



Forecasting ocean waves: Comparing a physics-based model with statistical models

Gordon Reikard ^{a,*}, W. Erick Rogers ^b

^a Statistics Department, Leap Wireless, CO, USA

^b Oceanography Division, Naval Research Laboratory, Stennis Space Center, MS, USA

ARTICLE INFO

Article history:

Received 18 June 2010

Received in revised form 8 November 2010

Accepted 20 December 2010

Available online 14 January 2011

Keywords:

Wave forecasting

Simulation models

Time series models

ABSTRACT

The literature on ocean wave forecasting falls into two categories, physics-based models and statistical methods. Since these two approaches have evolved independently, it is of interest to determine which approach can predict more accurately, and over what time horizons. This paper runs a comparative analysis of a well-known physics-based model for simulating waves near shore, SWAN, and two statistical techniques, time-varying parameter regression and a frequency domain algorithm. Forecasts are run for the significant wave height, over horizons ranging from the current period (i.e., the analysis time) to 15 h. Seven data sets, four from the Pacific Ocean and three from the Gulf of Mexico, are used to evaluate the forecasts. The statistical models do extremely well at short horizons, producing more accurate forecasts in the 1–5 hour range. The SWAN model is superior at longer horizons. The crossover point, at which the forecast error from the two methods converges, is in the area of 6 h. Based on these results, the choice of statistical versus physics-based models will depend on the uses to which the forecasts will be put. Utilities operating wave farms, which need to forecast at very short horizons, may prefer statistical techniques. Navies or shipping companies interested in oceanic conditions over longer horizons will prefer physics-based models.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

The literature on ocean wave forecasting falls into two broad categories, physics-based models and statistical techniques. In the physics literature, large-scale energy balance models have been the methodology of choice since the 1970s. This last major advance in this technology was the introduction of the third-generation (3G) wave model (Komen et al., 1984). These 3G models are based on theoretical and experimental studies by a number of researchers including Phillips (1957, 1958), Miles (1957), Snyder et al. (1981), Hasselmann et al. (1985), and Janssen (1991). The best-known is the WAM (Wave Model), which has historically been used primarily for large-scale, deep-water applications (WAMDIG, 1988; Komen et al., 1994). The SWAN (Simulating Waves Near shore) model is a more recently developed 3G model, created for the purpose of achieving economical solutions on high resolution grids (Booij et al., 1999). The model includes mechanisms relevant to near-shore processes (depth-induced breaking, bottom friction, parameterized triad interactions) in addition to the traditional mechanisms associated with global/regional wind wave models. SWAN has been used successfully in hindcast applications (Ris et al., 1999), and in real time wave prediction systems for coastal areas (Rogers et al., 2007), Dykes et al. (2009). For recent reviews of 3G

wave modeling, see Jensen et al. (2002), Tolman et al. (2002) and WISE (2007).

The statistical literature is more recent. The most popular technique has been neural networks (Deo and Naidu, 1998; Deo et al., 2001; Tsai et al., 2002; Deo and Jagdale, 2003; Makarynsky, 2004; Londhe and Panchang, 2006; Jain and Deo, 2007; Tseng et al., 2007, and Zamani et al., 2008). Ozger (2010) investigates wavelet transformations. Regression-based models have been less popular, but have also been used to good effect (Malmberg et al., 2005; Ho and Yim, 2006; Roulston et al., 2005). Gaur and Deo (2008) use genetic programming. Reikard (2009) extends regression and neural network techniques to predicting the wave energy flux.

Since these two approaches have evolved independently, it is of interest to determine which approach can predict more accurately, and over what time horizons. A reasonable critique of statistical methods is that they do not capture the underlying physics. However, tests in fields ranging from the sciences to financial economics have demonstrated that time series models can often forecast quite accurately over short horizons.

A further reason to compare the two approaches has to do with the uses of the forecasts. Large-scale simulation models were initially developed primarily to satisfy the requirements of national navies, although application in civilian forecasting centers is now widespread. For instance, forecasts are also used by commercial shipping lines, and typically involve prediction of wave heights over longer horizons, on the order of several hours to a few days. Recently, however, the technology has been developed to make wave farms

* Corresponding author. Tel.: +1 303 734 7768.

E-mail address: Reikarsen@msn.com (G. Reikard).

commercially viable. Since electricity is perishable, utilities will be interested in forecasting over much shorter horizons, on the order of 1–6 h, consistent with generating and selling power.

This paper runs a comparison of the SWAN model with two statistical techniques. The most commonly predicted variable, in the context of a 3G wave model such as SWAN, is the significant wave height, H_{st} , which is the measure of the total energy (i.e., the integrated wave spectrum). Predictions are denoted $H_s(t + \tau)$, where t is the analysis time and τ is the forecast horizon, or look-ahead period. Seven data sets, from the Pacific Ocean and Gulf of Mexico, are used to evaluate the models. Section 2 consists of a review of the models. The setup of the forecasting tests is discussed in Section 3, and results are presented in Section 4. Section 5 concludes.

2. The forecasting models

The SWAN model represents surface waves with the two-dimensional wave action density spectrum $N(\sigma, \theta) = E(\sigma, \theta)/\sigma$, where σ is the intrinsic frequency (i.e., the wave frequency measured from a frame of reference which may include the motion of a current), θ is the wave direction, and E is the spectral energy density. The wave spectrum is described by the spectral action balance equation:

$$\frac{\partial N}{\partial t} + \nabla \cdot CN = \frac{S}{\sigma} \quad (1)$$

where ∇ is the gradient operator in x, y, θ , and σ ; C is the propagation speed, which in the absence of currents is the wave group velocity. The left hand side of this governing equation represents conservative phenomena such as refraction and shoaling. The right-hand side contains the source terms. The total source term S is given in most general form as: $S = S_{input} + S_{diss} + S_{nl}$ i.e., input, dissipation, and nonlinear interactions respectively. Specific source terms include depth-limited wave breaking S_{br} , steepness-limited wave breaking (white-capping) S_{wc} , dissipation by bottom friction S_{bf} , four-wave nonlinear interaction S_{nl4} , and input from the local wind S_{in} . For a discussion of the other source terms, see Booij et al. (1999) and SWAN (2008).

