Available online at www.sciencedirect.com
SCIENCE DIRECT®
Ocean Modelling 8 (2005) 253–278
Ocean Modelling
www.elsevier.com/locate/ocemod

Neural network approximations for nonlinear interactions in wind wave spectra: direct mapping for wind seas in deep water ☆

Hendrik L. Tolman [a,*], Vladimir M. Krasnopolsky [a], Dmitry V. Chalikov [1,2]

[a] *SAIC-GSO at NOAA/NCEP/EMC, Marine Modeling and Analysis Branch, 5200 Auth Road Room 209, Camp Springs, MD 20746, USA*

Received 26 August 2003; received in revised form 12 October 2003; accepted 23 December 2003
Available online 21 January 2004

Abstract

The potential of Neural Networks (NN) to provide accurate estimates of nonlinear interactions for wind wave spectra by means of direct mapping is considered. Expanding on a previously reported feasibility study, an Empirical Orthogonal Functions (EOF) based NN for single peaked spectra is shown to be much more accurate than the well known Discrete Interaction Approximation (DIA), at the expense of a moderate increase of computational costs. This Neural Network Interaction Approximation (NNIA) gives reasonable results for modeled wave spectra, but is not yet capable of providing acceptable model integrations. Methods to expand the NNIA to be suitable for model integration are discussed.

Keywords: Wind waves; Nonlinear interactions; Neural networks; Numerical modeling; Wave model physics; Empirical orthogonal functions

1. Introduction

Wind waves on oceans, seas, and lakes are generally described with a wave energy or action spectrum. Such spectra describe the distribution of wave energy over various wavenumbers,

☆ MMAB contribution Nr. 234.

* Corresponding author. Tel.: +1-301-763-8000x7253; fax: +1-301-763-8545.
E-mail addresses: hendrik.tolman@noaa.gov (H.L. Tolman), vladimir.krasnopolsky@noaa.gov (V.M. Krasnopolsky), dchalikov@essic.umd.edu (D.V. Chalikov).
[1] UCAR Visiting Scientist at NOAA/NCEP/EMC/MMAB.
[2] Now at University of Maryland, ESSIC.

1463-5003/$ - see front matter
doi:10.1016/j.ocemod.2003.12.008

frequencies and directions. Following Hasselmann (1960), the spectral conservation equation for wind waves generally has the form

$$\frac{\mathrm{D}F}{\mathrm{D}t} = S_{\mathrm{tot}} = S_{\mathrm{in}} + S_{\mathrm{nl}} + S_{\mathrm{ds}}, \tag{1}$$

where F is the spectrum, and S_{tot} represents the sources and sinks, consisting of a wind input (S_{in}), nonlinear interactions (S_{nl}) and dissipation (S_{ds}). The spectrum F is a density function describing the distribution of wave energy or variance in spectral space. The spectral space is defined by the wavenumber vector $\vec{k}$ and the corresponding (radian) frequency σ, and hence is in principle a three-dimensional space. However, using an inherently linear spectral description, the wavenumber $k = |\vec{k}|$ and the frequency σ satisfy the linear dispersion relation

$$\sigma^2 = gk \tanh kd, \tag{2}$$

where d is the mean water depth. Hence the spectral space becomes two dimensional, and spectra are typically defined as $F(\vec{k})$, $F(k,\theta)$, $F(\sigma,\theta)$ or $F(f,\theta)$, where θ is the direction of $\vec{k}$ and $f = \sigma/2\pi$. The different spectral forms are related by straightforward Jacobian transformations. Source terms are defined for the same spectral domain that is used to define F, and represent $\partial F/\partial t$ for the corresponding physical process. Following common practice in operational wave modeling, we will present results here in terms of the variance spectrum [3] $F(f,\theta)$ and its accompanying source term. The units of this spectrum are m^2/Hz/rad or m^{2}s, whereas the units of the source terms hence are m^2.

Because Eq. (1) is inherently linear, nonlinear contributions (S_{nl}) are represented as a source term on the right hand side. These nonlinear interactions do not generate or dissipate momentum, energy or action, but re-distribute it over the spectrum. For practical purposes S_{nl} is negligible if swell propagation over large distances is considered (e.g., Snodgrass et al., 1966). However, the pioneering theoretical work of Phillips (1960) and Hasselmann (1962, 1963a,b) and experimental validation in the JONSWAP project (Hasselmann et al., 1973), have identified the crucial role of S_{nl} in wave growth. The nonlinear interactions provide the lowest order mechanism to shift wave energy to longer waves, and provide a stabilization mechanism for the shape of the spectrum. Reviews of the interactions and their impact can be found in Phillips (1981), Young and Van Vledder (1993) or Komen et al. (1994).

Nonlinear interactions as identified by Hasselmann (1962, 1963a), describe the resonant exchange of energy, momentum and action between four spectral components with wavenumber vectors $\vec{k}_1$–$\vec{k}_4$ and (radian) frequencies σ_1–σ_4. Such a group of wave components is generally called a quadruplet, and has to satisfy the resonance conditions

$$\vec{k}_1 + \vec{k}_2 = \vec{k}_3 + \vec{k}_4, \tag{3}$$

$$\sigma_1 + \sigma_2 = \sigma_3 + \sigma_4. \tag{4}$$

The interactions are conventionally expressed in terms of the rate of change of the action spectrum $n \equiv F/\sigma$ as a function of the wavenumber vector $\vec{k}$, as

[3] Erroneously named the energy spectrum in most wave literature.

$$\frac{\partial n_1}{\partial t} = \int\int\int G(\vec{k}_1,\vec{k}_2,\vec{k}_3,\vec{k}_4)\delta_k\delta_\sigma[n_1n_3(n_4-n_2)+n_2n_4(n_3-n_1)]\,\mathrm{d}\vec{k}_2\,\mathrm{d}\vec{k}_3\,\mathrm{d}\vec{k}_4, \quad (5)$$

where n_i is the action density at component i, $n_i = n(\vec{k}_i) = F(\vec{k}_i)/\sigma$, G is a complicated coupling coefficient (Webb, 1978; Herterich and Hasselmann, 1980), and δ_k and δ_σ are delta functions corresponding to the resonance conditions (3) and (4), respectively.

The numerical solution of Eq. (5) is computationally expensive for two reasons: First, a six-dimensional integration is required to obtain an estimate of S_{nl} for a single wavenumber vector $\vec{k}$. Second, the integration is complicated by singularities in G. Although major strides have been made in optimizing calculations (see reviews by Young and Van Vledder, 1993; Hashimoto and Kawaguchi, 2001; Van Vledder, 2002a), evaluation of S_{nl} according to Eq. (5) still takes on the order of 10^4 times more computational effort than all the other parts of a typical wave model combined. This is simply unacceptable for operational wave models, which in essence require the computational effort expended on S_{nl} to be of the same order of magnitude as the effort expended on the remainder of the model (or less).

To circumvent this problem, most early wave models used highly simplified approximations for S_{nl} (see review by the SWAMP group, 1985). Such simplified approaches generally focus on the shift of energy to lower frequencies, but fail to capture the stabilizing influence of S_{nl} on the spectral shape. This implies that spectral shapes have to be explicitly assigned, and corresponding wave models are generally denoted as second generation models (SWAMP group, 1985).

A major breakthrough in wave modeling was achieved with the development of the Discrete Interaction Approximation (DIA, Hasselmann et al., 1985). In this parameterization of S_{nl} only a very small sub-set of resonant quadruplets is considered, for which a discrete analogue of Eq. (5) is evaluated. This effectively replaces the six-fold integral in Eq. (5) with a single summation of contributions from a small part of the spectral space, and also replaces the complicated function G with a representative constant. The resulting algorithm proved to be sufficiently economical for application in practical wave models, which made the development of the WAM model (WAMDIG, 1988; Komen et al., 1994) possible. The WAM model was the first so-called third-generation wind wave model. In this class of models, Eq. (1) is integrated without pre-assigned spectral shapes. Since the development of WAM, several other third-generation wave models have been built around the DIA (e.g., Tolman, 1991; Abdalla and Ozhan, 1993; Van Vledder and Dee, 1994; Booij et al., 1999).

When the DIA was introduced, it was clear that its approximation of the actual nonlinear transfer had some systematic shortcomings. For example, it represents shifting of energy to lower frequencies reasonably well, but has large systematic errors at higher frequencies (Hasselmann et al., 1985, Fig. 7). The latter results in spectra with unrealistically high energy levels at high frequencies, and with unrealistically broad directional distributions (Hasselmann et al., 1985, Fig. 12). Such deficiencies were acceptable in the context of our limited understanding of other source terms when WAM was originally developed. Nowadays, however, our increased understanding of the source term balance makes these errors less acceptable. This has resulted in a recently increased effort to find more accurate (yet economical) parameterizations of S_{nl} (e.g., Van Vledder et al., 2000).

