Coastal Engineering 56 (2009) 759–772

Contents lists available at ScienceDirect

Coastal Engineering

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

Wind-wave variability in a shallow tidal sea—Spectral modelling combined with neural network methods

Agnieszka Herman [a,*], Ralf Kaiser [b], Hanz D. Niemeyer [b]

[a] Institute of Oceanography, University of Gdańsk, Pilsudskiego 46, 81-378 Gdynia, Poland
[b] Coastal Research Station, Lower Saxony Water Management, Coastal Defence and Nature Conservation Agency, An der Mühle 5, 26548 Norderney, Germany

ARTICLE INFO

Article history:
Received 28 February 2008
Received in revised form 11 February 2009
Accepted 16 February 2009
Available online 17 March 2009

Keywords:
Tidal inlets
Spectral wave modelling
Principal component analysis
Neural networks
North Sea
Wadden Sea

ABSTRACT

In this paper the wind-wave variability in the tidal basins of the German Wadden Sea is modelled with combined numerical and neural-network (NN) methods. First, the wave propagation and transformation in the study area are modelled with the state-of-the-art third-generation spectral wave model SWAN. The ability of SWAN to accurately reproduce the phenomena of interest in nonstationary conditions governed by highly variable winds, water levels and currents is shown by comparisons of the modelled and measured mean wave parameters at four stations. The principal component analysis of the SWAN results is then used to reveal the dominating spatial patterns in the data and to reduce their dimensionality, thus enabling an efficient and relatively straightforward NN modelling of mean wave parameters in the whole study area. It is shown that the data produced with the approach developed in this work have statistical properties (discrete probability distributions of the mean wave parameters etc.) very close to the properties of the data obtained with SWAN, thus proving that this approach can be used as a reliable tool for wind wave simulation in coastal areas, complementary to (often computationally demanding) spectral wave models.

1. Introduction

The knowledge concerning the short- and medium-term variability of waves and currents is an important prerequisite for the studies of many important aspects of the coastal zone processes, e.g. for the analysis of sediment transport or the safety assessment of coastal protection structures and the coast itself. A prerequisite for such investigations is the availability of a sufficient amount of data spanning time periods important for the processes under study. The work presented here is part of a research project MOSES ("Modelling of the medium-term wave climatology at the German North Sea coast"), one of the purposes of which is to produce a medium-term (~40 years) database of water levels, currents and mean wave parameters for a coastal area in the German Wadden Sea. Although nowadays the state-of-the-art hydrodynamic and wave models are able to reproduce the wave and current processes in shallow tidal seas with high accuracy, their application with sufficiently high temporal and spatial resolution over longer periods of time still requires substantial computer resources. Thus, computationally cheap and effective alternatives to this kind of modelling—like the technique described in this work—are still of practical use and have many advantages in practical applications. Also, an important aspect is the proper reduction of the amount of data without loss of information crucial for the understanding of the processes involved—an approach known as 'input reduction' in morphodynamic studies, see e.g. Latteux (1995) or a review paper by de Vriend et al. (1993).

The hydrodynamic modelling in the study area is described in detail in Herman et al. (2007). In the present study the results of high-resolution wind-wave simulations are used as a starting point for the development of a neural-network-based modelling system, which enables fast and accurate prediction of temporal and spatial patterns of significant wave height (H_S), energy period ($T_e \equiv T_{-1,0}$) and mean wave direction (θ_m) in the study area—a set of parameters used e.g. in studies concerning design wave conditions or the safety of the coastal zone (see Appendix A.1 for the definitions of these parameters). The main idea behind the approach developed is to decompose the dataset into a (small) number of fixed modes, assumed 'universal' over the four decades studied, and to model the time evolution of these modes, thus reducing the dimensionality of the problem by more than three orders of magnitude. The results presented in this paper show that the assumption of the decadal 'time-universality' of the modes is justified, provided they are determined from data covering a sufficiently long period of time (in this case, 2 years).

The usage of artificial neural networks (NNs) in oceanic and atmospheric applications has increased rapidly in recent years. A broad review of various aspects of these methods, their possibilities and limitations is given in Cherkassky et al. (2006) in their paper introducing a special issue of the *Neural Networks* journal devoted to these problems. One important group of applications of NNs in

* Corresponding author.
E-mail addresses: oceagah@univ.gda.pl (A. Herman), ralf.kaiser@nlwkn-ny.niedersachsen.de (R. Kaiser), hanz-dieter.niemeyer@nlwkn-ny.niedersachsen.de (H.D. Niemeyer).

0378-3839/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.coastaleng.2009.02.007

oceanic and atmospheric simulations are the so-called hybrid models, in which some—usually computationally most expensive—parts of the code are replaced with a trained NN, enabling accurate and in some cases a few orders of magnitude more efficient estimation of modelled quantities or processes. Krasnopolsky and Chevallier (2003) discuss examples of hybrid models, the most interesting from the point of view of the issues discussed here being the Neural Network Interaction Approximation (NNIA; see also Tolman et al., 2005), a NN-based algorithm designed to replace the Discrete Interaction Approximation (DIA) of the nonlinear four–four wave interactions in spectral wave models. Another class of approaches represents neural networks applied directly to predict the time series of wave parameters at the location of interest (usually nearshore) from the times series of other available data (usually off-shore), e.g. from satellite measurements (Kalra et al., 2005), wave buoy measurements (Tsai et al., 2002; Makarynskyy, 2005; Makarynskyy et al., 2005b), large-scale modelling results (Browne et al., 2006) or wind data (Deo et al., 2001; Rao and Mandal, 2005). The NN technique has also been used for forecasting future values of wave parameters at a given location based on the measured values of those parameters (see e.g. Deo and Naidu, 1999; Makarynskyy, 2004; Londhe and Panchang, 2006; Makarynskyy and Makarynska, 2007), for the interpolation of the missing values (Makarynskyy et al., 2005a), for studies concerning the interdependency of various wave parameters (Agrawal and Deo, 2004; Deo et al., 2002) and to improve the accuracy of the results obtained with a numerical wave model (Makarynskyy, 2005; Zhang et al., 2006).

Contrary to the works listed above, the approach used here, combining the data reduction and pattern recognition with NN techniques, enables to reproduce the mean wave parameters in the *whole study area*, not only at *single locations* as in the cited papers. The results presented here indicate that the linear principal component analysis (PCA; see e.g. Preisendorfer, 1988; Jolliffe, 2002) of the modelling results is a promising input-reduction tool e.g. for morphodynamic and other studies, for which the knowledge of the temporal and spatial variation of wave processes is a prerequisite. The ideas of combining PCA with NNs are similar to those used e.g. in van der Merwe et al. (2007) or Tolman et al. (2005). The same technique has been used successfully by Herman et al. (2007) to model the water levels and currents in the German Wadden Sea.

The structure of the paper is as follows: in Section 2 a brief description of the study area is given, followed in Section 3 by the description of the set-up of the third-generation spectral wave model SWAN (Simulating WAves Nearshore) used in the simulations. The verification of the results is presented as well. Section 4 discusses the results of the PCA of H_s, T_e and θ_m, with emphasis given to aspects crucial for the further stages of the project realization. The results of PCA, together with the time series of water level, wind velocity components and wave parameters in a chosen location, are then used to set up and train a feed-forward NN, as described in detail in Section 5. The trained NN is shown to accurately reproduce the test data and hence to be applicable as a tool for producing the data in periods that could not be modelled with SWAN because of time constraints. The role of wind and tidal processes in shaping the wave climate in the study area is analyzed in Section 6. Finally, Section 7 summarizes and discusses the results.

2. Area description

The main object of the study are the catchment areas of the tidal inlets between the islands of Borkum, Juist and Norderney, belonging to the chain of the East Frisian Islands separating the German Wadden Sea from the North Sea. The location of the study area, together with its bottom topography, is shown in Fig. 1. Fig. 2 shows a detailed view

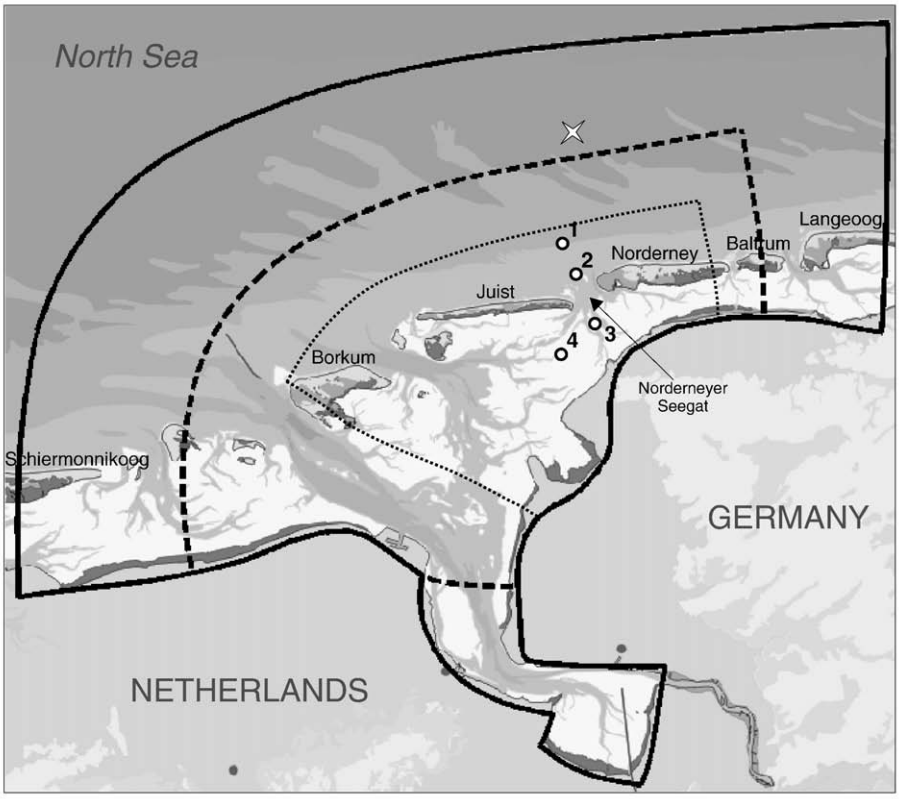

Fig. 1. Location of the study area on the German North Sea coast. Continuous line: boundary of the Delft3D grid; dashed line: boundary of the SWAN grid; dotted line: boundary of the area, in which the results are analyzed and modelled with NNs; numbered dots: locations in which the statistical distribution of wave parameters is analyzed (Section 5); a star: location of an input point for the wind atlas and the NN modelling (see text for detailed description).