There is a wide range of statistical methods available in the literature, but in the interests of tractability, only two are evaluated. One is a regression with time-varying coefficients. This approach is attractive for two reasons. First, any nonlinear model can in principle be approximated by a stochastic coefficient model (Granger, 2008). Second, empirical tests have often found that time-varying parameter regressions yield the most accurate forecasts for volatile data (see for instance Bunn, 2004). Let $\ln$ denote natural logs, let ω denote a coefficient, let the t -subscript denote time variation, and let ε_t denote the residual. The model is of the form:

$$\ln H_{st} = \omega_{0t} + \omega_{1t} \ln H_{st-1} + \omega_{2t} \ln H_{st-2} + \omega_{3t} \ln H_{st-3} + \omega_{4t} \ln H_{st-4} + \varepsilon_t \quad \varepsilon_t \sim P(0, \nu^2) \quad (2)$$

where P is the probability distribution, and ν^2 is the residual variance. Here the regression includes four lags. With hourly data, all the lags were statistically significant, and this specification was found to predict more accurately than models with shorter or longer lag structures.

The second is a frequency domain method. The reason for going to the frequency domain is that the sea surface can be represented as a spectrum, while waves can be represented as a superposition of cycles at different frequencies. The specific approach used here involves taking the Fourier transform of the moving average representation of the series (Koopmans, 1974, 235–237). Because this approach is not widely known, the mathematical derivation is presented, followed by the steps of the algorithm used in the forecasts. Readers not interested

in the mathematics may wish to omit Eqs. (3)–(7), and proceed directly to the algorithm.

The moving average representation models a time series as a function of its innovations:

$$H_{st} = \beta(L)\varepsilon_t \quad (3)$$

where L is the backshift operator, $\beta(0) = 1$, and ε_t is the residual. This representation is commonly used in the ARIMA framework of Box and Jenkins (1976), where it is modeled in the time domain using a rational polynomial. Let $\phi(L)$ be the autoregressive operator, represented as a polynomial in the backshift operator: $\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$, and let $\theta(L)$ be the moving average operator: $\theta(L) = 1 + \theta_1 L + \dots + \theta_q L^q$. Ignoring constants, the moving average representation is then given by the ratio of the moving average and autoregressive polynomials: $\beta = [\theta(L) / \phi(L)] \varepsilon_t$.

In the frequency domain, spectral methods can be used to calculate the Fourier transform of β , which can be used to forecast. The spectral density (F) of H_{st} can be expressed as a function of the z -transform:

$$F_{Hst} = \beta(z)\beta(z^{-1})\nu^2 \quad (4)$$

The z -transform converts the discrete values of a time series into a complex frequency domain representation, and is given by:

$$z = \Lambda \exp(i\varphi) = \Lambda(\cos \varphi + i \sin \varphi) \quad (5)$$

where Λ is the magnitude of z and φ is the complex argument, or phase, in radians.

Let γ denote a one-sided polynomial in positive powers of z . Then the log spectral density can be expressed as:

$$\ln F_{Hst} = \gamma(z) + \gamma(z^{-1}) + \gamma_0 \quad (6)$$

Taking antilogs, and setting Eq. (4) equal to Eq. (6):

$$\beta(z)\beta(z^{-1})\nu^2 = \exp[\gamma(z)] + \exp[\gamma(z^{-1})] + \exp(\gamma_0) \quad (7)$$

Eq. (7) provides a frequency domain estimate of β , which can then be projected outside the range of the data set.

The steps of the forecasting algorithm are as follows:

- Compute the log spectral density of the time series to be forecasted.
- Inverse Fourier transform, to mask the negative powers.
- Fourier transform again, and take anti-logs in the frequency domain.
- Inverse Fourier transform to estimate β . Normalize so that $\beta(0) = 1$.
- Fourier transform, and filter the transformed series by $1 / \beta(L)$; this computes the residuals.
- Inverse Fourier transform, and mask the residuals outside the range of the data.
- Fourier transform and filter the residuals by $\beta(L)$ in the frequency domain.
- Inverse Fourier transform the filtered residual series, extending the date range to encompass the forecast range.
- Define an equation in the time domain such that the series to be forecasted is a deterministic function of the series defined in the previous step, and forecast this equation.

In the Regression Analysis of Time Series software package used to run the statistical models, the frequency domain algorithm is available as an automated routine, while the code can also be extracted for more flexible programming. The website of this software package is: <http://www.estima.com>.

One crucial issue in the statistical models is the method of estimating the time-varying coefficients. Time-varying parameters can be estimated using a Kalman filter (Kalman, 1960), or by sliding

window techniques. When the Kalman filter is unrestricted, the coefficients behave like a random walk. In highly variable time series, however, the unrestricted Kalman filter is susceptible to finding spurious patterns in random events. The degree of variation in the coefficients can be reduced either by restricting the filter, or by adjusting the widths of the moving window. Preliminary experiments were run with the data sets below to determine the optimal degree of restriction on the Kalman filter, by running forecasting tests and comparing the errors. Generally, the tests favored high degrees of restriction.

A second set of experiments were run using a moving window, and comparing the errors associated with various window widths. At short widths, on the order of 100 h or less, the results were similar to the unrestricted Kalman filter. Over wider widths, on the order of 600 to 1200 h, the models became too inertial, and did not respond well to changing conditions. The optimal window widths from the standpoint of overall forecast accuracy were in the range of 400 to 600 h. On the basis of the tests, a width of 480 observations was used. The spectral algorithm was found to be less sensitive to changes in the window widths, but here also 480 h yielded good results. Additional tests were run for data sets in a variety of other locations, including deep water sites in the Atlantic and Pacific. Again, the most accurate forecasts were obtained when the window widths were in the range of 400 to 600 h.

Finally, a persistence forecast – setting the predicted value equal to the most recently observed actual value – is also reported. This method is commonly used to provide a benchmark forecast in econometrics (Theil, 1971).

3. The forecasting tests

The physics-based numerical model outputs are all taken from archived output from real time forecasting systems designed and operated the Naval Research Laboratory, using SWAN. Inputs to these systems were official operational wind and wave products. Thus the wave predictions from these systems can be fairly portrayed as completely “blindfolded”. Further, they did not benefit from post facto improvements to forcing fields, as might be the case for hindcast simulations driven by observational data and/or reanalysis wind fields. The first forecasting system is the Coastal Northwest (CNW) system for the Washington and Oregon coastlines (<http://www7320.nrlssc.navy.mil/CNW/>). This system operated from October 2004 to May 2009. The initial setup was performed for the NOAA Coastal Storms Program. The second system is the north central Gulf of Mexico (GMEX) system, which was initiated in October 2006 and continues to operate at time of writing. The initial setup was performed for the CenGOOS (Central Gulf Ocean Observing System) program (<http://www7320.nrlssc.navy.mil/CenGOOS/>).