Within the above context, we are investigating the potential of NN to provide accurate yet economical estimates for the nonlinear interactions. Development of a useful Neural Network

Interaction Approximation (NNIA) is a complex interdisciplinary undertaking which involves solving many particular problems both in wave modeling and in machine learning (NN) fields, as well as developing an interface between the two fields. It is unrealistic to assume that all problems encountered can be solved at once, and hence an iterative development approach is appropriate. This paper therefore only presents results of the first step of research in a broader research plan. To design an iterative development strategy, it is paramount to recognize main issues in both fields.

From the wave modeling perspective, several critical issues can be identified. At first the economy of the resulting algorithm has already been identified above. A second major issue, is that NNIA should not only be able to produce accurate interactions for individual spectra, but also be able to result in stable model integration in conditions describing wave growth. The exact interactions tend to stabilize the spectral shape for frequencies above the spectral peak frequency by smoothing local spectral perturbations at relatively short time scale (e.g., Resio and Perrie, 1991; Young and Van Vledder, 1993). The inability of many early parameterizations of the interactions to reproduce this aspect makes them unsuitable for application in third-generation wave models (e.g., Hasselmann et al., 1985). A third issue is the shear complexity of wave fields in a general purpose wave model. Apart from an actively wind driven wave field (the wind sea), there may be one or many wave fields present that have been generated elsewhere in the past (swell). Wind seas and swells interact through the nonlinear interactions. Although generally, effects of interactions on swell are small, they will also have to be realistic for a model to give good results.

From a machine learning (NN) perspective, several major issues can also be identified. The first issue is that a direct mapping between spectra and nonlinear interactions would consider too many degrees of freedom to be feasible, and would, by definition, be applicable only to one discrete spectral resolution. To mitigate these problems, the number of degrees of freedom in F and S_{nl} needs to be reduced (although it remains to be seen if a NNIA can be made sufficiently flexible to be applicable to multiple spectral resolutions). The second issue is to establish the necessary complexity of the NN, which depends on the resulting number of degrees of freedom remaining in the algorithm. The third issue is the joint optimization of the reduction of degrees of freedom and the NN, with the corresponding balance between accuracy and economy. Interspersed with the first three steps/issue, is the fourth issue of selecting a proper training data set. Considering the natural complexity of wave spectra and the corresponding nonlinear interactions, it is not clear if it is possible to generate a truly complete training data set. This implies that generalization ability of the NN (or robustness of the complete NNIA) is expected to be a fifth major issue.

The development of a practical NNIA is clearly a daunting task. To assess if an NNIA has any potential at all, a feasibility study was performed (Krasnopolsky et al., 2001, 2002). In this study, the spectra and source terms were decomposed using Legendre and Fourier decompositions, and a simple NN (a feedforward, fully connected multi-layer perceptron) was developed using a training data set consisting of a superpositions of Pierson–Moskowitz spectra (Pierson and Moskowitz, 1964). The following conclusions can be drawn from this study: (i) An NNIA can indeed reproduce nonlinear interactions for complex multi-peaked spectra more accurately than the DIA, at computational costs comparable to those of the DIA. (ii) The large majority of computational costs in the NNIA is expended in the de- and re-composition of the spectra and source terms, with typically less than 10% of the time expended on the actual NN. (iii) Due to the ad hoc nature of

the training data set, and the fact that the appropriate scaling behavior was not built into the NNIA, the resulting NNIA proved to have a very limited applicability.

Considering the above, several critical requirements for the development of a practical NNIA can be identified. Probably the most critical aspect of an NNIA would be the capability to reproduce wave growth for wind sea conditions. Furthermore, this NNIA will have to be sufficiently robust to be applicable to arbitrary growth conditions. Once this capability has been established, the NNIA needs to be generalized for application to arbitrary combinations of wind seas and swell. However, before an NNIA can be developed specifically for wave growth conditions, the basic mapping procedures for the NNIA have to be established, and need to be made applicable to fairly general wind sea spectra. This is the subject of the present study.

Section 2 describes the basic design considerations for the NNIA, as well as the construction of the training data set. For a basic description of Neural Networks in this context, reference is made to Krasnopolsky et al. (2002). Results for several NNIA designs are presented in Section 3. These results include testing of the NNIA on spectra not used for the training, and spectra generated by a numerical model. The testing is expanded by performing limited model integration with the most promising NNIA. This model integration is not the main focus of the present study. It has been added for completeness, and to facilitate the discussion of future research in Section 4.

2. Neural network design and training

The present study is a continuation of work presented in Krasnopolsky et al. (2002, henceforth denoted as KCT). A basic description of NN and the corresponding terminology used here can be found in Appendix A of KCT. In the present approach, the NN is used to generate a direct mapping algorithm from the spectrum F to the source term S_{nl}

$$S_{\mathrm{nl}} = T'(F), \tag{6}$$

where T' is a nonlinear operator. In KCT, the mapping is applied directly between the spectrum and the source term. Here we apply the mapping to nondimensional spectra $\widetilde{F}$ and source terms $\widetilde{S}_{\mathrm{nl}}$, using a normalization frequency f_n, direction θ_n and energy density F_n

$$\tilde{f} = f_n^{-1} f, \tag{7}$$

$$\tilde{\theta} = \theta - \theta_n, \tag{8}$$

$$\widetilde{F}(\tilde{f}, \tilde{\theta}) = F_n^{-1} F(f, \theta), \tag{9}$$

$$\widetilde{S}_{\mathrm{nl}}(\tilde{f}, \tilde{\theta}) = g^4 F_n^{-3} f_n^{-11} S_{\mathrm{nl}}(f, \theta). \tag{10}$$

Because we consider single peaked spectra only, it is natural to choose f_n and θ_n to be the peak frequency f_{p} and direction θ_{p}, and to choose $F_n = F(f_{\mathrm{p}}, \theta_{\mathrm{p}})$, so that $\widetilde{F}(1,0) \equiv 1$. Eq. (6) then becomes

$$\widetilde{S}_{\mathrm{nl}} = T(\widetilde{F}). \tag{11}$$

This normalization ensures proper scaling behavior of the NNIA, independent of the approximation of T. To reduce the size of the mapping problem, and to make the mapping less dependent

on the discrete spectral resolutions, the spectrum $\widetilde{F}$ and source term $\widetilde{S}_{\mathrm{nl}}$ are represented by a linear superposition of two sets of orthogonal functions Φ_i and Ψ_q. That is

$$\widetilde{F} \approx \sum_{i=1}^{n} x_i \Phi_i, \quad \widetilde{S}_{\mathrm{nl}} \approx \sum_{i=1}^{m} y_i \Psi_i, \tag{12}$$

where x_i and y_q are the expansion coefficients. The orthogonality of Φ_i and Ψ_q implies that

$$\int\int \Phi_i \Phi_j = \delta_{ij} = \int\int \Psi_i \Psi_j, \tag{13}$$

where $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ otherwise, and where the double integral identifies integration over the spectral (f, θ) space. Given Eq. (13), the expansion coefficients in Eq. (12) are given by

$$x_i = \int\int F \Phi_i, \quad y_i = \int\int S_{\mathrm{nl}} \Psi_i. \tag{14}$$

Since each orthogonal sets of functions $\{\Phi_i\}_{i=1,\ldots,\infty}$ and $\{\Psi_q\}_{q=1,\ldots,\infty}$ constitute a complete set, they provide a basis to represent arbitrary functions in the spectral domain as assumed in Eq. (12). Then the accuracy of the approximation depends only on the truncation of the sets at n and m components in (12). Substituting (12) into Eq. (11) gives

$$\vec{Y} = T(\vec{X}), \tag{15}$$

which represents a mapping of the finite vectors $\vec{X} \in \Re^n$ and $\vec{Y} \in \Re^m$, and where T still represents the full nonlinear interaction operator. This operator can be approximated with a NN with n inputs and m outputs and l internal components (see Appendix A of KCT)

$$\vec{Y} \approx T_{\mathrm{NN}}(\vec{X}). \tag{16}$$

The accuracy of the approximation T_{NN} is determined by l, and can generally be increased by increasing l. The NN approximation T_{NN} of T is developed using a process called 'training'. A training data set is generated with pairs of vectors $\vec{X}$ and $\vec{Y}$. This data set is then used to train the NN that is central to T_{NN} (see KCT). In our case, a representative set of spectra F_{p} has to be generated with corresponding (exact) interactions $S_{\mathrm{nl,p}}$. Our exact interactions are calculated according to the Webb–Resio–Tracy (WRT) method (Webb, 1978; Tracy and Resio, 1982; Resio and Perrie, 1991), as implemented in the portable exact interactions package developed by Van Vledder (2002b). [4] For each pair $(F, S_{\mathrm{nl}})_{\mathrm{p}}$, the corresponding vectors $(\vec{X}, \vec{Y})_{\mathrm{p}}$ are determined using Eq. (14). All pairs of vectors are then used to train the NN to obtain T_{NN}. Note that the training of a NN is generally very time consuming. However, it needs to be performed only once, after which the application of the NN is generally very fast.

After T_{NN} has been obtained by training, the resulting NN Interaction Approximation (NNIA) algorithm consists of five steps:

[4] Version 3 of the software used here.