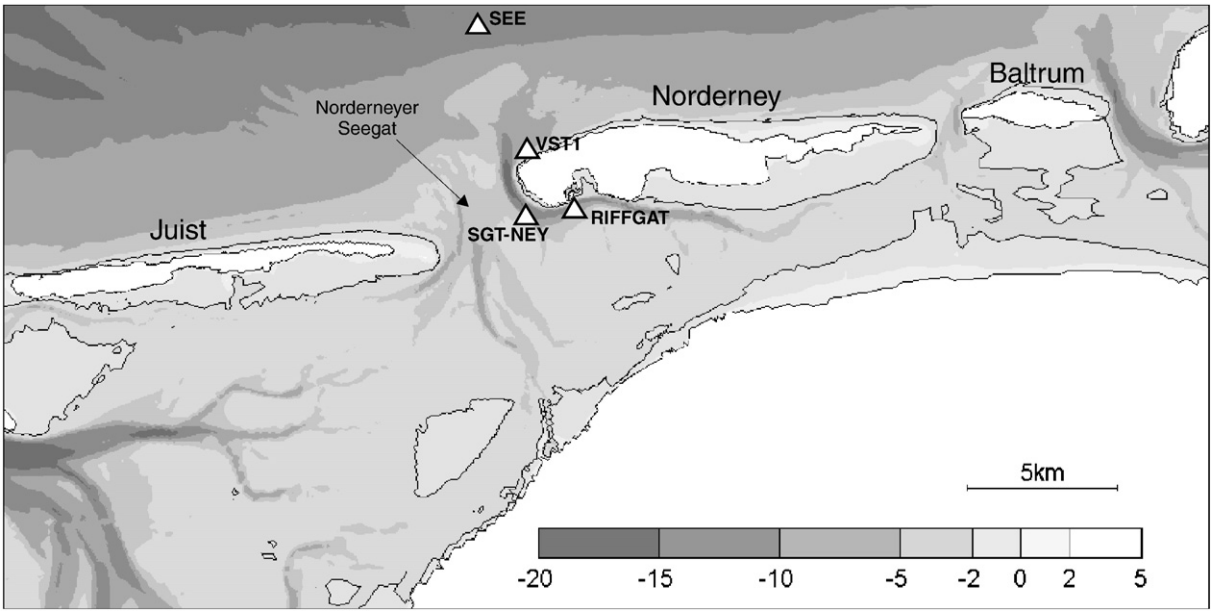

Fig. 2. Bathymetry (in meters relative to the German datum) of the tidal inlet Norderneyer Seegat and its catchment area and location of the measurement stations used for the verification of the modelling results.

of the tidal inlet Norderneyer Seegat, the main area of interest within the MOSES project. The mean tidal range in the study area (within the dashed line in Fig. 1) varies from 2.3 m to 2.7 m. According to the classification of Hayes (1979) the region is subject to upper mesotidal conditions. Details of the hydrodynamic processes in the study area were investigated recently by Herman (2007) and Herman et al. (2007).

The long-term mean H_s and T_e at the open boundary of the model equal 1.1–1.2 m and 5.5–5.6 s, respectively (values calculated from the data used as input in this study, see next section). Seawards from the islands the wave conditions are determined primarily by the larger-scale situation in the North Sea, which is in turn shaped by mesoscale atmospheric pressure and wind fields over the North Sea region. With mid-latitude low pressure systems typically traveling to the north of the German coast, during storm events the analyzed coast section experiences strong north-westerly and northerly winds, resulting in very rough sea states directly at the coast of the islands. Due to high water levels by storm events (storm surges of over 2.5 m were recorded in the harbour of Norderney; the surges at the mainland coast can be even 50 cm higher) the waves can penetrate through the tidal inlets into the Wadden Sea, where their interaction with the bottom is the dominating process in shaping the wave field. On the contrary, by typical calm or moderate weather conditions the local wind and water level (and thus the phase of the tidal cycle) play a dominant role.

3. Model set-up and verification

As mentioned in the introduction, the work presented here is part of a larger project, the main purpose of which was to produce a medium-term (1962–2002) database of water level, current and wave data for the area of interest. As described in detail in Herman et al. (2007) the hydrodynamic modelling was done by means of the two-dimensional version of the Delft3D model (Delft Hydraulics, 2003), set up on a curvilinear grid larger than the grid for the wave model (thick continuous line in Fig. 1). The wave modelling, which is the main subject of this paper, was performed with the parallel version 40.41AB of the third generation spectral wave model SWAN (Booij et al., 1999, 2004). The curvilinear orthogonal wave-model grid has spatial resolution varying from ~30 m in the tidal inlet between Juist and Norderney, called Norderneyer Seegat (the main area of interest in the MOSES project), to ~200 m at the north–west model boundary (see Fig. 1). The bottom topography was compiled from various data sources, collected during a number of measuring campaigns in different years. In Section 7 we discuss briefly the possible influence of using constant bathymetry on the modelling results.

In the nonstationary mode of SWAN there are two important parameters governing the time increment: the time step Δt and the maximal number of iterations per time step n_t. The tests performed for Δt in the range 5–30 min with 1, 2 or 3 iterations per time step showed that $\Delta t > 10$ min required large n_t values to properly propagate the changes at the open boundary through the model grid. On the other hand, $n_t > 2$ resulted in a sequence of stationary wave fields fitted to the instant conditions, which is unrealistic in situations with rapid changes of the forcing. Two sets of parameters gave highly accurate and comparable results: $\Delta t = 5$ min with $n_t = 1$ and $\Delta t = 10$ min with $n_t = 2$. For the reasons of computational economy the second version was used in the final model set-up (in over 40% of the time the model went over to the next time step after just one iteration). In the spectral domain 36 directions and 41 logarithmically distributed frequencies ranging from 0.04 Hz to 1.0 Hz were used, with an analytical diagnostic tail above the cut-off frequency. All source terms—wave generation by wind, dissipation due to whitecapping, bottom friction and depth-induced breaking, and nonlinear quadruplet and triad wave–wave interactions—were activated. Some details concerning the SWAN set-up are provided in Table 1. The choice of many of the model parameters was based on the earlier studies in the Norderneyer Seegat area (e.g. Kaiser and Niemeyer, 1999, 2001; Niemeyer and Kaiser, 2001).

Table 1
Some of the numerical parameters of the SWAN set-up (Booij et al., 1999, 2004).

Parameter	Value/formulation	Other values tested
Time step (min)	10	5, 20, 30
Max. No. of iterations	2	1,3
No. of frequencies	41	34
Frequency range (Hz)	0.04–1.00	0.04–0.6
No. of directions	36	–
Propagation scheme	BSBT	–

Table 2
SWAN performance at the four DWR measurement stations in the study area in autumn 2002.

	SEE	VST1	SGT-NEY	RIFFGAT
Buoy operation period	01.09–30.11	01.09–12.09, 17.09–09.10	01.10–30.11	01.10–30.11
No. of data points	1405	504	875	600
Mean water depth (m)	12.7	6.0	7.0	6.6
H_s (0.04–0.6 Hz)				
Mean measured value (m)	1.02	0.66	0.25	0.13
Mean modelled value (m)	1.07	0.55	0.29	0.18
Std. dev. of diff. (m)	0.58	0.32	0.15	0.11
Mean diff. (meas.-mod.; m)	0.01	0.11	−0.04	−0.05
T_e (0.04–0.6 Hz)				
Mean measured value (s)	5.53	5.76	3.46	3.22
Mean modelled value (s)	4.58	3.91	2.89	2.25
Std. dev. of diff. (s)	1.17	1.42	0.82	0.26
Mean diff. (meas.-mod.; s)	0.95	1.86	0.56	0.13

The model was driven by:

- 3-hourly time series of the 2D wave energy spectra from the database produced within the European Union Project HIPOCAS (Weisse et al., 2002)—results of the large-scale North Sea simulations with the WAM model (Günther et al., 1992); the spectra were interpolated onto the model boundary from a regular grid with spatial resolution 0.1°×0.05° (~5.5 km in each direction);
- hourly wind fields interpolated from the wind atlas of the German Weather Service (DWD);
- hourly water level and current fields calculated with Delft3D (Herman et al., 2007), linearly interpolated onto the SWAN grid.

The wind atlas, produced within MOSES and designed to fit its purposes, contains high resolution unidirectional wind fields for a number of wind speed, wind direction and water level classes. The wind fields were calculated by means of a wind model in which variations of the surface friction due to drying/flooding of the tidal flats and due to changes in the sea state (estimated by a simple statistical wave model) were taken into account. As an input for the wind atlas the time series of the HIPOCAS water level and wind speed and direction in location marked with a star in Fig. 1 were used.

The ability of SWAN run in a stationary mode to reproduce the transformation of wave energy in shallow tidal areas has been proved in a number of earlier studies, see e.g. Ris (1997), Ris et al. (1999), Kaiser and Niemeyer (1999, 2001), Niemeyer and Kaiser (2001). Here we test the ability of the model to accurately simulate both the temporal and spatial variability of waves in the study area subject to highly nonuniform, time-varying forcing. In a test period Sep.–Nov. 2002 the SWAN results are compared with measured H_s and T_e at four stations in the neighbourhood of the island of Norderney, shown in Fig. 2. At all stations the measurements were performed with the Datawell Directional Waverider (DWR) buoys in periods listed in Table 2. The raw data files, saved hourly, contain 1536 data samples each, collected during 20-minute periods of data acquisition. From those time series the wave energy spectra were calculated for the frequency range 0.025–0.58 Hz, with resolution $\Delta f = 0.005$ Hz within the 0.025–0.1 Hz range and $\Delta f = 0.01$ Hz within the 0.11–0.58 Hz range. The scatter plots of measured H_s and T_e at all four stations are shown in Fig. 3. One strong wind event took place in the analyzed period (end of October), with H_s up to 6 m at the station SEE and up to 1 m at the sheltered station RIFFGAT to the south of the island. Also, a calm period with low-amplitude swell coming from the North Sea took place (the second week of October), particularly visible at the station VST1. The remaining part of the three months analyzed was dominated by weak to moderate winds from various directions. At the station RIFFGAT numerous measurements with very low H_s and very high T_e values were made (Fig. 3d), corresponding always to low water independently of wind conditions and thus suggesting some disturbances in the recorded time series—or, at least, the manifestation of processes that cannot be accounted for in the SWAN model. Thus, the data points with measured H_s values lower than 0.1 m were excluded from the statistical analysis of the modelling results.

Examples of measured and modelled time series of H_s and T_e at the station SGT-NEY and RIFFGAT are shown in Fig. 4. In the case of the modelled values, two 'versions' of H_e and T_e are shown: calculated based on the energy from the whole range of frequencies (0.04–1.0 Hz) with the high-frequency tail taken into account and from the narrower range of frequencies cut at 0.6 Hz to allow the direct comparison with the measured data. The influence of the high-frequency energy on the H_s and T_e values at the stations SEE and VST1, located seawards from the island, is negligible (not shown), but at the sheltered stations SGT-NEY and RIFFGAT (Fig. 4) the differences are significant, especially in the case of the wave periods: the strong underestimation of periods—a known and often discussed problem in the SWAN model, see e.g. Ris (1997)—is largely reduced, especially in periods with low significant wave heights. The amplitudes of the tidal fluctuations of H_s and T_e calculated from the spectra with frequencies cut at 0.6 Hz correspond well to the measured amplitudes (e.g. days 37–43 in Fig. 4).