Table 1 reports the locations of the sites, the dates of the forecasts. The grid domains for these two systems are shown in Figs. 1 and 2. Some features of the model setup are:

- Spherical coordinates were used, with grid spacing as shown in Table 2.
- The directional resolution $\Delta\theta = 10^\circ$ in all cases.
- The frequency grid is logarithmic, with 34 frequencies from 0.0418 Hz to 1.0 Hz.
- The default parameterizations in SWAN are that of WAM, Cycle 3, sometimes referred to in the literature as “WAM3 physics”. In this study, the default parameterizations for S_{in} , S_{ds} , S_{nlq} are used, except that the integer used for the weighting of relative wave number was increased by 1.0 (see also Rogers et al., 2003 and Janssen et al. 1989).
- For the fully non-stationary models, a time step of 10 min is used.
- For the pseudo non-stationary models (see below), the default convergence criterion for iterations was used.
- The interval of wind input is 3 h in all cases.

Table 1
The SWAN forecasts and locations.

NDBC Buoy	SWAN Forecasts	Location	Latitude, Longitude
46050	714 values, Jan. 2008 to May 2009	Stonewall Banks, Oregon, 20 NM West	44.641 N 124.500 W
46041	714 values, Jan. 2008 to May 2009	Cape Elizabeth, Washington, 45 NM Northwest	47.353 N 124.731 W
46029	714 values, Jan. 2008 to May 2009	Columbia River, Oregon, 20 NM West	46.144 N 124.510 W
46211*	714 values, Jan. 2008 to May 2009	Gray's Harbor, Washington, 4.5 NM Southwest	46.515 N 124.146 W
42007	1679 values, Oct. 2007 to May 2010	Biloxi, Mississippi, 22 NM South Southeast	30.090 N 88.769 W
42040	1679 values, Oct. 2007 to May 2010	Mobile, Alabama, 64 NM South	29.212 N 88.207 W
42039	306 values, Dec. 2009 to May 2010	Pensacola, Florida, 115 NM East Southeast	28.791 N 86.008 W

NDBC: National data buoy center (<http://www.ndbc.noaa.gov>).

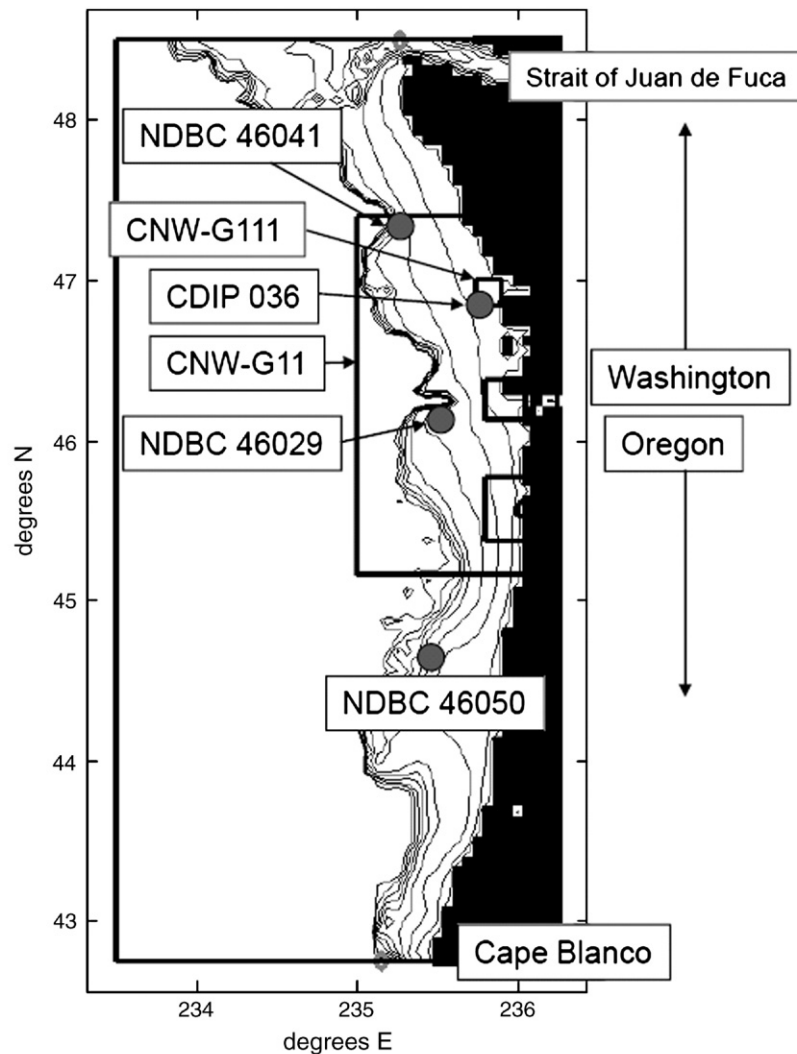
*Equivalent to Coastal Data Information Program site 036 (<http://www.cdip.ucsd.edu>).
NM = Nautical miles.

- In the case of the CNW system, the resolution of the forcing provided to the outer SWAN nest (CNW-G1) case is at 1° intervals along the boundary, and this was obtained from the operational NCEP WW3 Eastern North Pacific (ENP) model, which has a resolution of 0.25° .
- The Gulf of Mexico outer grid, denoted GMEX-G1 in Table 2, includes the entire Gulf and the Caribbean north of 18° N. Boundary forcing is not applied to this grid. Swells from east of the Florida Straits or from the southern Caribbean are not included in the forecast, since they typically have a negligible influence in the northern Gulf.
- In all cases, boundary forcings are provided as full directional spectra (i.e. not parameterized).
- The JONSWAP bottom friction formulation is used (see SWAN manual). A non-default friction coefficient is used in the Gulf of Mexico SWAN models. The value is based on an unpublished document by Dr. Hendrik Tolman (NCEP).
- For the CNW SWAN models, the default setting for depth-limited breaking is used. For the GMEX SWAN models, depth-limited breaking is disabled.
- Default propagation schemes are used in all cases. The default schemes for geographic propagation in SWAN are second order accurate. For stationary computations, numerical diffusion is second order. For nonstationary computations, numerical diffusion is third order (see Rogers et al., 2002).
- The GMEX-G11 SWAN model includes surface currents in its forcing. These fields are taken from a $1/25^\circ$ implementation of the HYCOM model, also operated by NRL.