(i) Convert F to $\widetilde{F}$ using Eq. (9).
(ii) Decompose $\widetilde{F}$ by applying Eq. (14) to calculate $\vec{X}$.
(iii) Estimate $\vec{Y}$ from $\vec{X}$ using Eq. (16).
(iv) Compose $\widetilde{S}_{\mathrm{nl}}$ from $\vec{Y}$ using Eq. (12).
(v) Convert $\widetilde{S}$ to $\widetilde{S}_{\mathrm{nl}}$ using Eq. (10).

In the design of such an NNIA, two main issues need to be addressed: The first is the selection of the type and number of basis functions. The second is the construction of the training data set. Because a training data set is also required to assess the accuracy of selected basis functions, the training data set will be discussed first.

The training data set has to provide an envelope of realistic spectra, to assure that the resulting NNIA 'interpolates' rather than extrapolates. Sources for training spectra could be observations, models, or parametric spectral descriptions. Considering that in the present study we seek for a natural envelope of spectra, observations might be considered appropriate. However, observed spectra are less suitable, because the directional properties of such spectra are generally poorly known, and are dependent on the data analysis techniques used.

Because the NNIA is ultimately intended to replace the exact interactions in a wave model, modeled wave spectra could be considered ideal for developing the NNIA. However, two practical aspects hamper generating such a training data set. The first is the prohibitive cost of running a wave model with exact interactions, the second is that there are no properly tuned and validated input and dissipation parameterizations available to be used with the exact interactions. Finally, the nonlinear interactions are known to strongly regulate the spectral shape. We therefore expect that a set of modeled spectra and source terms form a less broad envelope of conditions than can be generated with parametric spectra. It is expected that a broader envelope of training conditions will result in a more expensive NNIA. For this reason, a training data set composed out of parametric spectra has been considered here.

It is unavoidable that a parameterically generated training data set cannot incorporate all typical peculiarities of model wave spectra. It may therefore be expected that a reasonable description of source terms for fairly arbitrary spectra can be found, but that such an NNIA cannot yet result in a stable model integration. Taking into account the intended scope of the present study, this is considered acceptable here.

The parametric spectra used here for training the NNIA consist of several components. A general description is given here, whereas details of the corresponding equations are given in Appendix A. The basic description of the frequency spectrum is provided by the commonly used and versatile JONSWAP formulation (see, e.g., Hasselmann et al., 1973). To obtain a full two-dimensional spectrum, directional distributions of Hasselmann et al. (1980) and Ewans (1998) are used. The former represents a classical directional distribution. The latter represents a more modern view, allowing for bimodality in the directional distribution. A review of recent directional distributions can also be found in the latter paper. To this parametric spectrum a directional shear is added. To make these spectra even more general, random noise is added with a given relative magnitude and with given scales in spectral space. With this parametric spectral description, sets of 15,000 spectra and source terms have been generated using Monte Carlo realizations of the relevant parameters of the spectrum (See Appendix A). 10,000 of these spectra were used for generating basis functions and for the training of the NN. The remaining 5000

spectra were used to control and stop the training, and to avoid overfitting (as will be discussed in detail below). An independent set of 10,000 spectra and source terms was generated to validate the resulting algorithm.

Two approaches have been used for the basis functions: The first is a mathematical basis taken from KCT, with the normalization as discussed above added for the present study. As is common in the parametric spectral description of wind waves, separable basis functions were chosen where the frequency and angular dependence were separated. For example, for Φ_i in Eq. (12) this implies

$$\Phi_i(f,\theta) \Rightarrow \Phi_{ij} = \phi_{f,i}(f)\phi_{\theta,j}(\theta). \tag{17}$$

A similar separation was used for Ψ_q. Considering the strongly suppressed behavior of F and S_{nl} for $f \to 0$ and the quickly decreasing asymptotic behavior for $f \to \infty$, generalized Laguerre's polynomials L_i^{α} (Abramowitz and Stegun, 1964) are used to define ϕ_f and ψ_f:

$$\phi_{f,i} = \left[\frac{i!}{\Gamma(\alpha+i+1)}\right]^{1/2} f^{\alpha/2}\mathrm{e}^{-f/2}L_i^{\alpha}(f), \tag{18}$$

where $\psi_{f,i}$ is defined similarly. Because in KCT no directional preferences existed in the training data set of F and S_{nl}, Fourier decomposition was used for ϕ_θ and ψ_θ.

The advantage of this choice of basis functions is the simplicity of the generation of the bases. Disadvantages are the slow convergence of the decompositions (12), as illustrated below, and the necessity to use large values for α to sufficiently suppress basis functions for $f \to 0$. The value of α is limited from above by the numerical stability of the procedure used to generate the basis. This choice of basis functions will henceforth be denoted as the 'mathematical' basis.

As an alternative, a second approach to the basis functions has been investigated in the present study. In this approach, Empirical Orthogonal Functions (EOFs) or principal components are used (Lorenz, 1956; Jolliffe, 1986). EOFs have previously been used in wave modeling to describe nonlinear interactions by Allender et al. (1985) and Hasselmann et al. (1985), but without the rigorous optimization and fitting procedures presently available in the NN field. EOFs are created from the fields to be decomposed and form a statistically optimal basis. In the present case, the basis functions Φ_i and Ψ_j are functions of two variables f and θ. The set of spectra F and source terms S_{nl}, which are used for the training of the NN, are also used to generate the EOFs for decomposing F and S_{nl}. When using EOFs, the basis generation procedure is computationally expensive, with the cost increasing as the resolution of F and S_{nl} increases. However, like the NN training, the basis generation needs to be performed only once. Stored results can be used without the need for re-calculation in a practical NNIA.

The main advantage of EOFs is the fast convergence of the decomposition. This is illustrated in Fig. 1, which shows the average absolute value C_i of the ith component of the basis

$$C_i = \frac{1}{N}\sum_{j=1}^{N}|y_i^j|, \quad i = 0,1,\ldots,(m-1), \tag{19}$$

where N is the number of test spectra ($N = 10,000$ in the present study). The rapid and systematic decrease of C_i with increasing i for the EOF basis (solid line in Fig. 1) indicates rapid convergence. In contrast, the mathematical basis converges more slowly, and does not show monotonic convergence behavior.

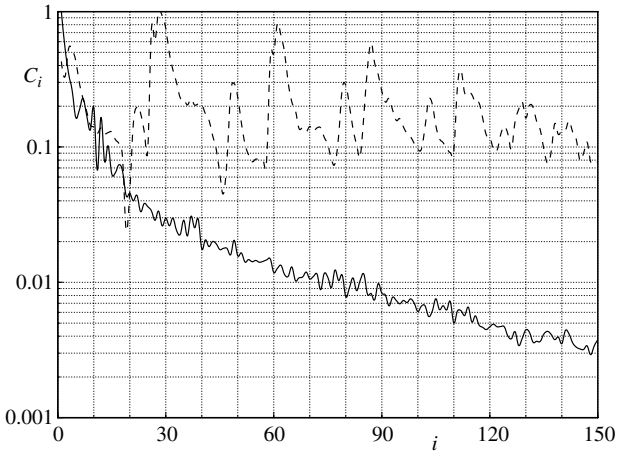


Fig. 1. Average absolute value of the ith component in the decomposition of S_{nl} for the entire training data set, for the mathematical basis (- - -) or the EOF basis (—).

The present NNIAs include three parameters that influence their accuracy: n, m and l. The parameter l represents the number of components inside the NN, and will be considered later. For the mathematical basis, the numbers of basis functions, $n = 50$ and $m = 150$, were selected (see Eq. (12)) based on the somewhat arbitrary criterion that $C_{(n,m)}/C_1 \approx 0.1$. Fig. 1 indicates that a similar accuracy can be achieved with the EOF basis with a much smaller number of components n and m. Alternatively, n and m can be taken identical to those used for the mathematical basis, to illustrate the impact of the basis selection on the accuracy of the truncated decomposition. The latter approach is used here. To quantify the relative accuracy of the EOF basis versus the mathematical basis, several statistics have been calculated.

The first set of statistics is based on rms errors ϵ_j for individual source terms or spectra

$$\epsilon_j = \left\{ \int\int [S^j_{\mathrm{nl}}(f,\theta) - \widehat{S}^j_{\mathrm{nl}}(f,\theta)]^2 \,\mathrm{d}f\,\mathrm{d}\theta \right\}^{1/2}, \tag{20}$$

where S represents the exact source term, and $\widehat{S}$ its approximation (the truncated composition of the basis functions). A similar error can be defined for the spectrum F. Note that in the following, these errors are always calculated for the normalized spectra and source terms $\widetilde{F}$ and $\widetilde{S}_{\mathrm{nl}}$. From the N individual errors ϵ_j, a mean individual error $\bar{\epsilon}_{\mathrm{ind}}$ and a standard deviation of individual errors σ_ϵ can be calculated in the conventional way.