The above statements are not meant to suggest that the problem of underestimation of wave periods in SWAN, reported by many authors and related to the formulation of the source terms in the model, is not present in the area studied. The results simply call the attention to the importance of the usage of consistent spectral frequency ranges when calculating the mean wave parameters for purposes of the model evaluation, especially at shallow, sheltered locations. That the underestimation of wave periods still persists if the frequency range

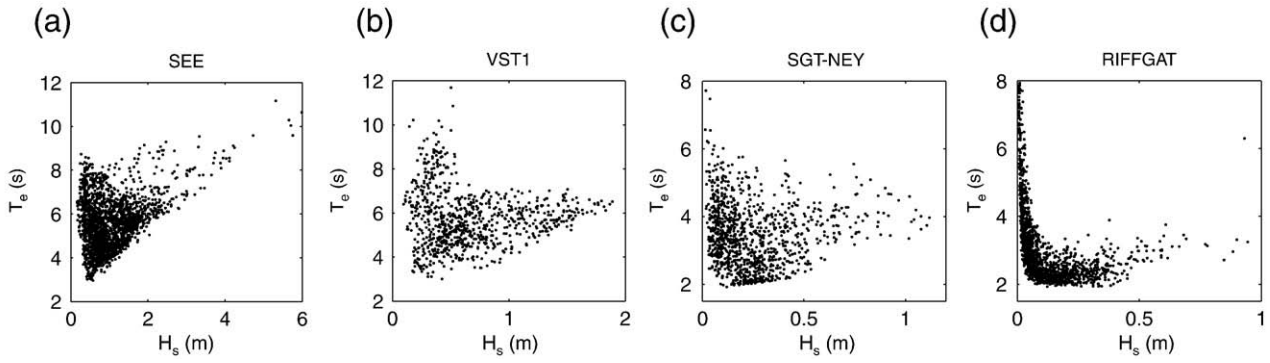

Fig. 3. Relationships between measured H_s and T_e at the four analyzed stations in Sep.–Nov. 2002.

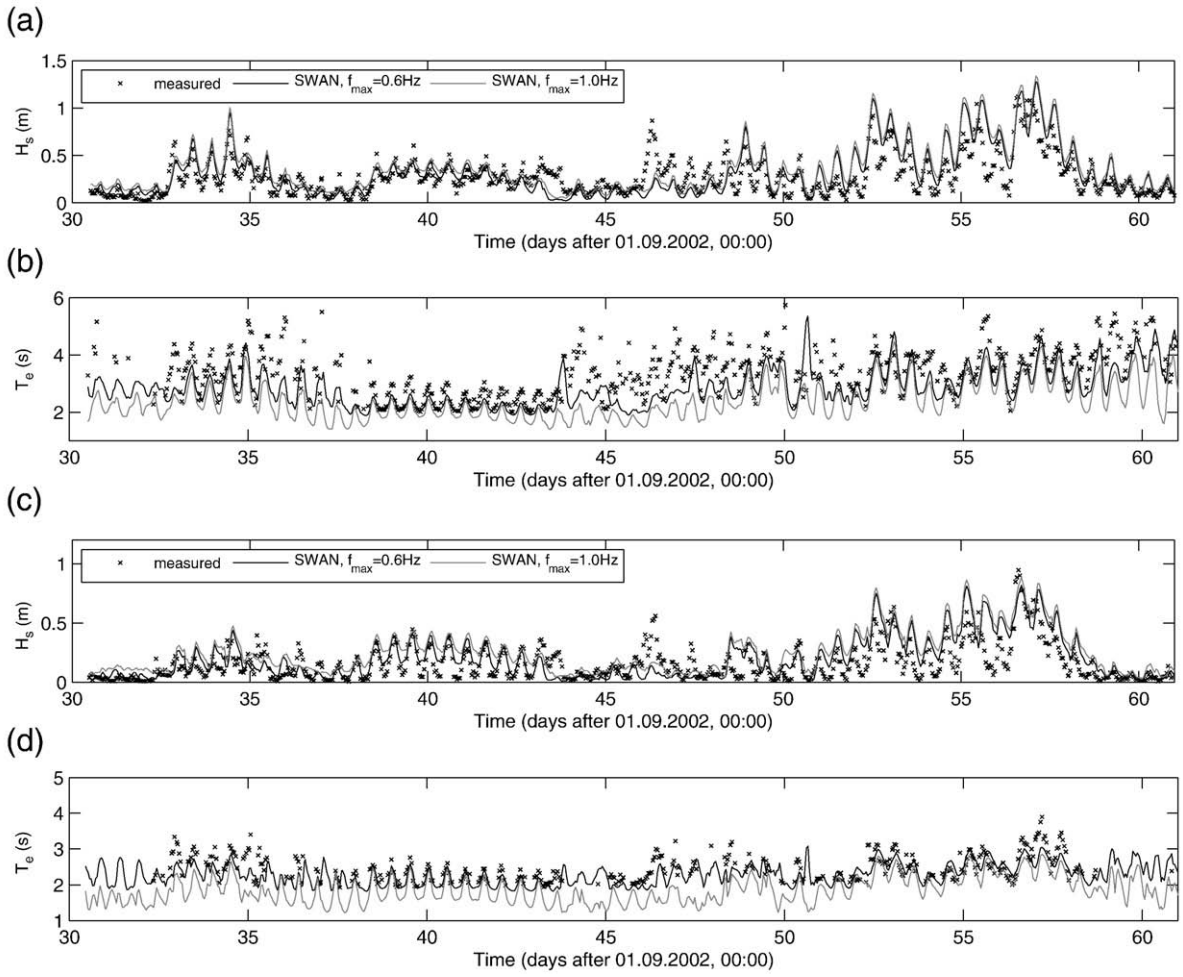

Fig. 4. Time series of measured and modelled H_s (a,c) and T_e (b,d) at the station SGT-NEY (a,b) and RIFFGAT (c,d) in October 2002.

0.04–0.6 Hz is considered can be seen from Table 2, summarizing the overall performance of SWAN at the four analyzed stations. The small values of bias of the significant wave heights at the stations SGT-NEY and RIFFGAT suggest that the amount of wave energy modelled by SWAN is consistent with measurements. The bias of energy periods shows that this energy is differently distributed among the frequencies: SWAN tends to underestimate the low frequency energy and to overestimate the high frequency energy.

The overall conclusion from the data presented in Fig. 4 and in Table 2 is that the modelling system built within the MOSES project—the configured SWAN model together with the input data—is capable of reliably reproducing the wave variability in the study area. The accuracy of the results lies well within the accuracy achievable with the state-of-the-art spectral wave models. Also, it must be remembered that what is important for the present study is the performance of the PCA/NN technique *given* the wave model results, as described further.

In the following the results from the first two years of the analyzed period (1962–1963) are used to develop a statistical model (based on PCA and NN) of the wave processes in the study area—similarly as in Herman et al. (2007) for water levels and currents. The performance of the trained NNs is then tested by comparing their predictions with the SWAN results from the year 1985, chosen based on the analysis of the wind and water level data: the distributions of the wind speeds and water levels at the model boundary in the year 1985 were very similar to those obtained for the whole analysis period 1962–2002.

4. PCA of the mean wave parameters

4.1. General remarks

The modelled fields of H_s, T_e and θ_m from the first two years of the period of study (1962–1963), saved hourly, were analyzed inside the area marked with a dotted line in Fig. 1. Every 7th grid point in each direction was taken into account, resulting in $N_p \times N_t$-point data matrices of H_s, T_e and θ_m, where $N_p = 3463$ is the number of data points (only those points, which remained wet during at least 10% of the simulation time were taken into account) and $N_t = 17521$ is the number of time points (hourly from 01.01.1962, 00:00 to 01.01.1964, 00:00). The reason for the reduction of the spatial resolution of the datasets prior to the PCA was pragmatic—to obtain manageable matrix sizes—but justified by an observation that the mesh resolution necessary to properly account for the influence of the bottom topography on the wave propagation in SWAN was much higher than the desired resolution of the data produced within MOSES. The sensitiveness of the PCA results to changes of the data resolution was thoroughly tested; no significant influence neither on the spatial structure of the leading modes nor on the percentage of the total variance they describe (changes <1%) could be noticed. Of course, in general it is the planned further application of the data that decides whether the reduction of the number of points is necessary and allowable.

The mean wave propagation directions were analyzed as vectors $\mathbf{r}_m$ of unit length pointing in the direction θ_m: $\mathbf{r}_m = \cos\theta_m + i \sin\theta_m$, where $i = \sqrt{-1}$ is the imaginary unit. However, because the quantity of

interest in the further analysis were not the vectors $\mathbf{r}_m$ as a whole, but their direction θ_m, all statistical parameters describing the modelling performance were calculated for θ_m, not for $\mathbf{r}_m$. All exception values in the analyzed datasets, corresponding to negative water depths, have been replaced with zeros. The following naming convention is used in this work: the time-invariant spatial patterns, denoted with $\mathcal{M}_{H_s,n}(\mathbf{x})$, $\mathcal{M}_{T_e,n}(\mathbf{x})$, $\mathcal{M}_{\theta_m,n}(\mathbf{x})$, are termed 'modes'. The time series describing their time evolution, denoted with $\mathcal{P}_{H_s,n}(\mathbf{t})$, $\mathcal{P}_{T_e,n}(\mathbf{t})$, $\mathcal{P}_{\theta_m,n}(\mathbf{t})$, are called 'principal components' (PCs). In the notation above $\mathbf{x}$ is a vector of locations of length N_p, $\mathbf{t}$ is a time vector of length N_t and $n=1,\ldots,N_p$.

4.2. Basic properties of the modes

The percentage of variance described by the ten most energetic mode/PC pairs of H_s, T_e and θ_m is shown in Table 3. A striking feature of the PCA results is the high amount of variance contained in the leading modes of T_e, θ_m and, especially, H_s. The spatial structure of these modes is shown in Fig. 5. The first H_s mode, $\mathcal{M}_{H_s,1}$ (Fig. 5a) reflects an 'average' wave height distribution in the study area, with highest values along the open boundary, gradually decreasing towards the shore and further to the tidal flats. A characteristic feature (present in $\mathcal{M}_{H_s,2}$ as well, see Fig. 5b) is that the contours over the ebb delta of the Norderneyer Seegat—a region of the highest H_s gradients—correspond very well to the contours of the bottom topography (Fig. 2). The distribution of variance described by the first two leading modes of H_s is not spatially uniform (not shown), e.g. $\mathcal{M}_{H_s,1}$ accounts for over 97% of the variance seawards from Juist and Norderney, 60–80% in the deep tidal channels and less than 50% over the tidal flats. To the contrary, $\mathcal{M}_{H_s,2}$, 'responsible' for 3.35% of the total variance, describes over 40% of the variance at shallow locations. The behaviour of the modes of T_e and θ_m is similar: $\mathcal{M}_{T_e,1}$ and $\mathcal{M}_{\theta_m,1}$ account primarily for the variance to the north of the islands, while the further modes have their centers of action at the shallow locations in the Wadden Sea. Also, the spatial small-scale variability of the modes increases substantially with increasing n. A detailed analysis of $\mathcal{P}_{H_s,n}$, $\mathcal{P}_{T_e,n}$ and $\mathcal{P}_{\theta_m,n}$ will be performed further in Section 6, together with the discussion on the role of wind and tides upon wave variability in the study area.

4.3. PCA reconstruction of the SWAN data

In all three cases (H_s, T_e, θ_m) $N=5$ first mode/PC-pairs were used in the further analysis and in the NN modelling. Although the amount of variance reproduced with five leading PCs is different in the case of H_s (98.9%), T_e (92.6%) and θ_m (85.9%), the inclusion of the further components led in neither of the three cases to a noticeable improvement of the performance of the NN modelling: a very irregular time variation of the further PCs, which are significant only locally (almost exclusively in locations falling dry at low water), makes the training of neural networks prohibitively difficult and in some cases leads even to the deterioration of the quality of the results.

Table 3
Percentage of total variance described by the first ten linear modes of H_s, T_e and θ_m.