Additional details for these forecasts are given in Table 2. “Fully nonstationary” indicates that computations are made in full time stepping mode. “Pseudo-nonstationary” indicates that the model uses a time-sequence of computations that utilize the stationary assumption (see discussion in Rogers et al., 2007). The deepwater source term parameterizations used are not tuned for this simulation or for this area; rather they are the same as what are used in SWAN forecasting systems run at NRL for other areas.

Unlike the statistical model, the physics-based model does not use wave observations to produce its forecast. Instead, in SWAN, wave energy is introduced in two ways: Either by the model source term's response to wind forcing (i.e., the right hand side of Eq. (1)), or by inputs of directional spectral at the grid boundary, provided by a host wave model, as indicated in Table 2.

At the four Pacific Coast sites, the forecasts were run for selected dates spanning the period from January 1, 2008 through May 17, 2009, and for the following horizons: $\tau = 0, 3, 6, 9, 12$ and 15 h. For

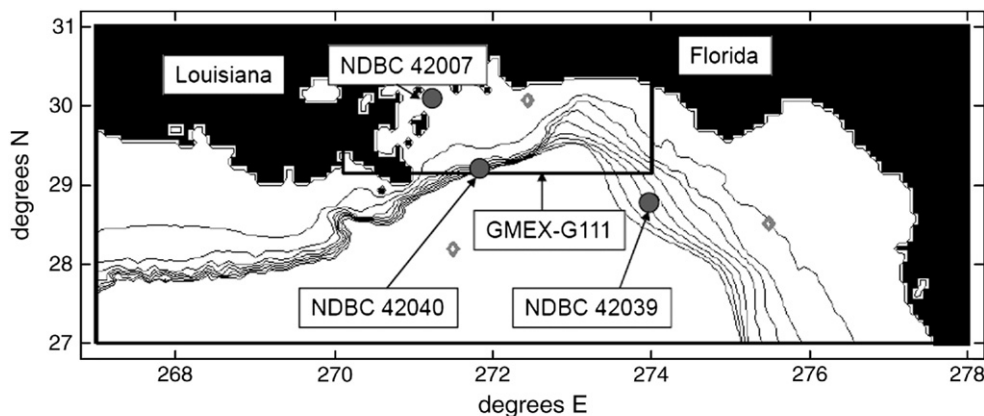


Notes: Depths contours are drawn at 50 m intervals, from 400 m to 50 m. CNW-G11 is the larger grid encompassing all the sites used here. CNW-G111 refers to a smaller grid close to the Gray's Harbor site.

Fig. 1. Diagram of Coastal Northwest Grid 1 (CNW-G1).

each value of τ , 714 forecasted values are available. Different time spans and forecast horizons were used at the Gulf sites. At the Biloxi buoy (NOAA NDBC 42007) the forecasts were also available for $\tau=0$,

3, 6, 9, 12, and 15 h. The time span is from October 25, 2007 through May 17, 2010, and 1679 predicted values are available for each value of τ . At the Mobile buoy (42040) the time span is the same, but the



Notes: Depths contours are drawn at 50 m intervals, from 400 m to 50 m.

Fig. 2. Diagram of Gulf of Mexico grid 11 (GMEX-G11).

Table 2
Additional Detail for the SWAN Forecasts.

Grid ID	CNW-G1	CNW-G11	CNW-G111	GMEX-G11	GMEX-G111
Grid descr.	Coastal Northwest (outer)	Coastal Northwest (shelf)	Grays Harbor	North-Central Gulf of Mexico (outer)	North-Central Gulf of Mexico (shelf)
Δx (longitude)	3.75' or 4.9 km	0.92' or 1.17 km	0.229' or 289 m	4'–6.5 km	1.17'–1.88 km
Δy (latitude)	3.75' or 6.9 km	0.92' or 1.71 km	0.232' or 430 m	4' = 7.4 km	1.17' = 2.18 km
Origin (°E, °N)	233.50 42.75	235.00 45.16	235.74 46.84	267.0 27.0	270.12 29.15
(# x-cells, # y-cells)	45 93	73 147	43 45	164 59	198 78
Bathymetry	1' NGDC	15" NGDC	3" NGDC	NRL DBDB (2')	NRL DBDB (1')
Computation mode	Fully nonstationary	Pseudo-nonstationary	Pseudo-nonstationary	Fully nonstationary	Pseudo-nonstationary
Bottom friction coefficient	$C_f = 0.067$ (default)			$C_f = 0.019$	
Output locations	46050 (depth = 123 m)	46041 (132 m) 46029 (135 m)	46211 (38 m) (CDIP-036)	42039 (307 m) 42040 (165 m)	42007 (14 m)
Boundary forcing from	NCEP WW3 ENP 15' × 15' resolution	SWAN CNW-G1	SWAN CNW-G11	SWAN GMEX-G1	SWAN GMEX-G11
Wind forcing	NCEP GFS ENP, 15' × 15' resolution			FNMOCCOAMPS CENAM, 12' × 12' resolution	
Boundary forcing to	CNW-G11	CNW-G111, G112, G113	None	GMEX-G111	none
Output interval	3 h			hourly	3 h
forecast range (days)	3	2.0	1.5	2.0	2.0

values of τ run continuously from zero through 15 h. There are 1679 forecasts for every value of τ . At the Pensacola buoy (42039) the time span is from December 12, 2009 through May 15, 2010. Again, the values of τ run from zero through 15 h, and 306 predictions are available for each value of τ .

The data sets used to validate the forecasts were downloaded from the National Oceanographic and Atmospheric Administration's National Data Buoy Center (<http://www.ndbc.noaa.gov>). Table 3 reports the available dates, along with other properties of the data. Fig. 3 shows the wave height over a 12-month period from one of the Pacific Coast sites, Cape Elizabeth. Data from the other three Pacific sites are comparable. At all of these sites, the mean wave height ranges from 2.09 to 2.27 m, and the standard deviation is slightly over 1 m. There is evidence of excess kurtosis. Fig. 4 shows one of the three Gulf sites, Pensacola. The Gulf data shows somewhat different properties. There are more outlying fluctuations, and greater seasonal volatility. As Table 3 indicates, the mean wave heights in the Gulf are lower, but the kurtosis is significantly greater.