The second set of statistics considers errors in spectral space for the entire data set, described here with the 'field' rms error ϵ_{F},

$$\epsilon_{\mathrm{F}}(f,\theta) = \left\{ \frac{1}{N} \sum_{j=1}^{N} [S^j_{\mathrm{nl}}(f,\theta) - \widehat{S}^j_{\mathrm{nl}}(f,\theta)]^2 \right\}^{1/2}, \tag{21}$$

from which the one-dimensional projections $\epsilon_{\mathrm{F}}(f)$ and $\epsilon_{\mathrm{F}}(\theta)$ and the mean field error $\bar{\epsilon}_{\mathrm{F}}$ can be calculated as

$$\left.\begin{aligned} \epsilon_{\mathrm{F}}(f) &= \int \epsilon_{\mathrm{F}}(f,\theta)\,\mathrm{d}\theta \\ \epsilon_{\mathrm{F}}(\theta) &= \int \epsilon_{\mathrm{F}}(f,\theta)\,\mathrm{d}f \\ \bar{\epsilon}_{\mathrm{F}} &= \int\int \epsilon_{\mathrm{F}}(f,\theta)\,\mathrm{d}f\,\mathrm{d}\theta \end{aligned}\right\}. \tag{22}$$

The corresponding field bias $\beta_{\mathrm{F}}(f,\theta)$ and variability (field standard deviation of the exact solution) $\sigma_{\mathrm{F}}(f,\theta)$ are calculated as

$$\beta_{\mathrm{F}}(f,\theta) = \frac{1}{N}\sum_{j=1}^{N}[\widehat{S}^{j}_{\mathrm{nl}}(f,\theta) - S^{j}_{\mathrm{nl}}(f,\theta)], \tag{23}$$

$$\sigma_{\mathrm{F}}(f,\theta) = \left\{ \frac{1}{N-1}\sum_{j=1}^{N}\left[S^{j}_{\mathrm{nl}}(f,\theta) - \frac{1}{N}\sum_{j=1}^{N} S^{j}_{\mathrm{nl}}(f,\theta)\right]^{2}\right\}^{1/2}. \tag{24}$$

Note that the latter parameter is not an error measure, but a measure for the variability of spectra and source terms in the training data set. A comparison of this variability with the field error ϵ_{F} provides a relative error measure. One-dimensional field biases $\beta_{\mathrm{F}}(f)$ and $\beta_{\mathrm{F}}(\theta)$ and variabilities $\sigma_{\mathrm{F}}(f)$ and $\sigma_{\mathrm{F}}(\theta)$ are calculated by integration, as in Eq. (22).

Table 1 shows the bulk statistics for the mathematical and EOF bases generated from the training data set of 10,000 cases. The statistics are calculated from the independent validation set of 10,000 spectra and source terms. For the spectrum F, the EOF basis results in an error that is a factor of 4.7 ($\bar{\epsilon}_{\mathrm{ind}}$) or 8.0 ($\bar{\epsilon}_{\mathrm{F}}$) smaller than for the mathematical basis. For the source term S_{nl} the factors of improvement are 8.3 and 13.3, clearly indicating the superior results obtained with the EOF bases.

With the above established basis functions, the NN can be trained and tested. The training was performed using a backpropagation algorithm (Rumelhart et al., 1986), and the Nguyen and Widrow (1990) algorithm was used to initialize the NN weights. As mentioned above, the accuracy of the mapping T_{NN} in Eq. (16) depends on the number of internal elements l in the NN. An increase in l, in principle, results in a more accurate NN. In practice, however, since the data include errors, an optimum l exists. In our case, truncation in the decomposition introduces errors. If a dataset includes errors, a choice of l that is too large results in 'over-fitting' of the data. In such a case, the NN tries to fit the errors in the data, rather than the useful content. Unfortunately, the optimum l can only be estimated by trial and error. In Table 2, the impact of l on the accuracy of the NN is investigated. A similar comparison of errors for the training and test data sets was used to determine the proper moment to stop the training for each choice of l.

Table 1
Mean individual rms error $\bar{\epsilon}_{\mathrm{ind}}$, standard deviation σ_{ϵ}, and mean field error $\bar{\epsilon}_{\mathrm{F}}$ for 10,000 validation spectra $\widetilde{F}$ and source term $\widetilde{S}_{\mathrm{nl}}$ due to the decomposition onto a truncated basis with $n = 50$ and $m = 150$ components, respectively

	Basis	$\bar{\epsilon}_{\mathrm{F}}$	$\bar{\epsilon}_{\mathrm{ind}}$	σ_{ϵ}
$\widetilde{F}$	Mathematical	0.14	0.0350	0.0160
	EOF	0.03	0.0044	0.0067
$\widetilde{S}_{\mathrm{nl}}$	Mathematical	0.5175	0.0425	0.1397
	EOF	0.0623	0.0032	0.0089

All errors are calculated using the independent validation data set.

Table 2
Biases and rms errors of components of the nonlinear interactions $\vec{Y}$ as estimated by the Neural Network (Eq. (16)) using the EOF bases as a function of the number of internal components l in the NN

l	Training set		Test set	
	bias	rms error	bias	rms error
25	0.08	131.0	0.31	141.7
30	0.07	124.1	0.12	131.8
35	0.10	111.8	0.06	126.0
40	0.16	101.6	0.10	121.9
45	0.09	102.4	0.16	122.0
50	0.04	100.7	0.12	119.1
55	0.06	99.8	0.10	123.2
60	0.07	98.4	−0.06	125.3
65	0.07	97.8	0.11	126.9

Errors are calculated per training or test case, and averaged over all cases.

Table 2 shows biases and rms errors of $\vec{Y}$ as estimated by the NN for the training data set and for an independent test data set. The EOF decomposition with $n = 50$ inputs and $m = 150$ outputs is considered, with l ranging from 25 to 65 with increments of 5. The biases do not depend significantly on l. The rms errors for the training dataset improves monotonically with increasing l. However, for the test data set, the rms error has a minimum for $l = 50$. This suggests that overfitting starts to occur for $l > 50$. For this reason, $l = 50$ has been used in the following sections.

3. Results

In this section, results for two versions of the NNIA will be discussed. For convenience, the NNIA using the mathematical basis will be denoted as NNIA-M, whereas the EOF based NNIA will be denoted as the NNIA-E. The two NNIAs are compared using bulk statistics, sample source terms for one of the independent validation spectra (belonging to the same envelope of conditions as the training data set), and sample source terms for a spectrum generated by a wave model (outside the envelope of training conditions). As discussed in Section 1, this covers the main scope of the present study. Looking forward to subsequent parts of this study, a simple model integration test is performed with the NNIA-E, and its economy is briefly assessed.

A first comparison of the original DIA, NNIA-M and NNIA-E is given in Table 3, in which the mean rms errors of Eqs. (20)–(22) for the three parameterizations are presented. In this context,

Table 3
The mean individual rms error $\bar{\epsilon}_{\mathrm{ind}}$, the corresponding standard deviation σ_ϵ, and the mean field error $\bar{\epsilon}_{\mathrm{F}}$ for 10,000 validation source terms estimated with the DIA, NNIA-M or NNIA-E

Algorithm	$\bar{\epsilon}_{\mathrm{F}}$	$\bar{\epsilon}_{\mathrm{ind}}$	σ_ϵ
DIA	1.3266	0.3124	0.5273
NNIA-M	0.6590	0.0882	0.1888
NNIA-E	0.2836	0.0354	0.0750

Errors are calculated using the independent validation data set.

$\widehat{S}_{\mathrm{nl}}$ represents the DIA or NNIA estimate of the source term (instead of the truncated representation considered in the previous section). Using the mean field error $\bar{\epsilon}_{\mathrm{F}}$ as the quantitative performance measure, the NNIA-M is 2.0 times more accurate than the DIA, and the NNIA-E is 4.7 times more accurate than the DIA. Using the mean individual rms error $\bar{\epsilon}_{\mathrm{ind}}$, these ratios are 3.5 and 8.8, respectively. For both error measures, the NNIA-M, therefore, clearly outperforms the DIA, and the NNIA-E is significantly more accurate than the NNIA-M.

A more detailed impression of the errors of the different parameterizations can be obtained by analyzing the rms field errors $\epsilon_{\mathrm{F}}(f,\theta)$ and $\epsilon_{\mathrm{F}}(f,\theta)$, as in Eqs. (21)–(23), in more detail. The corresponding one-dimensional errors as a function of frequency f or direction θ for the validation set of 10,000 spectra are presented in Fig. 2. These field errors are calculated for normalized spectra and source terms. Hence the peak of each individual spectrum is by definition at $\tilde{f}=1$ and $\tilde{\theta}=0$. It is therefore expected that these field errors are representative for distinct spectral ranges, such as near peak, equilibrium and low-frequency ranges.