Mode	H_s		T_e		θ_m	
	Separate	Cumulative	Separate	Cumulative	Separate	Cumulative
1.	93.96	93.96	77.81	77.81	70.59	70.59
2.	3.35	97.30	10.92	88.73	9.73	80.32
3.	0.87	98.17	1.51	90.24	2.55	82.87
4.	0.47	98.64	1.28	91.52	1.95	84.83
5.	0.26	98.90	1.05	92.57	1.05	85.87
6.	0.18	99.08	0.70	93.27	0.85	86.72
7.	0.16	99.25	0.57	93.84	0.69	87.41
8.	0.09	99.34	0.45	94.29	0.58	87.99
9.	0.08	99.42	0.41	94.70	0.52	88.51
10.	0.07	99.48	0.34	95.04	0.47	88.98

The spatial distribution of the percentage of variance δ_{var} reconstructed with the five leading PCs and of the standard deviation of differences σ_{std} between the original and the reconstructed data is shown in Fig. 6. Not surprisingly the quality of the reconstructed data is highest in the region seawards from the islands, where the variability of the wave conditions is relatively small. The δ_{var} values are very high especially in the area surrounding the ebb delta of the Norderneyer Seegat (where the concentration of data points was highest; hence the high 'weight' of those points in the PCA results): over 95% for all three analyzed quantities, with σ_{std} in the range 4–5 cm for H_s (Fig. 6b), 0.2–0.3 s for T_e (Fig. 6d) and ~20° for θ_m (Fig. 6f). Over the shallows south from the islands the picture is more complicated, with small amplitude fluctuations of δ_{var} and σ_{std}. Generally, the quality of the reconstructed data is worst over the tidal flats and in points located directly at the coast—to a high degree due to long periods of zero values in the time series, which for obvious reasons cannot be approximated accurately with periodically time-varying PCs.

5. Neural network modelling

5.1. Configuration of the NNs

The NN architecture used here is the multi-layer perceptron (MLP), consisting of two layers of weights providing full connection from the input units to the hidden units and from the hidden units to the output units. Thus, it is a so-called two-layer feed-forward network with error gradient backpropagation as a learning procedure. Basics of NNs can be found in a number of textbooks and will not be discussed here. The same type of neural networks were used e.g. by Tsai et al. (2002), Makarynskyy (2004), Kalra et al. (2005), Makarynskyy et al. (2005b) and many others. All simulations presented further were performed with the NETLAB package of MATLAB scripts developed by Nabney (2004). Some of the most important parameters of the MLP used in this study are: hyperbolic-tangent activation function of the hidden layer, linear activation function of the output layer (most appropriate choice for nonlinear regression problems) and a conjugate gradient optimization algorithm, see Nabney (2004) for details. Other optimization algorithms (scaled conjugate gradient and quasi-Newton), as well as other network architectures (three-layer MLPs with and without direct connections from the input layer to the output layer and the Elman networks) were tested, but gave worse results and/or made the training difficult and time consuming. They are therefore not considered further.

For each of the parameters H_s, T_e and θ_m an ensemble of 50 neural networks with a different number N_i and different combinations of input neurons was set up. The initial 'full' set of input neurons considered was:

$$\left[H_{s,\mathrm{HIP}}(\tilde{\mathbf{t}}), T_{e,\mathrm{HIP}}(\tilde{\mathbf{t}}), \theta_{m,\mathrm{HIP}}(\tilde{\mathbf{t}}), \xi_{\mathrm{HIP}}(\tilde{\mathbf{t}}), u_w(\tilde{\mathbf{t}}), v_w(\tilde{\mathbf{t}}), \overline{u}_w(t), \overline{v}_w(t)\right],$$

where $H_{s,\mathrm{HIP}}$, $T_{e,\mathrm{HIP}}$, $\theta_{m,\mathrm{HIP}}$ and ξ_{HIP} denote the HIPOCAS significant wave height, energy period, mean wave direction and water level in the input point (a star in Fig. 1), t denotes time, $\tilde{\mathbf{t}}=[t-3\Delta t, t-2\Delta t, t-\Delta t, t]$, $\Delta t=1$ h, $[u_w, v_w]$ is the HIPOCAS wind velocity vector in the input point and $[\overline{u}_w, \overline{v}_w]$ is the average of $[u_w, v_w]$ over $\tilde{\mathbf{t}}$. Thus, a maximal value of N_i considered was 26. As a selection criterion for the final network from each ensemble the global statistics were examined, i.e. the correlation coefficients r and the standard deviation of differences σ_{std} obtained for the test data (see below). Each network was assigned ranks in terms of its r and σ_{std} values. It turned out that among the ten best networks from each ensemble their r and σ_{std} ranks did never differ by more than two positions. An inspection of the Hinton diagrams of these best networks, together with a pairwise comparison of the results obtained with networks differing by one input neuron, revealed that some input parameters (e.g. $u_w(t-3\Delta t)$,

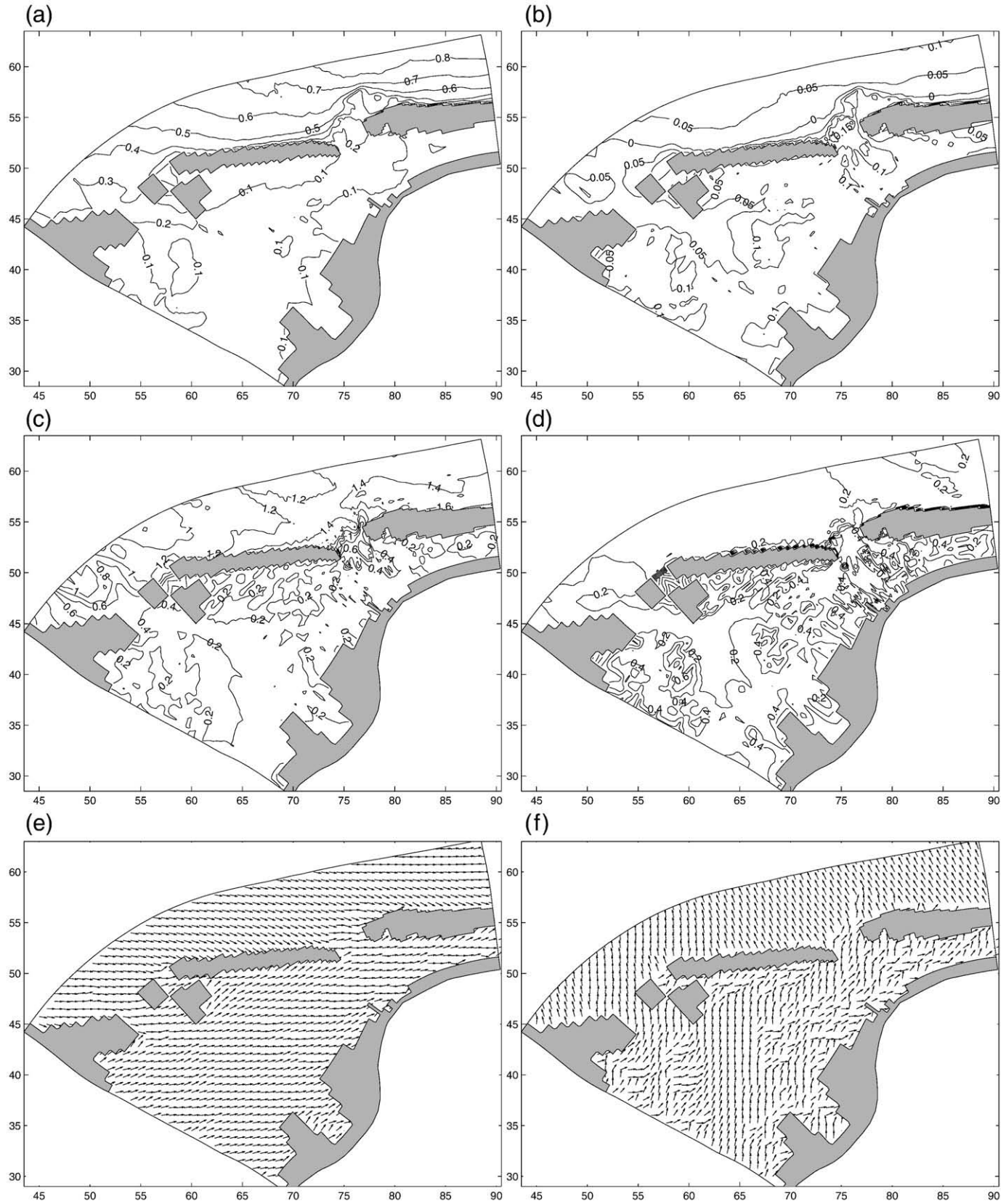

Fig. 5. The PCA results of H_s, T_e and θ_m: $\mathcal{M}_{H_s,1}$ (a), $\mathcal{M}_{H_s,2}$ (b), $\mathcal{M}_{T_e,1}$ (c), $\mathcal{M}_{T_e,2}$ (d), $\mathcal{M}_{\theta_m,1}$ (e) and $\mathcal{M}_{\theta_m,2}$ (f).

$v_w(t-3\Delta t)$) had almost no influence on the network performance. Elimination of networks containing these parameters as input resulted in 4, 7 and 5 remaining ensemble members for H_s, T_e and θ_m, respectively. The relative performance differences of these NNs (within each ensemble) in terms of r lied within 0.2%, and could be classified as insignificant. Hence, it was decided to choose the final

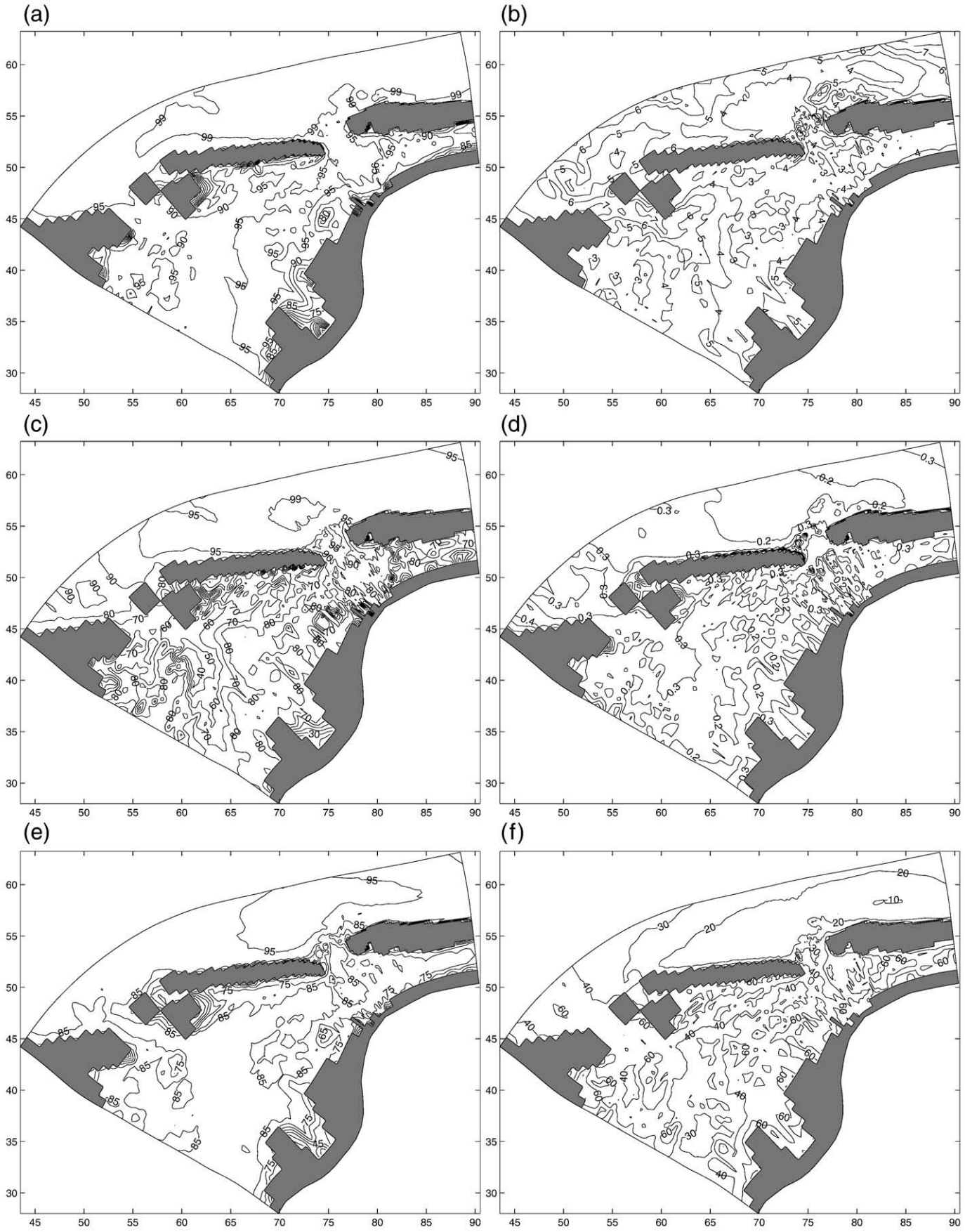

Fig. 6. Quality of the PCA reconstruction of H_s (a,b), T_e (c,d) and θ_m (e,f) with five leading PC/mode pairs: spatial distribution of the percentage of the reconstructed variance δ_{var} (a,c, e) and of the standard deviation of differences σ_{std} (b: in cm; d: in seconds; f: in degrees).