A major issue in the observational data is missing calendar dates, and missing values. The databases were interpolated in two steps. First, the date fields were used to create a continuous calendar.

Second, the missing values were interpolated using a polynomial. To verify that the interpolations produced realistic numbers, the statistical models were tested both over the interpolated data sets and on the original data, simply omitting any missing values. The forecast errors were extremely close, indicating that interpolation did not appreciably bias the results. However, in some cases, data were missing for several months at a time (see Table 3). These values were simply omitted.

The experiments using the statistical models were set up as follows. The forecast horizons were the same as for the SWAN model, except that 1–2 and 4–5 hour predictions were also generated. The forecasts were run over all the available data, and two sets of errors are reported, for the entire data set, and for the same values in the SWAN forecasts. The models were estimated over an initial set of starting values, and were then forecasted a given number of steps ahead. At the next step the models were re-estimated and forecasted, continuing through the end of the data set. All forecasting was dynamic, and the predicted values are true out-of-sample forecasts, in that they use only data prior to the start of the forecast horizon. The reported error for horizons beyond one period is for the forecast for that horizon only; intervening values are omitted. In other words, the error at $\tau = 3$ omits the errors at $\tau = 1$ and $\tau = 2$ h.

Table 3
Properties of the buoy data.

Wave height statistics						
NDBC Id / Site name	Water depth	Mean	Standard Deviation	Kurtosis	Dates Available	
46050 Stonewall Banks	123 m	2.21	1.09	2.69	Mar. 5, 2008 to May 31, 2009	
46041 Cape Elizabeth	132 m	2.27	1.18	1.64	Jan. 1, 2008 to May 31, 2009	
46029 Columbia River	135 m	2.25	1.13	2.75	Mar. 3, 2008 to May 31, 2009*	
46211 Gray's Harbor	40 m	2.04	1.08	2.21	Jan. 1, 2008 to Dec 31, 2009	
42007 Biloxi	35 m	0.76	0.51	17.06	Oct. 1, 2007 to Dec 31, 2009**	
42040 Mobile	161 m	1.06	0.77	17.45	Oct. 1, 2007 to Oct 5, 2009***	
42039 Pensacola	307 m	1.41	0.83	4.24	Dec. 1, 2009 to Mar. 31, 2010	

*The Columbia River site is missing data from April 15 through June 20, 2008.

**The Biloxi site is missing data from June 24 through August 31, 2008.

*** The Mobile site is missing data from February 15 through May 18, 2008.

4. Results

Table 4 reports the results for the Pacific Coast, and Table 5 reports the results for the Gulf. The first three columns report the errors from the SWAN forecasts and the statistical forecasts only for the same horizons and values as in the SWAN simulations. The error is measured as the mean absolute error, i.e., the absolute difference between the forecast and the actual value. For $\tau = 0$, the regression forecast is of course the in-sample fit. No frequency domain forecast is reported for $\tau = 0$, since the fit would be nearly perfect. Since the SWAN model does not utilize the observational data, its prediction for $\tau = 0$ is, by contrast, not perfect. The next three columns report the errors from the statistical models and persistence forecasts over all available data points.

In all the sites, the statistical models are found to predict more accurately at short horizons, while SWAN predicts more accurately over longer horizons. At $\tau = 1$ –5, the errors from the regression and spectral algorithms are significantly lower than in the SWAN forecasts. The statistical model errors over the entire data set are in the same overall range as the errors for the smaller data set for which

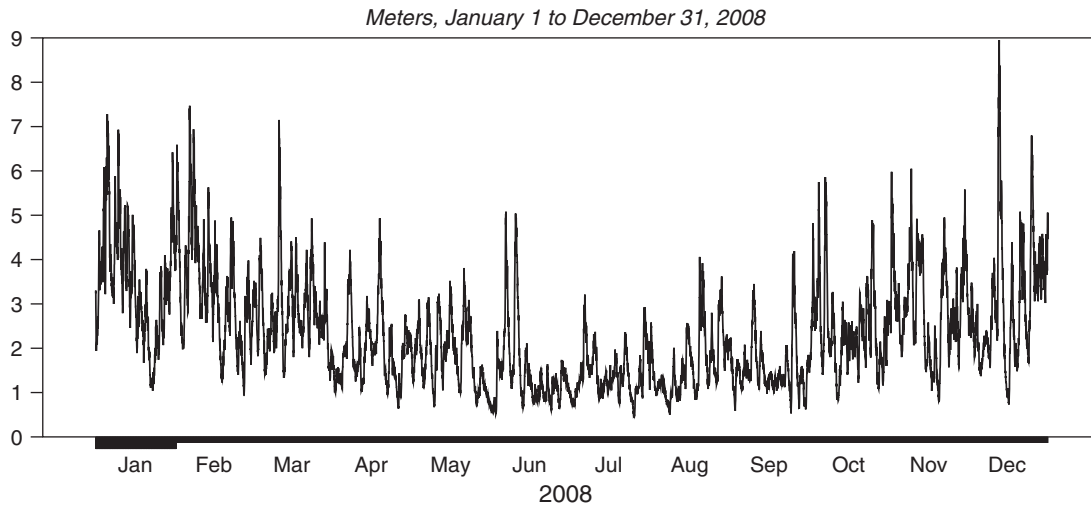


Fig. 3. The significant wave height, Cape Elizabeth.

the SWAN predictions are available — as would be expected if the values used in the SWAN forecasts represent a random sample of the larger data set. Among the statistical methods, the frequency domain algorithm is somewhat better than the regression. At all the Pacific Coast sites, both statistical models are superior to the persistence forecast. However, as the horizon increases, the accuracy of the statistical models deteriorates. At $\tau = 6$ h, the statistical models and SWAN show comparable forecast errors at nearly all the sites. At $\tau > 6$, SWAN is unambiguously superior.

The tests for the Mobile and Pensacola locations are of particular interest, since they include a larger number of forecast horizons for SWAN. In both instances, however, the results are similar. The errors from the statistical models significantly lower than the SWAN errors at very short horizons, but the gap between the two falls as the horizon increases, with the convergence points invariably occurring at or near 6 h. At $\tau = 7$, the statistical model errors start to exceed the SWAN error. In all three Gulf sites, the frequency domain algorithm is superior to the regression. Interestingly, the persistence forecast does better at the Gulf sites. At the Biloxi and Mobile sites, it is comparable to the regression, although the spectral algorithm is consistently better. At the Pensacola site, the persistence forecast is marginally better than all the other models for the first 5 h.