The bias of the DIA in frequency space (chain line in Fig. 2a) shows a remarkable similarity with the typical signature of $S_{\mathrm{nl}}(f)$. This implies that the DIA systematically overestimates the magnitude of the energy shifts in frequency space. The bias of the DIA in direction space (chain line Fig. 2b) is distinctly negative near the mean direction ($\tilde{\theta}\approx 0$) and distinctly positive at larger

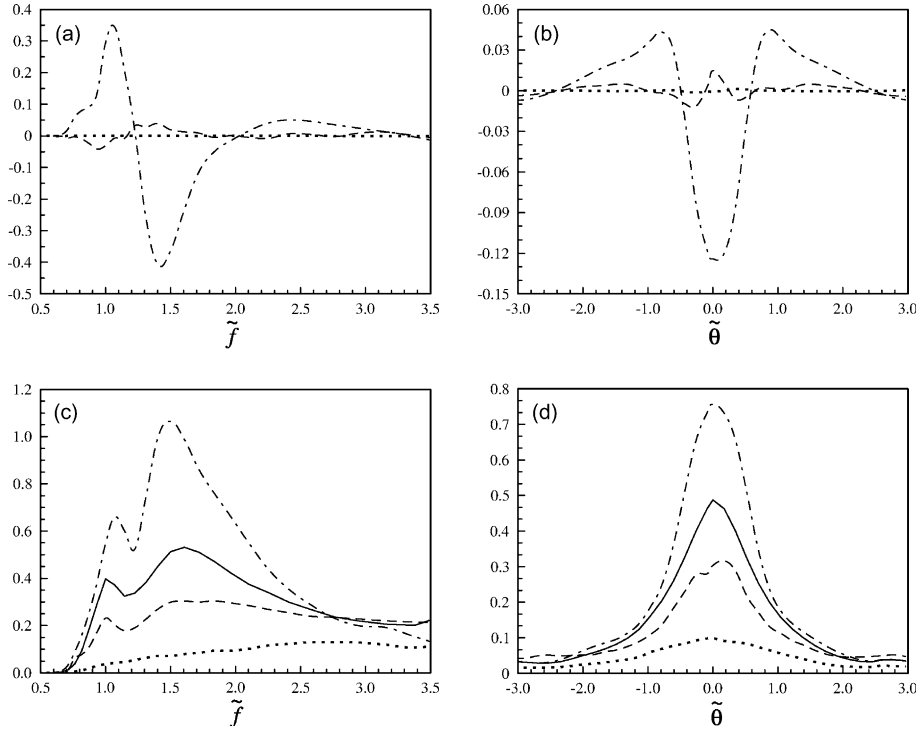


Fig. 2. One-dimensional field biases (upper panels) and rms errors (lower panels) as a function of frequency $\tilde{f}$ (left panels) or direction $\tilde{\theta}$ (right panels) for DIA (chain line) NNIA-M (- - -) and NNIA-E ($\cdots$). Solid line in lower panels represents the variability in the exact interactions (Eq. (24)).

absolute angles, implying that the DIA spuriously broadens the directional distribution. Both deficiencies were already identified by Hasselmann et al. (1985) for selected test spectra, and are confirmed here to be present for a broad range of spectral shapes.

The NNIA-M shows greatly reduced biases compared to the DIA (compare dashed and chain lines in Fig. 2a and b). In frequency space, the bias of the NNIA-M shows a behavior opposite to that of the exact interactions, with a slight negative lobe at low frequencies and a positive lobe for intermediate frequencies (dashed line in Fig. 2a). This suggests that the NNIA-M slightly underestimates the re-distribution of energy in frequency space. The biases in directional space are more complicated, without a clearly discernible pattern.

The field biases of the NNIA-E are negligible compared to the biases of either the DIA or the NNIA-M (dotted lines in Fig. 2a and b). For practical purposes, the NNIA-M hence appears to be free of systematic biases.

The rms field errors of all three parameterizations in frequency and direction space ($\epsilon_F(f)$ and $\epsilon_F(\theta)$, chain, dashed and dotted lines 2c and d), in broad lines show similar behavior as the corresponding variability in the exact interactions (solid lines). This suggests that relative errors in all parameterizations are fairly constant in spectral space.

The rms errors of the DIA (chain lines in Fig. 2c and d) are larger than the variability in the exact interactions in the test data set (solid lines). This is not surprising, because the DIA has been documented to overestimate interactions in this region by typically a factor of 2–3 (e.g., Hasselmann et al., 1985). The figures shown indicate again that the previously reported poor quantitative behavior of the DIA is prevalent for a large range of spectral shapes. The NNIA-M presents a clear improvement over the DIA for all but the highest frequencies ($\tilde{f} > 2.5$; compare chain and dashed lines in Fig. 2c). The NNIA-E provides an improvement over both the DIA and NNIA-M. Note that the improved behavior of the NNIA-E is mostly realized for frequencies $\tilde{f} < 2$. For higher frequencies, the NNIA-E shows the best performance of all three parameterizations, but with smaller relative improvements (dotted lines in Fig. 2c).

Although the bulk error statistics as shown above indicate the potential of an NNIA in terms of accuracy, they do not necessarily imply than the resulting interactions are satisfactorily realistic. To illustrate the latter, the resulting interactions for an arbitrary spectrum from the independent validation data set are presented in Fig. 3. These results were found to be fairly representative for the entire validation dataset.

Fig. 3a shows a moderately sheared wind sea spectrum with an asymmetric bimodal directional distribution at high frequencies. Large differences are found between the exact interaction for this spectrum (Fig. 3b) and the DIA (Fig. 3c), both in general shape of the interactions and in magnitude (note that the contours are logarithmically distributed with an increment factor of 2). The NNIA-M (Fig. 3d) is clearly much closer to the exact interactions (panel b) than the DIA (panel c). However, the NNIA-M has an unrealistic 'noisy' appearance, which can be attributed to the Gibbs effect as it occurs in Fourier re-composition with truncated series. In the NNIA-E (Fig. 3e), the mathematical decomposition has been replaced by EOFs. This results in a more accurate estimate of the interactions, as already addressed with bulk statistics above. Visual inspection confirms that it also results in a more realistic representation of the interactions, with no notable spurious noise.

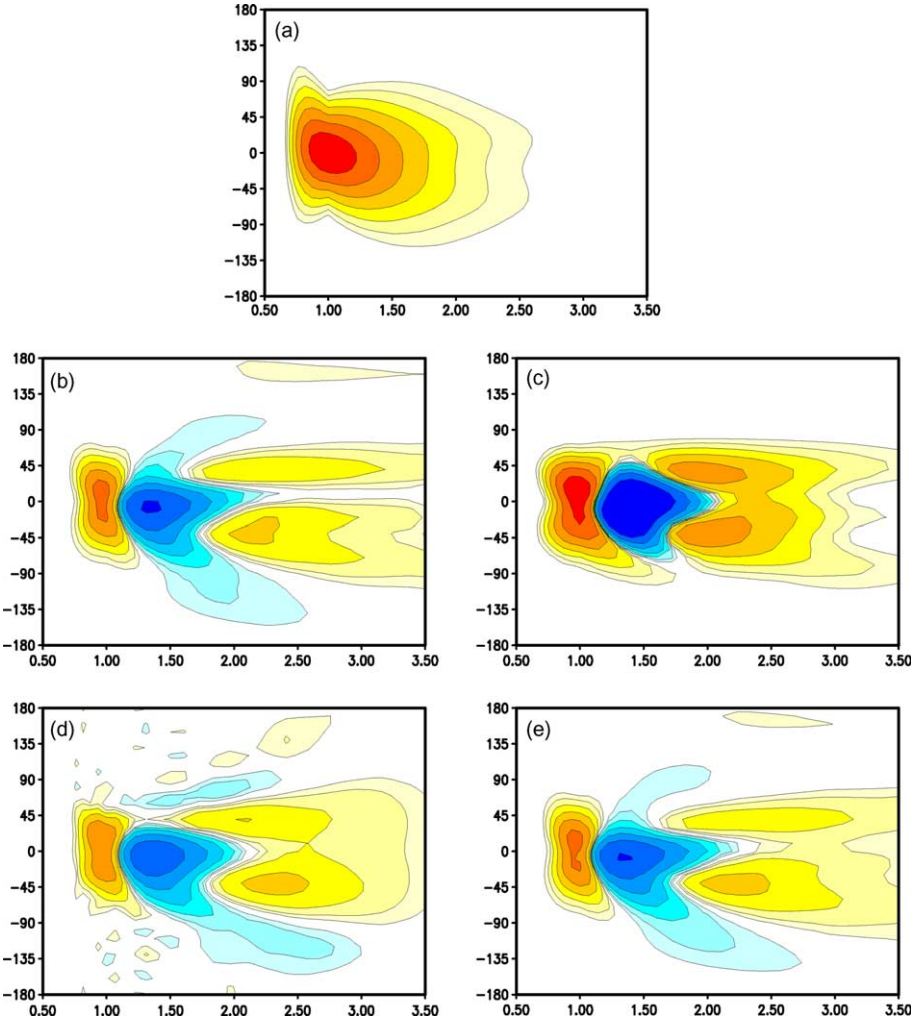


Fig. 3. Normalized parametric spectrum ($\widetilde{F}$, panel a) and nonlinear interactions ($\widetilde{S}_{\mathrm{nl}}$) according to the exact solution (WRT, panel b), the DIA (panel c), the NNIA-M (panel d), and the NNIA-E (panel e), as a function of the normalized frequency ($\tilde{f}$, horizontal axes) and direction ($\tilde{\theta}$, vertical axis). Logarithmic contours are at a factor 2 increment. The highest contour for the spectrum is 0.5. The lowest contour for the interactions is ±100. In the color version of this figure, yellow and red shading identifies positive values, and blue shading identifies negative values. In the black and white version solid and dashed lines represent positive and negative values, respectively.

