonymous reviewers are finally acknowledged for their comments and suggestions on an earlier version of this manuscript.

References

[1] Letcher TM. Future energy: Improved, sustainable and clean options for our planet. Amsterdam: Elsevier, 2008.
[2] Clement A, McCullen P, Falcao A, Fiorentino A, Gardner F, Hammarlund K, Lemonis G, Lewis T, Nielsen K, Petroncini S, Pontes MT, Schild P, Sjostrom BO, Sorensen HC, Thorpe T. Wave energy in Europe: current status and perspectives. Renew Sust Energy Rev 2002;6:405–1.
[3] Esteban M, Leary D. Current developments and future prospects of offshore wind and ocean energy. Appl Energ 2011 (available online).

[4] Matos MA, Bessa RJ. Setting the operating reserve using probabilistic wind power forecasts. IEEE Trans Power Syst 2011;26:594–603.
[5] Gneiting T. Quantiles as optimal point predictors. Int J Forecasting 2011;27:197–207.
[6] Reikard G, Rogers WE. Forecasting ocean waves: Comparing a physics-based model with statistical models. Coast Eng 2011;58:409–6.
[7] Reikard G, Pinson P, Bidlot J. Forecasting ocean wave energy: A comparison of the ECMWF wave model with time series methods. Ocean Eng 2011;38:1089–99.
[8] Sanchez I. Short-term prediction of wind energy production. Int J Forecasting 2006;22:43–56.
[9] Bacher P, Madsen H, Nielsen HAa. Online short-term solar power forecasting. Sol Energy 2009;83:1772–83.
[10] Pinson P. Very short-term probabilistic forecasting of wind power with generalized logit-Normal distributions. J Royal Stat Soc C 2012, available online.
[11] Gneiting T, Larson K, Westrick K, Genton MG, Aldrich E. Calibrated probabilistic forecasting at the stateline wind energy center - The regime-switching space-time method. J Am Stat Assoc 2006;101:968–79.
[12] Hering AS, Genton, MG. Powering up with space-time wind forecasting. J Am Stat Assoc 2009;105:96–104.
[13] Ljung L, Söderström T. Theory and practice of recursive estimation. Boston: MIT Press, 1983.
[14] Madsen H. Time series analysis. London: Chapman & Hall/CRC, 2007.
[15] Janssen PAEM. Quasi-linear theory of wind-wave generation applied to wave forecasting. J Phys Oceanogr 1991;21:1631–42.
[16] Komen GJ, Cavaleri L, Donelan M, Hasselmann K, Hasselmann S, Janssen PAEM. Dynamics and modelling of ocean waves. Cambridge: Cambridge University, 1994. Press: Cambridge.
[17] Janssen PAEM. Progress in ocean wave forecasting. J Comput Phys 2007;227:3572-94.
[18] Bidlot J-R, Li JG, Wittmann P, Faucher M, Chen H, Lefevre J-M, Bruns T, Greenslade D, Ardhuin F, Kohno N, Park S, Gomez M. Inter-comparison of operational wave forecasting systems. Proc. 10th International Workshop on Wave Hindcasting and Forecasting and Coastal Hazard Symposium, North Shore, Oahu, Hawaii, 2007.
[19] Stone M. Cross-validation and assessment of statistical predictions (with discussion). J Royal Stat Soc B 1974;36:111–47.
[20] Roulston MS, Smith LA. Evaluating probabilistic forecasts using information theory. Mon Weather Rev 2002;130:1653–60.
[21] Pinson P, Kariniotakis G. Conditional prediction intervals of wind power generation. IEEE Trans Pow Syst 2010;25:1845–56.
[22] Gneiting T, Balabdaoui F, Raftery AE. Probabilistic forecasts, calibration and sharpness. J Royal Stat Soc B 2007;69:243–68.
[23] Pinson P, McSharry P, Madsen H. Reliability diagrams for non-parametric density forecasts of continuous variables: Accounting for serial correlation. Q J Royal Meteor Soc 2010;136:77–90.
[24] Bröcker J, Smith LA. Increasing the reliability of reliability diagrams. Weather Forecast 2007;22:651–661.