5. Conclusions

The tests have yielded one central conclusion. The accuracy of the models is a function primarily of the forecast horizon. At very short horizons, up to 5 h, the statistical models achieve more accurate predictions. However, as the horizon extends, the accuracy of the statistical models falls off rapidly. At horizons beyond 6 h, the SWAN forecasts are uniformly more accurate. As a general rule, the crossover point, at which the accuracy of the two methods converges, in the area of 6 h. A notable feature of the SWAN forecasts is that they maintain the same level of accuracy at longer horizons. Additional tests of the SWAN model have found that the error only decays after longer period of time, on the order of several days.

One question that arises here is whether SWAN's performance at short horizons can be explained by the forcing terms used in the model. To evaluate this, the wind speed data at one of the sites (Mobile) was compared with the SWAN wind forcing values for the same location. The error is extremely similar over a range of horizons, similar to the performance of SWAN itself on wave height. When regression models were estimated for the Mobile wind speed, the error was found to be lower at the short horizons, and higher at longer horizons, with a crossover point in the area of 8 h.

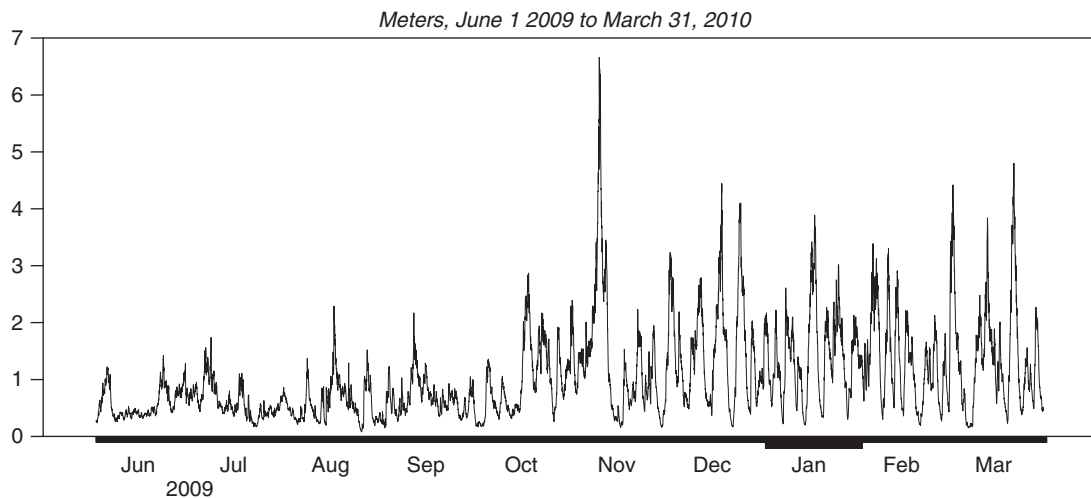


Fig. 4. The significant wave height, Pensacola.

Table 4

Comparison of forecast errors, Pacific Coast sites. Statistics are Mean Absolute Error.

Forecast Horizon	Tests for simulation values			Tests for all values		
	SWAN	Regression	Spectral	Regression	Spectral	Persistence
<i>Stonewall Banks</i>						
0	0.297	0.133	...	0.131	...	...
1	...	...	...	0.130	0.129	0.140
2	...	...	...	0.168	0.161	0.182
3	0.295	0.197	0.192	0.206	0.194	0.223
4	...	...	...	0.242	0.228	0.261
5	...	...	...	0.290	0.272	0.313
6	0.299	0.303	0.295	0.315	0.295	0.340
9	0.309	0.392	0.368	0.396	0.371	0.431
12	0.298	0.447	0.422	0.452	0.427	0.501
15	0.307	0.498	0.471	0.503	0.479	0.562
<i>Cape Elizabeth</i>						
0	0.306	0.132	...	0.127	...	...
1	...	...	...	0.126	0.119	0.131
2	...	...	...	0.165	0.158	0.172
3	0.331	0.205	0.187	0.208	0.201	0.217
4	...	...	...	0.247	0.238	0.258
5	...	...	...	0.301	0.288	0.314
6	0.327	0.311	0.293	0.328	0.314	0.343
9	0.334	0.397	0.369	0.422	0.401	0.441
12	0.356	0.455	0.421	0.495	0.428	0.521
15	0.351	0.511	0.491	0.553	0.538	0.592
<i>Columbia River</i>						
0	0.319	0.131	...	0.125	...	...
1	...	...	...	0.125	0.124	0.133
2	...	...	...	0.148	0.148	0.174
3	0.322	0.182	0.186	0.183	0.185	0.219
4	...	...	...	0.219	0.217	0.243
5	...	...	...	0.267	0.259	0.275
6	0.328	0.308	0.289	0.292	0.281	0.292
9	0.338	0.375	0.339	0.368	0.352	0.432
12	0.342	0.427	0.394	0.426	0.406	0.508
15	0.364	0.474	0.449	0.479	0.461	0.578
<i>Gray's Harbor</i>						
0	0.276	0.094	...	0.069	...	...
1	...	...	...	0.072	0.078	0.106
2	...	...	...	0.135	0.138	0.152
3	0.288	0.181	0.179	0.182	0.178	0.199
4	...	...	...	0.214	0.208	0.238
5	...	...	...	0.257	0.248	0.291
6	0.286	0.266	0.258	0.279	0.269	0.318
9	0.295	0.336	0.226	0.364	0.344	0.406
12	0.302	0.394	0.368	0.430	0.399	0.472
15	0.295	0.445	0.416	0.469	0.447	0.535

Using statistical forecasts of wind as an input to SWAN would however be problematic. In SWAN, waves are not generated by wind at a single site, but by wind over a range of locations. The wave model is a temporal and spatial integration, or stated another way, a function of wind histories at several different points. Using the statistical wind forecast would require a simplified version of SWAN that would not consider propagation, or spatial gradients. Further, the goal in this paper is not to evaluate individual inputs, but rather to compare the accuracy of the SWAN modeling system, in its entirety, to the statistical models. The SWAN model physics, wind forcing, bathymetry input, and boundary forcing are all part of the SWAN modeling system and all contribute to its performance.

Instead, the main reason for the differences between the two types of models lies with the data itself. At high frequencies, the data exhibit high degrees of dependence between time points (or in statistical terms, serial correlation). Over short horizons, the data are dominated by this dependence. The statistical models are able to parameterize this, and predict more accurately. However, at slightly lower frequencies, the dependence dissipates. At these longer horizons, the data is more dominated by the underlying signals. As a result, physics-based models are able to forecast more effectively.