Whereas Fig. 3 shows promising results for the NNIA-E in particular, we stress that the spectrum used here is similar to the spectra used to train the NNIA-E. When such an NNIA is applied to spectra from other sources more modest results may be expected, when such spectra include features not presently included in the training data set. As mentioned in the introduction, several such features might be expected in model wave spectra. To test the performance of the present NNIAs with respect to model spectra, several such spectra were generated using the WAVEWATCH III model (Tolman, 1991, 2002; Tolman et al., 2002), using the input and

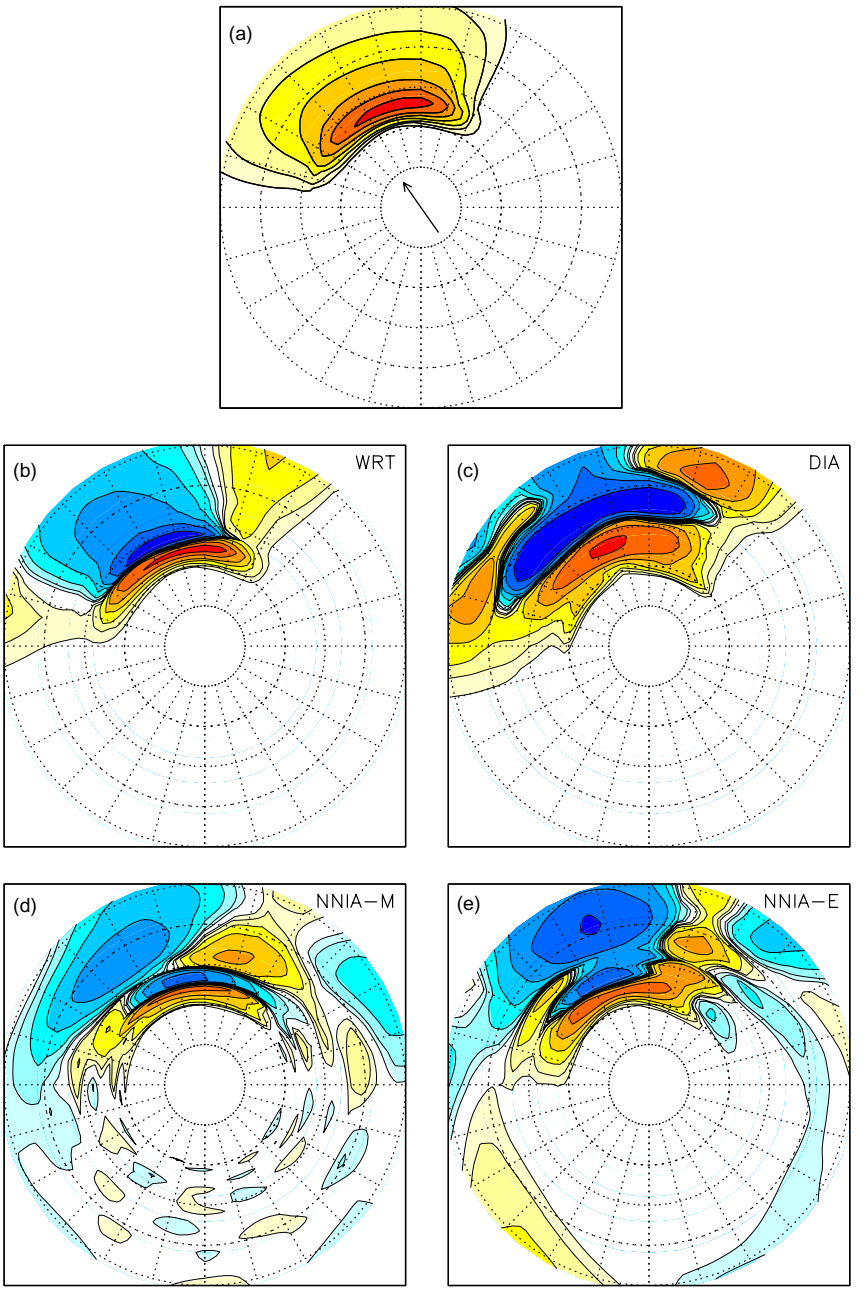


Fig. 4. Wave model spectrum (F, panel a) and nonlinear interactions (S_{nl}, panels b–e): exact solution (WRT, panel b), DIA (panel c), NNIA-M (panel d), and NNIA-E (panel e), as a function of the frequency (f, with $f = 0$ at center and $f = 0.25$ Hz at outer circle) and direction (θ, direction in which waves propagate from center of plot). Grid lines are at 15° and 0.05 Hz intervals. Logarithmic contours are at factor 2 increment. The lowest contour for the spectrum is 2×10^{-1} m^2 s. Lowest contour for interactions $\pm2\times10^{-5}$ m^2. Shading and line styles as in Fig. 3. The arrow in the center of panel a indicates the wind direction.

dissipation source terms of Tolman and Chalikov (1996) and the WRT interactions of Van Vledder (2002a). An arbitrary spectrum from the interactive test cases provided with WAVEWATCH III [5] is presented in Fig. 4.

The results in Fig. 4 were obtained with a spectral resolution of 36 directions ($\Delta\theta = 10°$) and 35 frequencies starting at 0.0418 Hz, and with an increment factor of 1.07. The wind speed considered is 20 m/s. To obtain the NNIA results for this spectrum, normalization is required. In this normalization, grids are shifted and the calculation grid generally does not cover the entire wave model grid and vice versa. Under such conditions, spectra and source terms not covered at the low frequency end of the grid are assumed to be zero. Spectra at the high frequency end of the discrete space are estimated using an f^{-5} parametric spectral shape. Following common practice in wave modeling, the results in Fig. 4 are presented using a polar representation.

The exact interactions shown in Fig. 4b have been used to calculate the spectrum shown in Fig. 4a, and are, therefore, consistent with this spectrum. These interactions are mostly smooth, with locally very sharp features. The estimates from the DIA (Fig. 4c) deviate significantly from the exact interactions, both in details of the source term and in magnitude (note again the logarithmic scaling of the contour levels). The results of the two NNIA algorithms (Fig. 4d and e) differ from each other mainly with respect to the spurious noise throughout the spectral space in the NNIA-M (Fig. 4d), which is consistent with the noise seen in Fig. 3d. This again clearly shows the NNIA-E to be superior to the NNIA-M.

The NNIA-E and the WRT interactions differ significantly for the model spectrum in Fig. 4a, compared to the results for the parametric spectrum shown in Fig. 3. Nevertheless, the NNIA-E reproduces several features of the exact WRT interactions much more accurately than the DIA. First, the positive lobe in the WRT solution around $f = 0.12$ Hz near the NNE direction is well reproduced by the NNIA-E in both magnitude and shape (compare Fig. 4b and e). The DIA reproduces this feature with the correct amplitude, but over a greatly extended area in spectral space. Furthermore, the positive-negative signature in the exact interaction for $f > 0.10$ Hz going from NNW to NE appears somewhat more realistic in the NNIA-E algorithm than in the DIA. The two main differences between the WRT and the NNIA-E algorithms are the spurious positive lobe in the NNIA-E at $f = 0.17$ Hz in the NE direction, and its much broader signature for $f > 0.15$ Hz. Considering the similar behavior in NNIA-M and NNIA-E, this is most likely due to features of the modeled spectrum that are not present in the training data set (as discussed in Section 2). This will be discussed in more detail in Section 4.

As a final test of the capability of the NNIA-E, it has been applied in the WAVEWATCH III model to provide an actual model integration. Using the same model setup as used for generating the model spectrum above, a homogeneous growth test (propagation terms switched off in Eq. (1)) is performed with a wind speed of 20 m/s and with an initial spectral peak frequency of 0.15 Hz. Results for the first 3 h of model integration with the default WAVEWATCH III model, which is based on a re-scaled DIA (Tolman and Chalikov, 1996), are presented in Fig. 5. The corresponding results obtained with the NNIA-E are presented in Fig. 6.

[5] i.e., the 12×12 grid 'out-of-the-box' test case.