Table 4
Statistical parameters describing the NN performance for the H_s, T_e and θ_m test datasets Jul.–Dec.1963.

	$PC_{H_s,1}$	$PC_{H_s,2}$	$PC_{H_s,3}$	$PC_{H_s,4}$	$PC_{H_s,5}$
Corr. coef. (%)	99.49	98.97	96.37	88.10	95.01
σ_{std} (–)	0.111	0.141	0.266	0.461	0.299
Mean difference (–)	0.007	0.000	0.001	0.025	0.006
	$PC_{T_e,1}$	$PC_{T_e,2}$	$PC_{T_e,3}$	$PC_{T_e,4}$	$PC_{T_e,5}$
Corr. coef. (%)	98.04	98.94	93.50	98.66	95.55
σ_{std} (–)	0.200	0.147	0.371	0.164	0.305
Mean difference (–)	−0.020	0.014	−0.034	−0.007	0.029
	$\theta(PC_{r_m,1})$	$\theta(PC_{r_m,2})$	$\theta(PC_{r_m,3})$	$\theta(PC_{r_m,4})$	$\theta(PC_{r_m,5})$
Corr. coef. (%)	99.35	98.23	97.78	98.15	96.38
σ_{std} (degr)	14.02	20.11	22.52	20.55	29.59
Mean difference (degr)	0.18	0.92	0.22	0.97	0.28

networks so that the differences in input vectors between the ensembles were minimal. Thus, the input vectors were set to:

$$\left[H_{s,HIP}(\tilde{\mathbf{t}}), T_{e,HIP}(t), \theta_{m,HIP}(t), \xi_{HIP}(\tilde{\mathbf{t}}), u_w(t), v_w(t), \overline{u}_w(t), \overline{v}_w(t)\right],$$

$$\left[H_{s,HIP}(t), T_{e,HIP}(\tilde{\mathbf{t}}), \theta_{m,HIP}(t), \xi_{HIP}(\tilde{\mathbf{t}}), u_w(t), v_w(t), \overline{u}_w(t), \overline{v}_w(t)\right]$$

and:

$$\left[H_{s,HIP}(t), T_{e,HIP}(t), \theta_{m,HIP}(\tilde{\mathbf{t}}), \xi_{HIP}(\tilde{\mathbf{t}}), u_w(t), v_w(t), \overline{u}_w(t), \overline{v}_w(t)\right]$$

for the H_s, T_e and θ_m network, respectively, resulting in $N_i = 14$. It is a noteworthy fact that all three wave parameters can be modelled with similar sets of input variables.

The necessary number of hidden units N_h was established by means of sensitivity tests with N_h varied from 50 to 150 with a step of 10. In the range 50–80 a strong performance improvement (again in terms of r and σ_{std} values) could be noticed; for $N_h \geq 90$ the relative differences of r and σ_{std} remained within the 0.3% and 0.6%, respectively. Thus, $N_h = 90$ was finally chosen. The output of the networks was $\left[\mathcal{P}_{H_s,n}(\mathbf{t})\right]_{n=1,\ldots,5}$, $\left[\mathcal{P}_{T_e,n}(\mathbf{t})\right]_{n=1,\ldots,5}$ and $\left[\mathcal{P}_{\theta_m,n}(\mathbf{t})\right]_{n=1,\ldots,5}$, respectively, resulting in $N_o = 5$ of output units in the case of H_s and T_e and $N_o = 10$ in the case of θ_m. Similarly as in the case of the water levels and currents (Herman et al., 2007) the dataset 1962–1963 was divided into the training data (Jan. 1962–Jun. 1963) and the validation data (Jul.–Dec. 1963).

5.2. Reconstruction of the PCs

The first two PCs of H_s, describing over 97.3% of the total variance (Table 3), are reconstructed very precisely with the NN, with the correlation coefficient r of 99.49% and 98.97% and with σ_{std} of 0.11 and 0.14, respectively (Table 4). The scatter plots for $\left[\mathcal{P}_{H_s,n}(\mathbf{t})\right]_{n=1,\ldots,5}$ are shown in Fig. 7a. The quality of the reconstruction is worst in the case of $\mathcal{P}_{H_s,4}$ (correlation coefficient of 88.10% and σ_{std} of 0.46), especially in situations in which this PC reaches high values. It must be stressed, however, that $\mathcal{P}_{H_s,3} - \mathcal{P}_{H_s,5}$ together describe only 1.6% of the total variance, so that their reconstruction errors have a significantly smaller influence on the overall quality of the NN modelling—a notion valid for the PCs of T_e and θ_m as well.

If the two most energetic PCs are concerned, the results of the NN modelling are worst in the case of T_e (Fig. 7b): although r exceeds 98% (98.04% for $\mathcal{P}_{T_e,1}$ and 98.94% for $\mathcal{P}_{T_e,2}$, see Table 4), the scatter of the

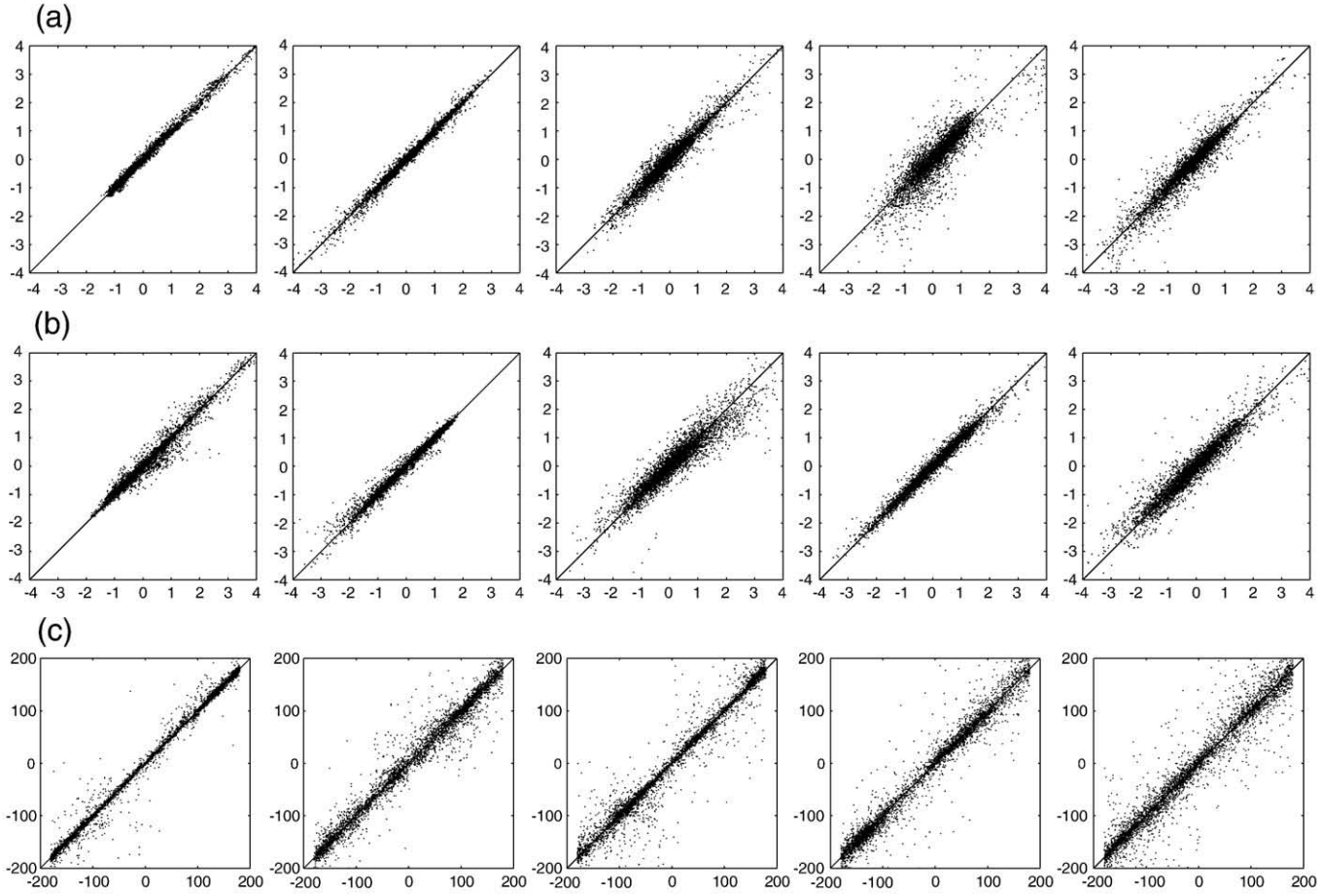

Fig. 7. Scatter plots of the original (horizontal axis) and the NN-reconstructed (vertical axis) five leading PCs of H_s (a), T), T_e (b) and θ_m (c) in the test period Jul.–Dec. 1963.

Table 5
NN modelling performance in points 1–4 (Fig. 1) in the year 1985.

Point	1	2	3	4
H_s				
Corr. coefficient (%)	98.9	98.1	96.5	96.4
σ_{std} (m)	0.11	0.09	0.05	0.04
Mean diff. (SWAN-NN; m)	0.04	0.06	0.00	0.00
T_e				
Corr. coefficient (%)	93.3	93.0	84.8	72.5
σ_{std} (s)	0.47	0.59	0.32	0.28
Mean diff. (SWAN-NN; s)	0.18	0.35	0.06	−0.01
θ_m				
Corr. coefficient (%)	80.9	80.0	81.5	86.3
σ_{std} (degr)	41.3	42.1	57.1	46.4
Mean diff. (SWAN-NN; degr)	2.6	0.7	−5.8	−2.6

points is relatively strong, with the values of σ_{std} 0.20 and 0.15, respectively.

Obviously in the case of θ_m, the NN performance has been evaluated not for the $\mathcal{P}_{\theta_m,n}$ vectors directly, but for their directions $\theta(\mathcal{P}_{\theta_m,n})$ measured in degrees, as shown in Table 4 and Fig. 7c.