Table 5

Comparison of forecast errors, Gulf of Mexico sites. Statistics are Mean Absolute Error.

Forecast Horizon	Test for simulation values			Tests for all values		
	SWAN	Regression	Spectral	Regression	Spectral	Persistence
<i>Biloxi</i>						
0	0.186	0.059	...	0.065	...	...
1	...	...	...	0.068	0.061	0.066
2	...	...	...	0.087	0.082	0.085
3	0.169	0.110	0.096	0.112	0.098	0.110
4	...	...	...	0.134	0.116	0.131
5	...	...	...	0.163	0.141	0.159
6	0.174	0.179	0.156	0.178	0.154	0.174
9	0.178	0.202	0.192	0.203	0.194	0.224
12	0.182	0.227	0.227	0.233	0.228	0.262
15	0.189	0.251	0.254	0.255	0.256	0.292
<i>Mobile</i>						
0	0.200	0.068	...	0.069	...	...
1	0.199	0.069	0.064	0.071	0.065	0.070
2	0.199	0.096	0.093	0.095	0.094	0.100
3	0.204	0.129	0.122	0.128	0.124	0.131
4	0.207	0.156	0.149	0.158	0.151	0.159
5	0.203	0.178	0.176	0.181	0.179	0.188
6	0.200	0.201	0.200	0.204	0.201	0.210
7	0.201	0.220	0.209	0.217	0.203	0.233
8	0.201	0.241	0.226	0.238	0.221	0.253
9	0.203	0.263	0.245	0.265	0.249	0.272
10	0.198	0.282	0.260	0.284	0.262	0.289
11	0.199	0.301	0.276	0.305	0.279	0.306
12	0.198	0.311	0.283	0.309	0.283	0.321
13	0.205	0.331	0.312	0.329	0.308	0.335
14	0.200	0.349	0.318	0.347	0.317	0.349
15	0.206	0.371	0.338	0.373	0.339	0.362
<i>Pensacola</i>						
0	0.295	0.084	...	0.092	...	...
1	0.280	0.085	0.076	0.093	0.083	0.076
2	0.272	0.129	0.110	0.124	0.119	0.112
3	0.273	0.171	0.152	0.170	0.147	0.146
4	0.270	0.206	0.190	0.208	0.188	0.182
5	0.266	0.245	0.220	0.238	0.212	0.214
6	0.270	0.263	0.235	0.276	0.241	0.249
7	0.272	0.318	0.276	0.308	0.273	0.272
8	0.267	0.339	0.302	0.337	0.303	0.298
9	0.278	0.367	0.321	0.363	0.323	0.322
10	0.286	0.398	0.358	0.389	0.348	0.344
11	0.293	0.427	0.391	0.412	0.372	0.366
12	0.295	0.446	0.387	0.435	0.385	0.386
13	0.281	0.467	0.394	0.456	0.402	0.405
14	0.277	0.469	0.399	0.464	0.421	0.423
15	0.283	0.487	0.415	0.492	0.448	0.439

These results suggest that the choice of statistical versus physics-based models for forecasting will depend on the intended uses. Utilities operating wave farms will need to predict power flows at horizons of only a few hours, and for this reason may prefer to rely on statistical techniques. Navies or shipping companies interested in oceanic conditions over longer horizons, and in denied areas or other locations where *in situ* data are not available to drive statistical forecasts, will prefer physics-based models.

The results also point to some directions for further research. For certain applications, it may be possible to combine physics-based and statistical models. The combined prediction from a hybrid model combining elements of both approaches may be superior to either method individually. Another extension is prediction of the wave energy flux, rather than the wave height alone. The next stage in this research is a combination of physics-based and statistical models for wave energy.

Acronyms and abbreviations

CenGOOS	Central Gulf Ocean Observing System
CDIP	Coastal Data Information Program
CNW	Coastal Northwest SWAN forecasting system

COAMPS	Coupled Ocean/Atmosphere Mesoscale Prediction System (Hodur, 1997)
DBDB	NRL Digital Bathymetric Data Base
ENP	Eastern North Pacific, referring to a WW3 implementation at NCEP
FNMOG	Fleet Numerical Meteorology and Oceanography Center
GMEX	north central Gulf of Mexico forecasting system
HYCOM	HYbrid Coordinate Ocean Model (Bleck, 2002).
JONSWAP	Joint North Sea Wave Project, (Hasselmann et al., 1973)
NCEP	National Centers for Environmental Prediction
NDBC	National Data Buoy Center
NGDC	National Geophysical Data Center
NOAA	National Oceanic and Atmospheric Administration
SWAN	Simulating WAVes Nearshore (Booij et al., 1999)
WAM	WAVE Model (WAMDIG, 1988; Komen et al., 1994)
WW3	WAVEWATCH III® (Tolman, 1991; Tolman, 2002)