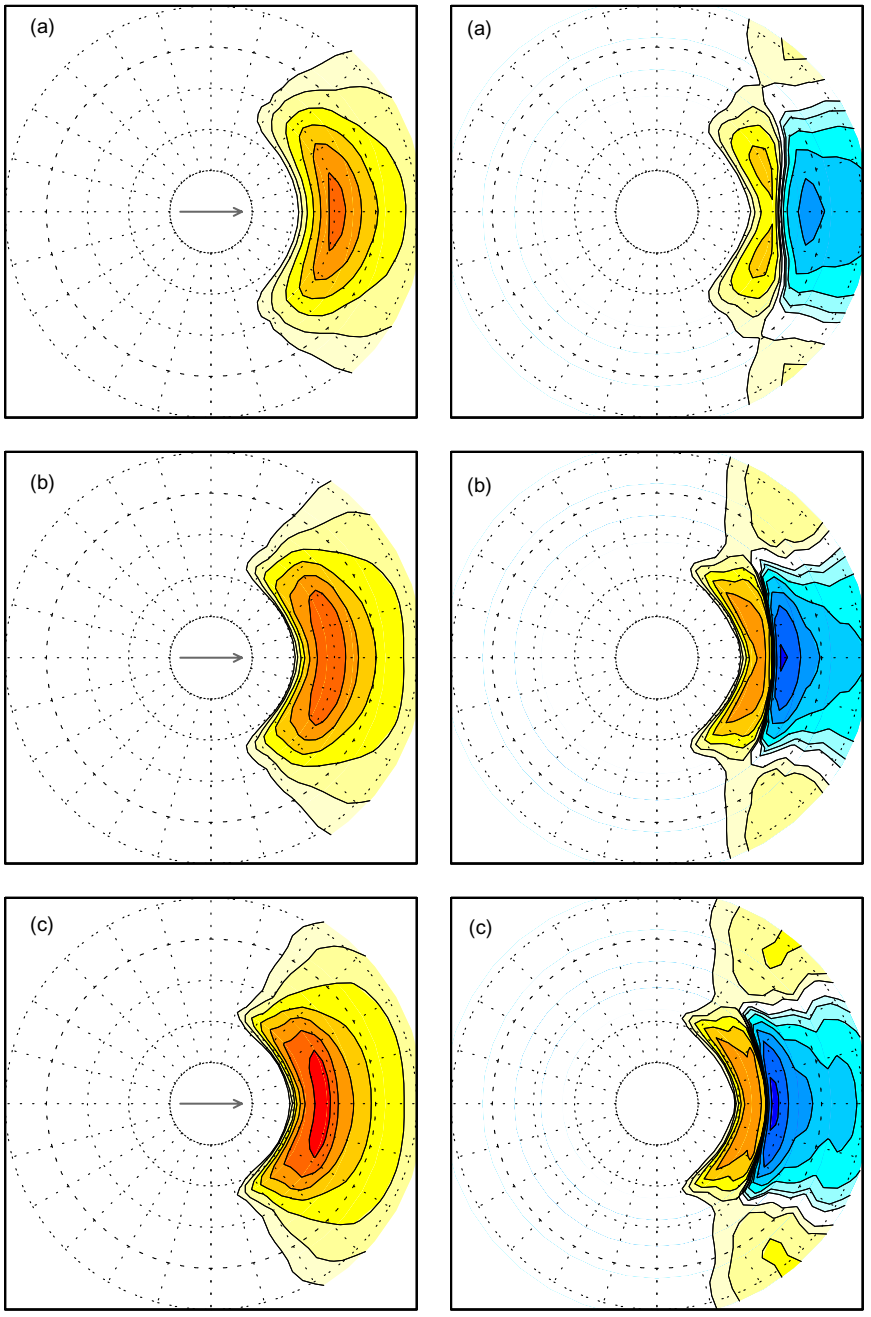


Fig. 5. Model spectra (left panels) and nonlinear interactions (right panels) after (a) 1 h, (b) 2 h and (c) 3 h calculated using the default WAVEWATCH III model (DIA). Spectral representation and legend as in Fig. 4. Lowest contour for spectrum 2×10^{-5} m^2 s. Lowest contours for interactions $\pm5\times10^{-5}$ m^2. The initial conditions consist of a JONSWAP spectrum with $f_p = 0.15$ Hz, $\gamma = 2$, and the Hasselmann et al. (1980) directional distribution. The wind is constant at 20 m/s.

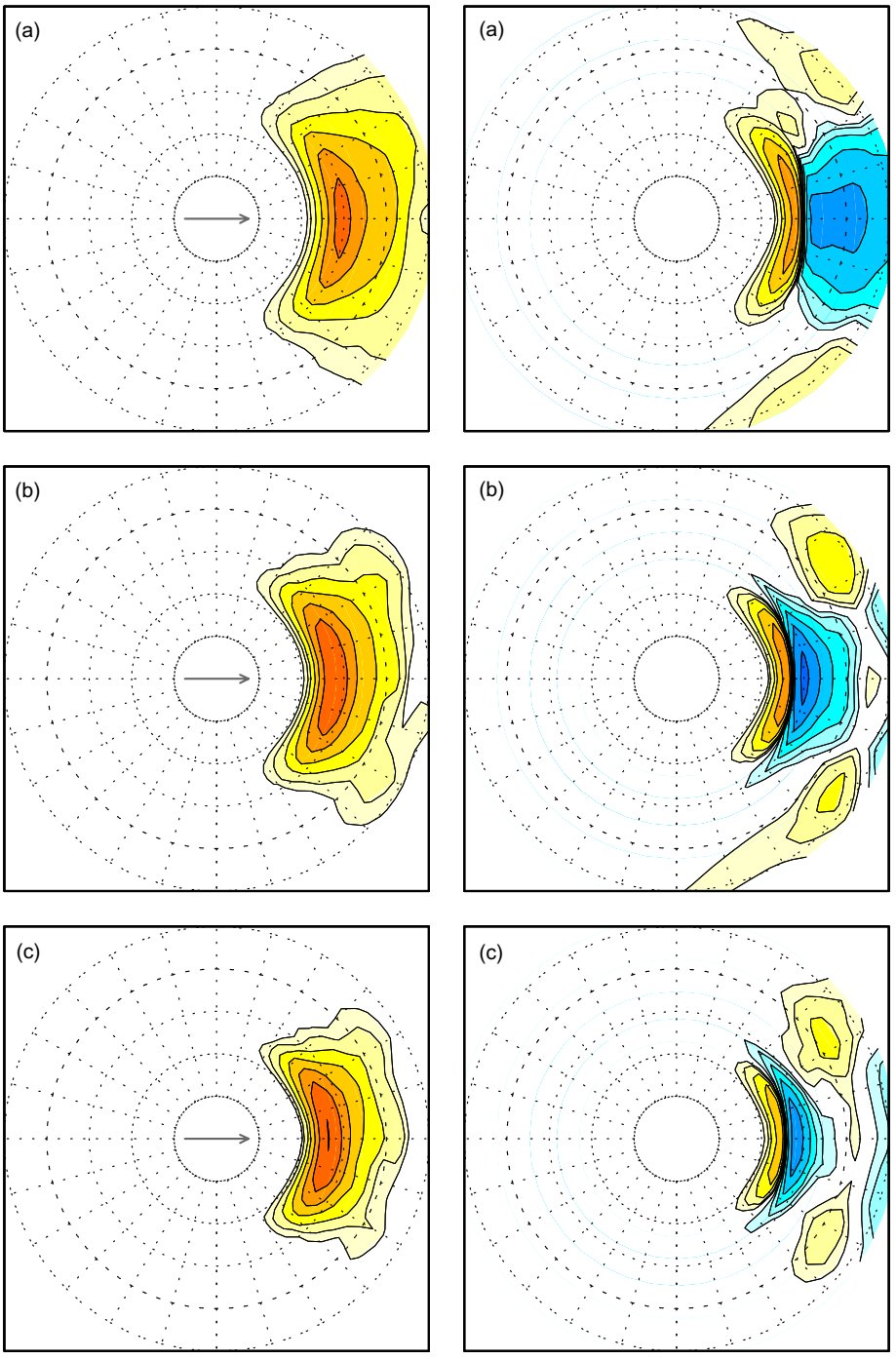


Fig. 6. Like Fig. 5 with the NNIA-E algorithm for nonlinear interactions.

The default WAVEWATCH III model (Fig. 5), shows the expected shift of the spectral peak to lower frequencies, with the corresponding increase energy near the spectral peak. During growth, a smooth and slow decrease of energy with increasing frequency is found for $f > f_p$ in the wind direction and adjacent directions. The model using the NNIA-E (Fig. 6), is not able to reproduce consistent wave growth. The two most obvious deficiencies are the fact the NNIA-E does not retain directional symmetry, and that the energy at high frequencies is suppressed (compared to the standard model results in Fig. 5). The latter significantly reduces the magnitude of the interactions, particularly after 3 h of model integration (compare Figs. 5c and 6c). Positive aspects of the integration experiment are the consistent shape of the spectral peak, and the retention of the sharp features of the nonlinear interactions for $f < f_p$. Considering the deficiencies of the NNIA-E when applied to model spectra shown in Fig. 4, the inability of this algorithm to produce realistic model integration is not surprising.

Finally, the relative costs of algorithms are important when applied in operational wave models. The DIA as implemented in WAVEWATCH III is highly optimized. Considering that the present NNIA-E is not a 'final product', we have not gone through the process of fully optimizing this algorithm. On NCEP's IBM SP2, the NNIA-E algorithm proved to be about 4.7 times more expensive than the DIA. A conservative estimate of the potential speed up of the NNIA-E due to straightforward optimization suggests that the NNIA-E needs to be no more expensive than about 2.5 times the costs of the DIA, with a potential of being even faster.