5.3. Reconstruction of the data from the year 1985

The performance of the NN modelling of the mean wave parameters in the study area can be evaluated in many different ways, depending on the further application of the modelling results. From the point of view of the goals of the MOSES project, one of the most important aspects concerns the ability of the NN model to accurately reproduce those statistical properties of the wave field, which are particularly important for morphodynamic studies and safety assessment of the coastal zone, e.g. the distribution curves of H_s, T_e and θ_m in various parts of the study area. The discrete probability density of these parameters was calculated for a number of locations both in the Wadden Sea and in the open sea region. The results for four chosen points, representing the whole range of wave conditions and marked with numbered dots in Fig. 1, are shown in the Supplementary Figs. 1 and 2. Table 5 summarizes the basic statistical parameters for H_s, T_e and θ_m at all four locations.

Although the shape of the probability density curves varies strongly from point to point, the NN was able to reproduce this variability with a high accuracy, not only in the points located seawards from the islands, but in the sheltered, dry-falling points (locations 3 and 4) as well. In the case of the energy periods, a small

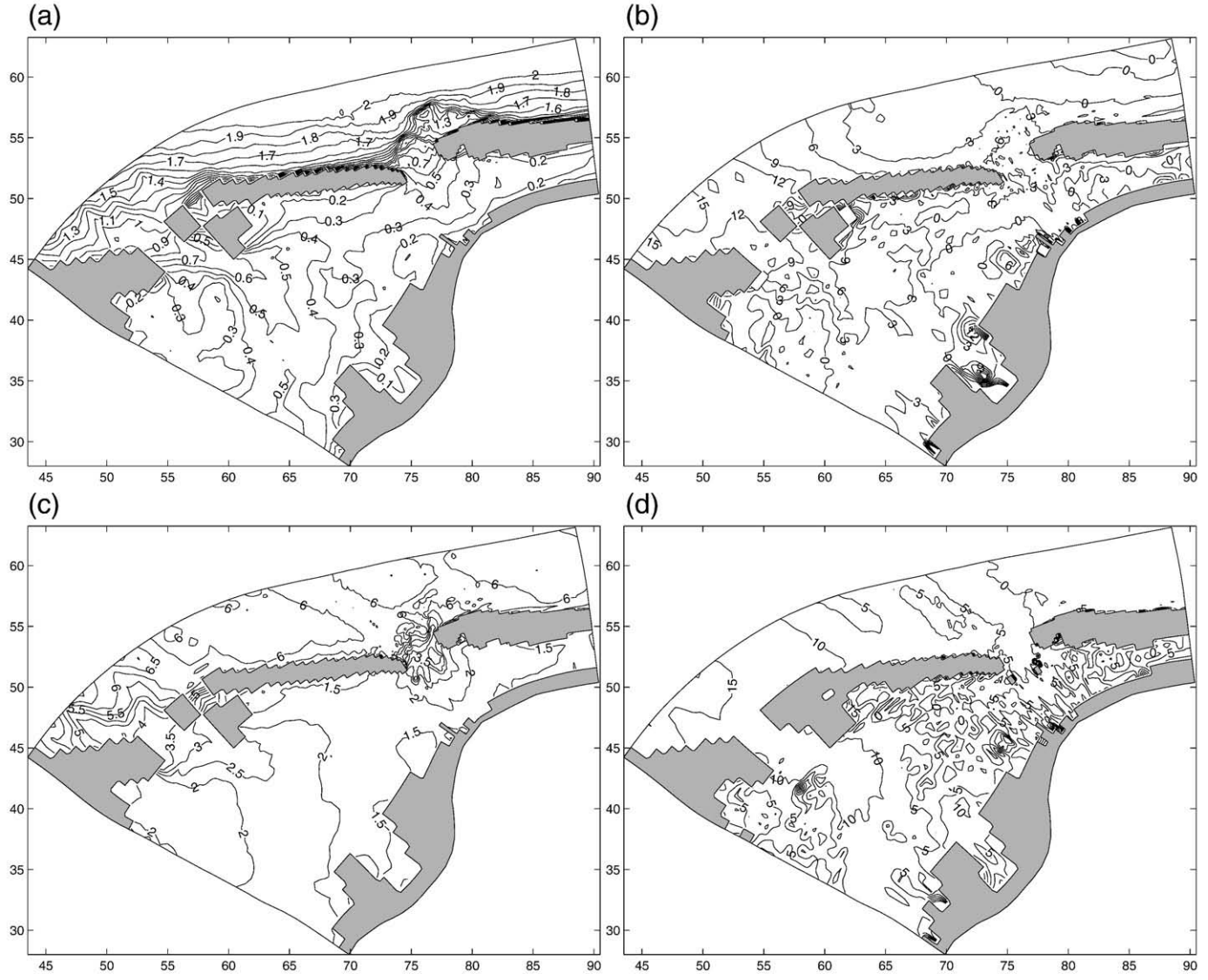

Fig. 8. Spatial distribution of the 90th percentile p_{90} of H_s (a; in m) and T_e (c; in s) calculated from the SWAN results from the year 1985. Right panels show the relative difference (in %) between p_{90} calculated based on the SWAN- and the NN-modelled H_s (b) and T_e (d).

bias in the NN results can be observed, especially at the location 2 (Supplementary Fig. 1f), which lies in the region of very strong T_e gradients at the edge of the ebb delta of the Norderneyer Seegat (Fig. 1).

The quality of the NN results—measured in terms of the ability to reproduce the statistical distributions of the analyzed parameters—is worst for the wave directions (Table 5). At locations seawards from the islands (points 1 and 2), where the wave directions concentrate around a single dominating one, the NN is able to reproduce this pattern, but gives too broad distributions (Supplementary Fig. 2a,b). In the Wadden Sea, where no dominating direction can be identified and the distribution of the wave propagating directions strongly resembles the distribution of wind directions (an effect of a local wind-wave generation), the NN results agree only approximately with the SWAN results (Supplementary Fig. 2c,d).

Finally, as another measure of the quality of the NN modelling, the spatial distribution of the 90th percentile of H_s and T_e, p_{90,H_s} and p_{90,T_e}, obtained from the SWAN and the NN results, are compared in Fig. 8. Although the NN tends to slightly underestimate both p_{90,H_s} and p_{90,T_e} values (positive relative differences in Fig. 8b,d), especially in the north-western part of the study area and along its northern boundary, the agreement between the results of the SWAN and the NN modelling is very good, especially in the main area of interest—the Norderneyer Seegat and its surroundings.

6. Wind-wave variability in the study area

The wind-wave variability in the study area is strongly dependent on the wind and the tides (i.e. changing water levels and currents) governing the local wave generation and propagation as well as the wave conditions at the open boundary of the region. These phenomena were studied thoroughly in (Herman et al., 2007); here we repeat the most important conclusions from that study. If the HIPOCAS hourly data is concerned, no statistically significant trends are present neither in the water level nor in the wind data. However, an increasing occurrence frequency of strong wind events in the cold time of the year can be noticed, manifesting itself in positive winter trends in seasonally averaged wind speeds. A very interesting feature is the existence of a 'dominant' wind-direction axis (approx. WNW–ESE) governing the non-tidal part of the water level and current variability in the study area. In the case of the wind-wave parameters this dominance is less pronounced (although present), but for the sake of consistency we are going to express the wind velocity vector components in the same way, as $[u_w, v_w]$, with u_w along the WNW–ESE axis and v_w in the perpendicular direction.

Before we proceed to a more detailed analysis it is useful to examine the basic properties of the PCs. Fragments of $\mathcal{P}_{H_s,n}$, $\mathcal{P}_{T_e,n}$ and $\mathcal{P}_{\theta_m,n}$ (for $n=1,2,3$) are shown in Fig. 9, together with the corresponding time series of the HIPOCAS data from the input point. As can be seen, the first PCs closely follow the HIPOCAS time series. Although some tidal modulation is present, especially in $\mathcal{P}_{\theta_m,1}$, they reflect mainly the general wave conditions, e.g. the magnitude of $\mathcal{P}_{H_s,1}$ and $\mathcal{P}_{T_e,1}$ is highest during strong wind events (e.g. hours 110–160 in Fig. 9a,b)—which is not surprising if one considers that the spatial structure of $\mathcal{M}_{H_s,1}$ and $\mathcal{M}_{T_e,1}$ (Fig. 5a,c) closely resembles the H_s and T_e distribution at high wind speeds. To the contrary, the second PCs are 'responsible' mainly for the tidal modulation of the mean wave parameters. Interestingly, the variance of $\mathcal{P}_{T_e,2}$ is lower in periods of high amplitude of $\mathcal{P}_{T_e,1}$, indicating the dominance of wind over the tidal processes.

To get a further insight into the properties of the PCs, and thus to investigate the relationships between the wind, the tidal forcing and the wave conditions, we follow a similar approach as the one used by (Herman et al., 2007). Thus, by means of the T_TIDE package of

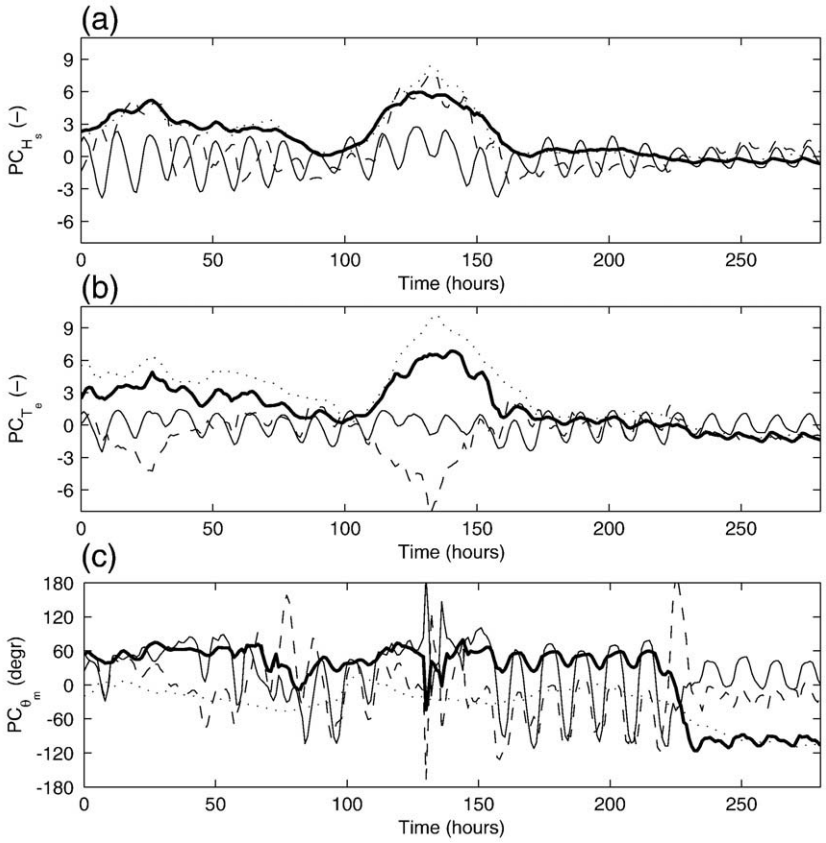

Fig. 9. Fragments of the three leading PCs of H_s (a), T), T_e (b) and θ_m (c). Thick continuous line: first PC, thin continuous line: second PC, thin dashed line: third PC. The HIPOCAS time series of H_s, T_e and θ_m, with the time means removed, are shown with dotted curves. In (c) the directions of $\mathcal{P}_{\theta_m,n}$ ($n=1,2,3$) instead of the PCs themselves are shown.