References

- Bleck, R., 2002. An oceanic general circulation model framed in hybrid isopycnic-cartesian coordinates. *Ocean Modelling* 4, 55–88.
- Booij, N., Ris, R.C., Holthuijsen, L.H., 1999. A third-generation model for coastal regions. Part I: Model description and validation. *Journal of Geophysical Research* 104, 7649–7666.
- Box, G.E.P., Jenkins, G.J., 1976. *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.
- Bunn, D.W., 2004. *Modelling Prices in Competitive Electricity Markets*. Wiley, New York.
- Deo, M.C., Jagdale, S.S., 2003. Prediction of breaking waves with neural networks. *Ocean Engineering* 30, 1163–1178.
- Deo, M.C., Naidu, C.S., 1998. Real time wave forecasting using neural networks. *Ocean Engineering* 26, 191–203.
- Deo, M.C., Jha, A., Chaphekar, A.S., Ravikant, K., 2001. Neural Networks for Wave Forecasting. *Ocean Engineering* 28, 889–898.
- Dykes, J.D., Wang, D.W., Book, J.W., 2009. An evaluation of a high-resolution operational wave forecasting system in the Adriatic Sea. *Journal of Marine Systems* 78, 255–271.
- Gaur, S., Deo, M.C., 2008. Real-time wave forecasting using genetic programming. *Ocean Engineering* 35, 1166–1172.
- Granger, C.W.J., 2008. Non-Linear Models: Where Do We Go Next – Time Varying Parameter Models? : *Studies in Nonlinear Dynamics and Econometrics* 12: Article 1 <http://www.bepress.com/snnde/vol12/iss3/art1>.
- Hasselmann, K., and Coauthors, 1973: Measurements of wind wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP). Herausgegeben vom Deutsch. Hydrograph. Institut., Reihe A, #12.
- Hasselmann, S., Hasselmann, K., Allender, J.H., Barnett, T.P., 1985. Computations and parameterizations of the non-linear energy transfer in a gravity wave spectrum. Part II: parameterizations of the non-linear energy transfer for application in wave models. *Journal of Physical Oceanography* 15, 1378–1391.
- Ho, P.C., Yim, J.Z., 2006. Wave height forecasting by the transfer function model. *Ocean Engineering* 33, 1230–1248.
- Hodur, R.M., 1997. The Naval Research Laboratory's Coupled Ocean/Atmospheric Mesoscale Prediction System (COAMPS). *Monthly Weather Review* 125, 1414–1430.
- Jain, P., Deo, M.C., 2007. Real-time wave forecasts off the western Indian coast. *Applied Ocean Research* 29, 72–79.
- Janssen, P., 1991. Quasi-linear theory of wind-wave generation applied to wave forecasting. *Journal of Physical Oceanography* 21, 1631–1642.
- Janssen, P., Lionello, P., Reistad, M., Hollingsworth, A., 1989. Hindcasts and data assimilation studies with the WAM model during the Seasat period. *Journal of Geophysical Research* C94, 973–993.
- Jensen, R.E., Wittmann, P.A., Dykes, J.D., 2002. Global and regional wave modeling activities. *Oceanography* 15, 57–66.
- Kalman, R.E., 1960. A new approach to linear filtering and prediction problems, *transactions of the American Society of Mechanical Engineers. Journal of Basic Engineering* 83D, 35–45.
- Komen, G.J., Hasselmann, S., Hasselmann, K., 1984. On the existence of a fully developed wind-sea spectrum. *Journal of Physical Oceanography* 14, 1271–1285.
- Komen, G.J., Cavaleri, L., Donelan, M., Hasselmann, K., Hasselmann, S., Janssen, P.A.E.M., 1994. *Dynamics and modelling of ocean waves*. Cambridge University Press.
- Londhe, S.N., Panchang, V., 2006. One-day wave forecasts based on artificial neural networks. *Journal of Atmospheric and Oceanic Technology* 23, 1593–1603.
- Makarynsky, O., 2004. Improving wave predictions with artificial neural networks. *Ocean Engineering* 31, 709–724.
- Malmberg, A., Holst, U., Holst, J., 2005. Forecasting near-surface ocean winds with Kalman filter techniques. *Ocean Engineering* 32, 273–291.
- Miles, J.W., 1957. On the generation of surface waves by shear flows. *Journal of Fluid Mechanics* 3, 185–204.
- Ozger, M., 2010. Significant wave height forecasting using wavelet fuzzy logic approach. *Ocean Engineering* 37, 1443–1451.
- Phillips, O.M., 1957. On the generation of waves by a turbulent wind. *Journal of Fluid Mechanics* 2, 417–445.
- Phillips, O.M., 1958. The equilibrium range in the spectrum of wind-generated waves. *Journal of Fluid Mechanics* 4, 426–434.
- Reikard, G., 2009. Forecasting ocean wave energy: tests of time series models. *Ocean Engineering* 36, 348–356.
- Ris, R.C., Booij, N., Holthuijsen, L.H., 1999. A third-generation wave model for coastal region, Part II: verification. *Journal of Geophysical Research* 104 (C4), 7667–7681.
- Roulston, M.S., Ellepola, J., von Hardenberg, J., Smith, L.A., 2005. Forecasting wave height probabilities with numerical weather prediction models. *Ocean Engineering* 32, 1841–1863.
- Rogers, W.E., Kaihatu, J.M., Petit, H.A.H., Booij, N., Holthuijsen, L.H., 2002. Diffusion reduction in an arbitrary scale third generation wind wave model. *Ocean Engineering* 29, 1357–1390.
- Rogers, W.E., Hwang, P.A., Wang, D.W., 2003. Investigation of wave growth and decay in the SWAN Model: three regional-scale applications. *Journal of Physical Oceanography* 33, 366–389.
- Rogers, W.E., Kaihatu, J.M., Hu, L., Jensen, R.E., Dykes, J.D., Holland, K.T., 2007. Forecasting and hindcasting waves with the SWAN model in the Southern California bight. *Coastal Engineering* 54, 1–15.
- Snyder, R.L., Dobson, F.W., Elliott, J.A., Long, R.B., 1981. Array measurements of atmospheric pressure fluctuations above surface gravity waves. *Journal of Fluid Mechanics* 102, 1–59.
- SWAN team, 2008. SWAN Technical Documentation. Delft University of Technology <http://www.swan.tudelft.nl>.
- Theil, H., 1971. *Principles of Econometrics*. Wiley, New York.
- Tolman, H.L., 1991. A third generation model for wind waves on slowly varying, unsteady and inhomogeneous depths and currents. *Journal of Physical Oceanography* 21, 782–797.
- Tolman, H.L., 2002. User manual and system documentation of WAVEWATCH-III version 2.2.2. : NCEP Technical Note <http://polar.ncep.noaa.gov/waves/references.html>.
- Tolman, H.L., Balasubramanian, B., Burroughs, L.D., Chalikov, D.V., Chao, Y.Y., Chen, H.S., Gerald, V.M., 2002. Development and implementation of wind-generated ocean surface wave models at NCEP. *Weather Forecasting* 17, 311–333.
- Tsai, C.P., Lin, C., Shen, J.N., 2002. Neural network for wave forecasting among multi-stations. *Ocean Engineering* 29, 1683–1695.
- Tseng, C.M., Jan, C.D., Wang, J.S., Wang, C.M., 2007. Application of artificial neural networks in typhoon surge forecasting. *Ocean Engineering* 34, 1757–1768.
- WAMDIG Group, 1988. The WAM model—a third generation ocean wave prediction model. *Journal of Physical Oceanography* 18, 1775–1810.
- The WISE Group, 2007. Wave Modelling – The state of the art. *Progress in Oceanography* 75, 603–674.
- Zamani, A., Solomatine, D., Azimian, A., Heemink, A., 2008. Learning from data for wind-wave forecasting. *Ocean Engineering* 35, 953–962.