Table 6
Variance of $\mathcal{P}_{H_s,n,t}$, $\mathcal{P}_{T_e,n,t}$ and $\mathcal{P}_{\theta_m,n,t}$ in percent of $\mathcal{P}_{H_s,n}$, $\mathcal{P}_{T_e,n}$ and $\mathcal{P}_{\theta_m,n}$, respectively.

	PC1	PC2	PC3	PC4	PC5	HIPOCAS
H_s	0.4	69.6	8.3	4.8	32.9	5.0
T_e	2.1	81.7	12.3	31.6	24.3	5.8
θ_m	1.4	57.0	17.6	37.7	16.3	3.6

In the last column the percentage of variance of the tidal part of the HIPOCAS time series is shown.

Pawlowicz et al. (2002) each of the PCs of H_s, T_e and θ_m is decomposed into a tidal and a residual part:

$$\mathcal{P}_{H_s,n} = \mathcal{P}_{H_s,n,t} + \mathcal{P}_{H_s,n,r}, \quad \mathcal{P}_{T_e,n} = \mathcal{P}_{T_e,n,t} + \mathcal{P}_{T_e,n,r}, \quad \mathcal{P}_{\theta_m,n} = \mathcal{P}_{\theta_m,n,t} + \mathcal{P}_{\theta_m,n,r}$$

for $n = 1,\ldots, 5$. In the harmonic analysis of the PCs the 25 most energetic tidal components were used. In the same way the decomposition of the HIPOCAS time series from the input point is performed, resulting in:

$$H_{s,HIP} = H_{s,HIP,t} + H_{s,HIP,r}, \quad T_{e,HIP} = T_{e,HIP,t} + T_{e,HIP,r},$$

and so on. The percentage of variance of the tidal components of the analyzed series (in % of the original series) is listed in Table 6. The variance distribution confirms the general conclusions drawn based on Fig. 9. In view of these properties of the modes and PCs it is not surprising that certain relationships between the considered quantities exist, as shown in Fig. 10. In the case of $\mathcal{P}_{H_s,1}$ and $\mathcal{P}_{T_e,1}$, a strong dependence on $H_{s,HIP}$ and $T_{e,HIP}$ is present, although its exact form depends on the sign of the wind component u_w (black and gray crosses in Fig. 10a,b). The values of $\mathcal{P}_{\theta_m,1}$ are concentrated around two bins, corresponding to the negative and positive u_w and clearly defined ranges of the HIPOCAS mean wave directions (Fig. 10c). If the second PCs are considered, their 'tidal' character is responsible for a strong (and almost linear, with an exception of the highest values) correlation with ξ_{HIP} (Fig. 10d–f).

Finally, it is worth noticing that, although no trends in the hourly PCs can be identified in the 40 years studied, the seasonally averaged time series reveal an analogous variability to the one present in the wind data, as discussed in Herman et al. (2007): the wintertime $\mathcal{P}_{H_s,1}$ and $\mathcal{P}_{T_e,1}$ exhibit a strong positive trend, contrary to the insignificant trend in the summertime $\mathcal{P}_{H_s,1}$ and $\mathcal{P}_{T_e,1}$ (Fig. 11). Similar positive trends can be found in wintertime maximum values—an important phenomenon e.g. from the point of view of the safety of the coastal protection structures in the study area. A more detailed analysis of these issues, concentrated on the extreme events, their course and occurrence frequency, will be presented elsewhere.

7. Discussion

The combination of the third-generation spectral wind-wave modelling with the NN/PCA technique, described in this paper, has been shown to provide an effective alternative to the 'classical' modelling approaches. An important aspect of this technique is its applicability to other similar regions. First tests performed for other areas at the German coast (unpublished) suggest that e.g. the same neural network structure and the same sets of input parameters are suitable for the NN simulation there—provided that the size of the modelled area is small enough so that the meteorological and tidal forcing is highly spatially correlated and can be represented by time

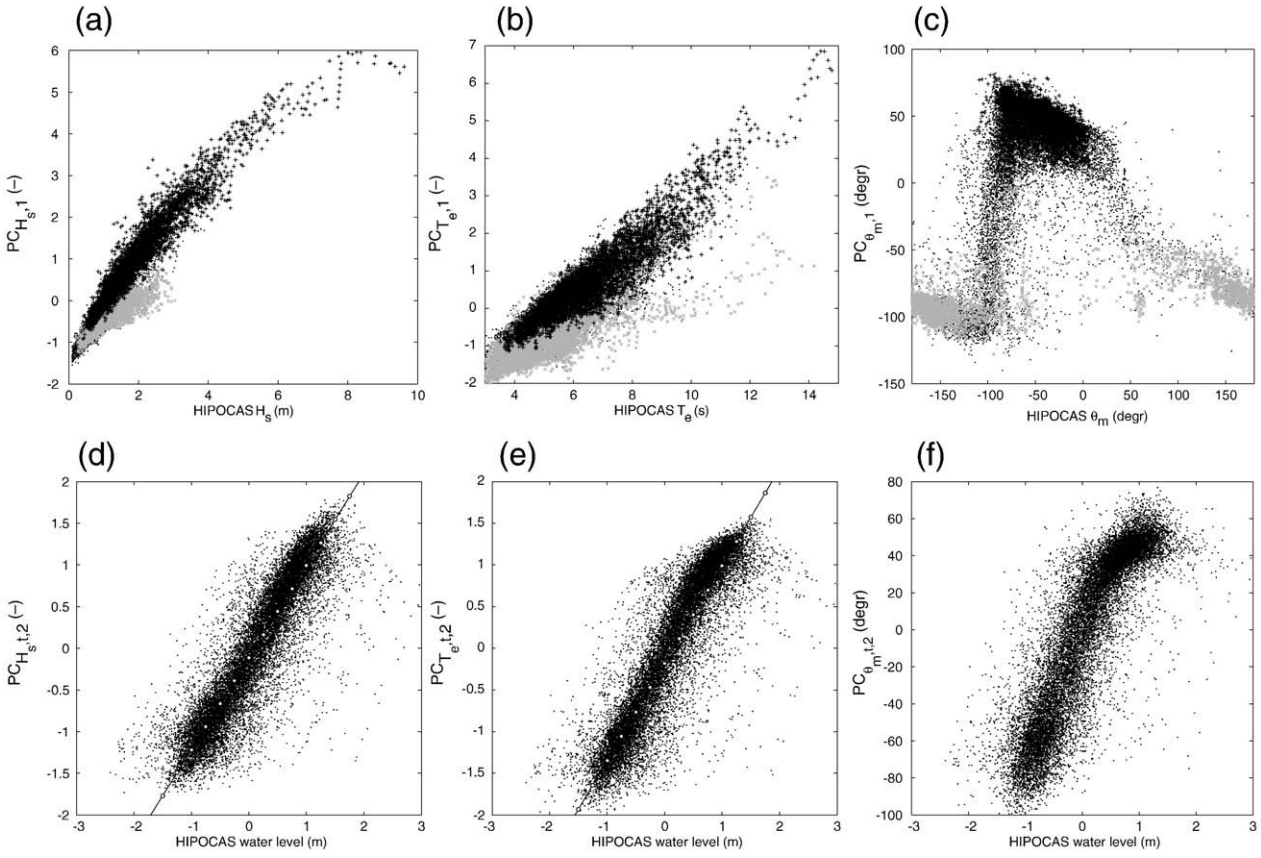

Fig. 10. Scatter plots of the chosen PCs against the HIPOCAS time series: $H_{s,HIP} - \mathcal{P}_{H_s,1}$ (a), $T_{e,HIP} - \mathcal{P}_{T_e,1}$ (b), $\theta_{m,HIP} - \mathcal{P}_{\theta_m,1}$ (c), $\xi_{HIP} - \mathcal{P}_{H_s,2}$ (d), $\xi_{HIP} - \mathcal{P}_{T_e,2}$ (e) and $\xi_{HIP} - \mathcal{P}_{\theta_m,2}$ (f). In the upper panels (a,b,c) data points for $u_w > 5$ m/s ($u_w < -5$ m/s) are shown with black (gray) crosses; the remaining points ($|u_w| \leq 5$ m/s) are marked with dots. Lines in (d,e) show the linear least-square fit.

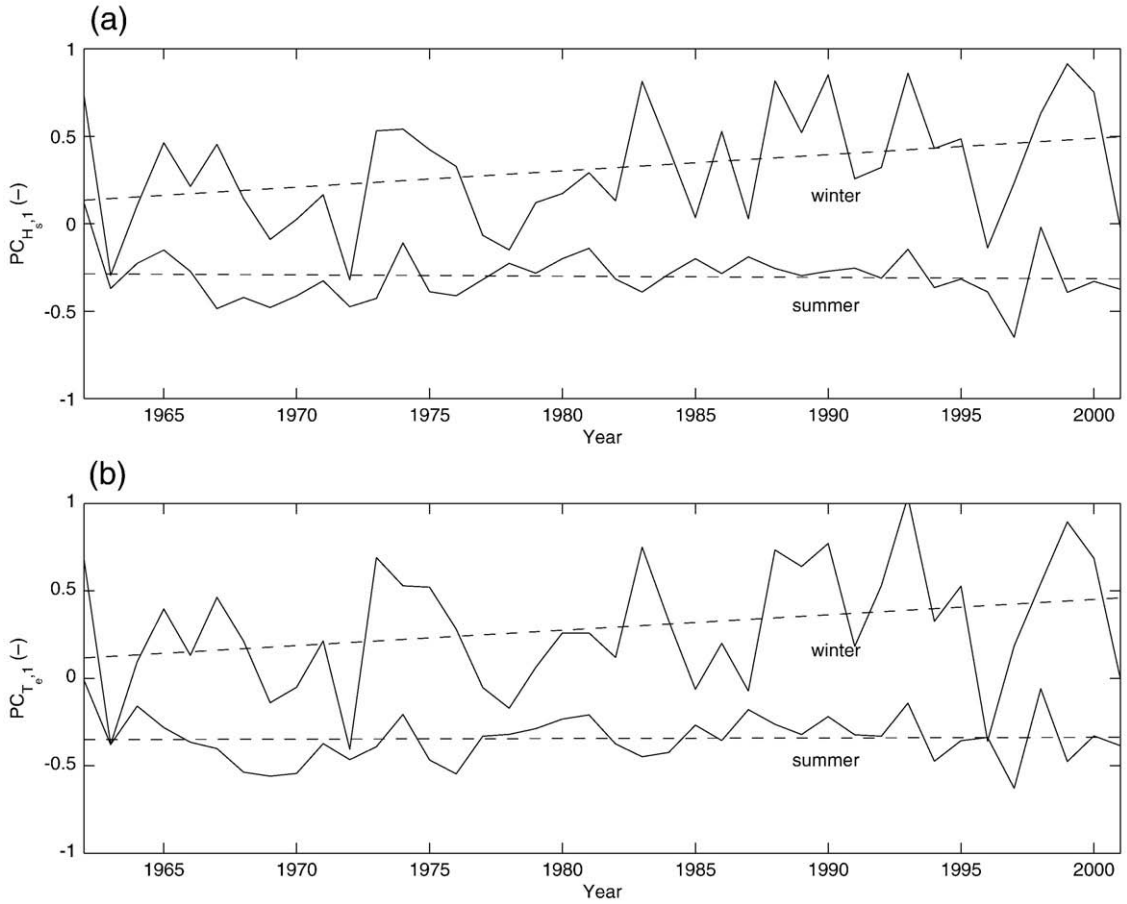

Fig. 11. Trends in the summer- (June–August) and wintertime (December–February) average values of $\mathcal{P}_{H_s,1}$ (a) and $\mathcal{P}_{T_e,1}$ (b).

series in one chosen location. This is an important aspect of the proposed technique. In search of possible methods that could further improve the quality of the NN-modelling results, we tested a possibility of using the data from two HIPOCAS input points instead of only one. Without changing the structure of the neural networks, time series of water levels and wave parameters from a HIPOCAS point located in the north-west part of the model domain were used as additional input neurons. Only very minor changes in the results were observed. For example, in the case of the PCs of H_s, the correlation coefficients improved by less than 0.2% and the standard deviation of differences by less than 0.02. This is not surprising after a brief examination of the input data. The wave conditions to the north of the islands are highly uniform. If a one-hour phase shift is taken into account, the correlation between the H_s and T_e in the two points considered equal 0.95 and 0.97, respectively. In the case of water levels the correlation is even higher, namely 0.996. Generally, to make the technique presented here applicable to regions with more nonuniform forcing, PCA of the forcing variables could be used to represent them as a few PCs used then as input for the neural networks.

Another issue worth some discussion—and one of the possible sources of errors in the modelling scheme described in this paper—is the fact that the constant bottom topography was used throughout the whole 1962–2002 period, as already mentioned in Section 3. It must be stressed that at the time scales analyzed here the outer parts of the ebb deltas are the only regions that can undergo substantial morphological changes. However, the amount of bathymetric data available is insufficient to properly account for these variability in the four decades studied, especially in the first part of this period. Sensitivity studies based on the bathymetric survey data from the years 1960 and 1995 (Kaiser and Niemeyer, 1999) show that the influence of the morphological changes within the ebb delta of the Norderneyer Seegat are significant under some conditions, but they are spatially limited. Further studies are necessary to analyze this issue in more detail. However, it does not affect the methodology presented in this paper, in the sense that one could treat the bottom topography as an additional (variable) forcing factor shaping the wave conditions—analogously to the way in which the meteorological and tidal forcing are handled.

Finally, the amount of data used for the training and validation of the neural networks should be commented upon. The performance of the NN modelling would be presumably better, especially by extreme/untypical situations, if longer training and validation time series were used. However, the amount of computer resources needed by SWAN made it impossible to produce the required model data in acceptable time spans. The calculations for the years 1962, 1963 and 1985 took over 13 months on a 16-processor cluster available for the project realization. These huge computational resources necessary for high-resolution spectral wave modelling provide one of the important arguments in favour of the technique described in this paper.

Acknowledgements

The work presented here was part of the research project MOSES ("Modelling of the medium-term wave climatology at the German North Sea coast"; project No. 03 KIS 040) financed by the German Federal Ministry for Education and Research (BMBF) under the umbrella of the German Coastal Engineering Research Council (KFKI). We would like to thank the anonymous reviewers for valuable comments that helped to improve the quality of this paper.

Appendix A

A.1. Definitions of the mean wave parameters

The mean wave parameters used in this study, i.e. the significant wave height H_s, the energy period T_e and the mean wave direction θ_m, are defined as:

$$H_s = 4\left[\int_0^{\infty}\int_0^{2\pi} E(\omega,\theta)\mathrm{d}\omega\mathrm{d}\theta\right]^{1/2},$$

$$T_e \equiv T_{-1,0} = 2\pi\int_0^{\infty}\int_0^{2\pi} \omega^{-1}E(\omega,\theta)\mathrm{d}\omega\mathrm{d}\theta\left[\int_0^{\infty}\int_0^{2\pi} E(\omega,\theta)\mathrm{d}\omega\mathrm{d}\theta\right]^{-1},$$

$$\theta_m = \arctan\left(\int_0^{\infty}\int_0^{2\pi} \sin\theta E(\omega,\theta)\mathrm{d}\omega\mathrm{d}\theta\left[\int_0^{\infty}\int_0^{2\pi} \cos\theta E(\omega,\theta)\mathrm{d}\omega\mathrm{d}\theta\right]^{-1}\right),$$

where E denotes the spectral energy density, ω—the wave frequency and θ—the wave propagation direction.

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.coastaleng.2009.02.007.

References

Agrawal, J.D., Deo, M.C., 2004. Wave parameter estimation using neural networks. Mar. Struct. 17, 536–550.

Booij, N., Ris, R.C., Holthuijsen, L.H., 1999. A third-generation wave model for coastal regions. Part I: model description and validation. J. Geophys. Res. 104 (C4), 7649–7666.

Booij, N., Haagsma, I.J.G., Holthuijsen, L.H., Kieftenburg, A.T.M.M., Ris, R.C., van der Westhuysen, A.J., Zijlema, M., 2004. SWAN Cycle III Version 40.41. User manual. Technical report, Delft University of Technology. 115 pp.

Browne, M., Strauss, D., Castelle, B., Blumenstein, M., Tomlinson, R., Lane, C., 2006. Empirical estimation of nearshore waves from a global deep-water wave model. IEEE Geosci. Remote Sens. Soc. Newsl. 3, 462–466.

Cherkassky, V., Krasnopolsky, V., Solomatine, D.P., Valdes, J., 2006. Computational intelligence in Earth sciences and environmental applications: issues and challenges. Neural Netw. 19, 113–121.

de Vriend, H.J., Capobianco, M., Chesher, T., de Swart, H.E., Latteux, B., Stive, M.J.F., 1993. Approaches to long-term modelling of coastal morphology: a review. Coast. Eng. 21, 225–269.

Delft Hydraulics, 2003. User manual of Delft3D-Flow: simulation of multi-dimensional hydrodynamic flows and transport phenomena, including sediments. Technical Report, Delft Hydraulics. 497 pp.

Deo, M.C., Naidu, C.S., 1999. Real time wave forecasting using neural networks. Ocean Eng. 26, 191–203.

Deo, M.C., Jha, A., Chaphekar, A.S., Ravikant, K., 2001. Neural networks for wave forecasting. Ocean Eng. 28, 889–898.

Deo, M.C., Gondane, D.S., Kumar, V.S., 2002. Analysis of wave directional spreading using neural networks. J. Waterw. Port Coast. Ocean Eng. 128, 30–37.

Günther, H., Hasselmann, S., Janssen, P.A.E.M., 1992. The WAM model cycle 4. Technical Report, Deutsch Klim. Rechenzentrum, Hamburg, Germany.

Hayes, M., 1979. Barrier island morphology as a function of tidal and wave regime. In: Leatherman, S.P. (Ed.), Barrier Islands from the Gulf of St. Lawrence to the Gulf of Mexico. Academic Press, New York, pp. 1–29.

Herman, A., 2007. Numerical modelling of water transport processes in partially connected tidal basins. Coast. Eng. 54, 297–320.

Herman, A., Kaiser, R., Niemeyer, H.-D., 2007. Modelling of a medium-term dynamics in a shallow tidal sea, based on combined physical and neural network methods. Ocean Model. 17, 277–299.

Jolliffe, I.T., 2002. Principal Component Analysis, 2nd ed. Springer. 487 pp.

Kaiser, R., Niemeyer, H.-D., 1999. Changing of local wave climate due to ebb-delta migration. Proc. 26th Int. Conf. Coastal Engng. ASCE, Copenhagen, Denmark.

Kaiser, R., Niemeyer, H.-D., 2001. Analysis of directional spectra in shallow environment—comparison of field data and results from mathematical modelling. Proc. Int. Conf. Ocean Wave Measurement and Analysis. Reston/Va., USA.

Kalra, R., Deo, M.C., Kumar, R., Agrawal, V.K., 2005. Artificial neural network to translate offshore satellite wave data to coastal locations. Ocean Eng. 32, 1917–1932.

Krasnopolsky, V.M., Chevallier, F., 2003. Some neural network applications in environmental sciences. Part II: advancing computational efficiency of environmental numerical models. Neural Netw. 16, 335–348.

Latteux, B., 1995. Techniques for long-term morphological simulation under tidal action. Mar. Geol. 126, 129–141.

Londhe, S.N., Panchang, V., 2006. One-day wave forecasts based on artificial neural networks. J. Atmos. Ocean Technol. 23, 1593–1603.

Makarynskyy, O., 2004. Improving wave predictions with artificial neural networks. Ocean Eng. 31, 709–724.

Makarynskyy, O., 2005. Artificial neural networks for wave tracking, retrieval and prediction. Pacific Oceanogr. 3, 21–30.

Makarynskyy, O., Makarynska, D., 2007. Wave prediction and data supplementation with artifical neural networks. J. Coast. Res. 23, 951–960.

Makarynskyy, O., Makarynska, D., Rusu, E., Gavrilov, A., 2005a. Filling gaps in wave records with artificial neural networks. Proc. Maritime Transportation and Exploitation of Ocean and Coastal Resources, IMAM 2005, 26–30.09.2005, Lisbon, Portugal.

Makarynskyy, O., Pires-Silva, A., Makarynska, D., Ventura-Soares, C., 2005b. Artificial neural networks in wave predictions at the west coast of Portugal. Comput. Geosci. 31, 415–424.

Nabney, I.T., 2004. NETLAB. Algorithms for pattern recognition, Advances in Pattern Recognition, 4th ed. Springer-Verlag, UK. 420 pp.

Niemeyer, H.-D., Kaiser, R., 2001. Design wave evaluation for coastal protection structures in the Wadden Sea. Proc. Int. Conf. Ocean Wave Measurement and Analysis. Reston/Va., USA.

Pawlowicz, R., Beardsley, B., Lentz, S., 2002. Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE. Comput. Geosci. 28, 929–937.

Preisendorfer, R.W., 1988. Principal Component Analysis in Meteorology and Oceanography. Elsevier Science, New York. 425 pp.

Rao, S., Mandal, S., 2005. Hindcasting of storm waves using neural networks. Ocean Eng. 32, 667–684.

Ris, R.C., 1997. Spectral modelling of wind waves in coastal areas. Communications on Hydraulic and Geotechnical Engineering. Delft University of Technology. 160 pp.

Ris, R.C., Booij, N., Holthuijsen, L.H., 1999. A third-generation wave model for coastal regions. Part II: verification. J. Geophys. Res. 104 (C4), 7667–7681.

Tolman, H.L., Krasnopolsky, V.M., Chalikov, D.V., 2005. Neural network approximations for nonlinear interactions in wind wave spectra: direct mapping for wind seas in deep water. Ocean Model. 8, 253–278.

Tsai, C.-P., Lin, C., Shen, J.-N., 2002. Neural network for wave forecasting among multi-stations. Ocean Eng. 29, 1683–1695.

van der Merwe, R., Leen, T.K., Lu, Z., Frolov, S., Baptista, A.M., 2007. Fast neural network surrogates for very high dimensional physics-based models in computational oceanography. Neural Netw. 20, 462–478.

Weisse, R., Feser, F., Günther, H., 2002. A 40-year high-resolution wind and wave hindcast for the Southern North Sea. Proc. 7th International Workshop on Wave Hindcasting and Forecasting, Banff, Alberta, Canada, pp. 97–104.

Zhang, Z., Li, C.-W., Qi, Y., Li, Y.-S., 2006. Incorporation of artificial neural networks and data assimilation techniques into a third-generation wind-wave model for wave forecasting. J. Hydroinform. 8, 65–76.