4. Discussion

This study presents a continuation of early NNIA experiments as reported in Krasnopolsky et al. (2002, KCT). This study represents the first part of a set of studies designed to develop a NNIA algorithm for practical wave models. In the present study, NNIA techniques for mapping nonlinear interactions for individual spectra are explored. Continuing the work presented in KCT, several more capable NNIAs have been developed by explicitly including scaling laws into the NNIA, by modifying decomposition methods, and by generating a more universal training data set. The first such NNIA (NNIA-M) uses the same basis functions as KCT (Legendre polynomials and Fourier components). Unlike the NNIA presented in KCT, the NNIA-M algorithm gives reasonable estimates of the nonlinear interactions for fairly arbitrary wind sea spectra. Its main drawback is that the NNIA-M generates spurious noise throughout spectral space, consistent with the Gibbs effect in Fourier analysis.

A second NN based interaction algorithm (NNIA-E) uses a decomposition of the spectrum and source terms in EOF. The NNIA-E is more accurate that the NNIA-M, probably due to the more convergent basis functions (Fig. 1). More importantly, it does not include spurious noise as displayed by the NNIA-M. For independent validation spectra not used in the training of the NN but belonging to the same envelope of conditions, the NNIA-E reduces errors by an order of magnitude compared to the DIA. This improvement is achieved at a conservatively estimated factor of 2.5 or less increase in computational costs for the NNIA-E compared to the DIA. This confirms the potential of an NNIA to eventually replace the DIA in practical wave models, and justifies the start of the second part of this study.

The NNIA-E can provide a very accurate estimate of instantaneous interactions, if the spectrum considered is well represented by the spectra used in the training. This is clearly demonstrated in Fig. 3. However, because only parametric spectra have been used in the training data set, it is unlikely that all expected spectral shapes are included. Modeled wave spectra can be expected to include features not found in the training data set. It is, therefore, not surprising that when applied to spectra generated by the numerical wave model WAVEWATCH III, the NNIA-E does not perform as well as with the above parametric spectra (Figs. 3 and 4). In this context, however, it is very encouraging to note that the NNIA-E results in fairly realistic interactions for modeled spectra, with several important features of the exact (WRT) interactions well represented (Fig. 4 and the corresponding discussion in the previous section).

A major remaining issue is the robustness of the NNIA-E when applied to model integration, or in other words, the generalization of the NNIA-E for truly arbitrary wind sea spectra. One way of making the NN more generally applicable is the use of a more sophisticated NN approach. In the present study we use the simplest NN possible (a feedforward, fully connected perceptron with a single hidden layer). It is well known that more advanced NN architectures and/or application of special techniques (e.g., Chen and Hagan, 1999) enhance generalization capabilities of NNs. An advanced generic machine learning approach (like Support Vector Machines) (Vapnik, 1995) can be used to generate optimal approximations in term of architecture and generalization ability. However, nonlinear interactions in wind waves represent a highly complex and highly nonlinear mechanism, with different behavior in different sub-domains (e.g., Hasselmann, 1963b; Hasselmann and Hasselmann, 1985). It is not realistic to expect that any more sophisticated NN approaches could enhance the capability of a NN to such a degree that it could generalize for physical processes not fully included in the training data set.

One obvious approach to make the NNIA-E suitable for model integration is to expand the present training data set to include features peculiar to spectra generated by wave models. Two such features that are not included in the present training data set are (i) the addition of a given parametric tail above a dynamically adjusted cut-off frequency and (ii) the fact that sharply peaked JONSWAP spectra tend to generate two minima in the negative lobe, whereas modeled spectra generally result in a single minimum. This suggests that there are subtle yet important inconsistencies between JONSWAP type spectra and the detailed source term balance in wind wave models. The former is likely to be the cause for some of the directional discrepancies between the results of the exact and NNIA-E algorithms at high frequencies in Fig. 4. The latter is possibly related to the secondary extrema in the NNIA at intermediate frequencies.

It is possible to attempt to add the above identified discrepancies to the parametric spectra of the training data set. This will undoubtedly lead to a more accurate and more generally applicable NNIA. However, including major discrepancies is also expected to highlight other discrepancies that have previously been overlooked. Expanding the present parametric training data set is, therefore, expected to result in a slowly converging process, if convergence to a general solution can be reached at all. On the other hand, the limited success of short term time integration of the NNIA-E in a wave model opens a more natural and potentially more powerful technique for expanding the training data set.

If the parametric spectra are used as the initial conditions for short time model integration, the resulting integrated spectra will naturally include spectral perturbations for which the NNIA needs to have accurate interactions in order to reproduce realistic spectral shapes. Alternatively, a

new training data set could be built entirely from modeled wave spectra. Although such approaches are also expected to require some iterations, the resulting training data sets will by definition form a true envelope of expected model spectra. With such an approach, we therefore expect to be able to generate an NNIA based on NNIA-E that will be accurate, and will be capable of resulting in realistic model integration for wind seas.

Other approaches are available to increase the robustness of the NNIA. One approach is the use of more advanced NN techniques (as discussed above). A second approach would be to attempt to estimate the accuracy of the resulting NNIA objectively, using an inverse NN (Krasnopolsky and Schiller, 2003). This, in principle, makes it possible to revert to alternative solutions (like a DIA) in cases where the NNIA is apparently inaccurate. If such alternatives are used sparingly, the model is expected to benefit from the accuracy and economy of the NNIA, while circumventing robustness issues. The development of such a hybrid NNIA furthermore provides a natural way to iteratively expand a training data set with spectra and source terms that are apparently not inside the previous envelope of conditions.

A third way of dealing with the robustness issue is to include more of the nonlinear physics into the NNIA. One approach could be to use some form of the DIA to generate basis functions for the nonlinear interactions. When combined with a NN, this in essence would result in a dynamically adjusted multiple DIA. As a preparation for designing such a hybrid NNDIA, the potential mapping accuracy of various modifications to the DIA is addressed in Tolman (2003, 2004).

5. Conclusions

Neural Network (NN) approximations for mapping of nonlinear interactions for wind waves for given spectra are discussed. A Neural Network Interaction Approximation (NNIA) for wind seas is presented. For fairly arbitrary single peaked spectra, this NNIA proved an order of magnitude more accurate than the conventional Discrete Interaction Approximation (DIA), at an estimated 2.5 times the computational cost or better. Because the parametric training data sets used to develop this NNIA do not include all peculiarities of modeled wave spectra, this NNIA cannot yet be used in operational wave models. Further research aimed at reaching the latter goal is discussed.

Acknowledgements

The authors would like to thank Henrique Alves, D.B. Rao, Mary Hart and the anonymous reviewers for their comments on early drafts of this manuscript. This study was supported by funding from the Office of Naval Research under grant N00014-00-F-0332, and by funding from the NOAA High Performance Computing and Communication (HPCC) office.

Appendix A. Training data sets

The parametric frequency spectrum on which the training data set is based is defined as

$$F(f) = \phi_1 \phi_2 \phi_3, \tag{A.1}$$

$$\phi_1 = \frac{ag^2}{(2\pi)^4 f_p^5}\left(\frac{f_p}{f}\right)^m, \quad \phi_2 = \exp\left[-\frac{m}{n}\left(\frac{f}{f_p}\right)^{-n}\right], \quad \phi_3 = \gamma^{\exp\left[-\frac{1}{2}\left\{\frac{f-f_p}{\sigma f_m}\right\}^2\right]},$$

where a is a proportionality constant, ϕ_1 describes the high-frequency flank of the spectrum, ϕ_2 describes the low-frequency flank of the spectrum, and ϕ_3 is the so-called peak-enhancement factor (see, e.g., Hasselmann et al., 1973); $m = 5$ and $n = 4$ gives the traditional JONSWAP spectrum. A full two-dimensional spectrum is obtained by adding a directional distribution $D(f, \theta)$:

$$F(f, \theta) = F(f) D(F, \theta), \tag{A.2}$$

where D is typically a slowly varying function of f. The Hasselmann et al. (1980) directional distribution (D_H) is defined as

$$D_H(f, \theta) = \cos^{2s}[0.5(\theta - \bar{\theta}(f))], \tag{A.3}$$

$$s = \alpha_1 \begin{cases} 6.97\left[1 + \alpha_2\left(\frac{f}{f_p} - 1\right)\right]^{4.06} & \text{for } f < 1.05 f_p, \\ 9.77\left[1 + \alpha_2\left(\frac{f}{f_p} - 1\right)\right]^{-2.34} & \text{for } f \geqslant 1.05 f_p, \end{cases} \tag{A.4}$$

for simplicity omitting the dependence of the bottom line of (A.4) on the wave age. The factors α_1 and α_2 are added to allow for the generation of an envelope of possible distributions, instead of a best fit ($\alpha_1 \equiv \alpha_2 \equiv 1$). The first factor influences the general width of the distribution, the second the rate of change of the width with f. The bimodal Ewans (1998) distribution (D_E) is given as

$$D_E(f, \theta) = \frac{1}{\sigma(f)\sqrt{8\pi}} \sum_{n=-\infty}^{\infty} \left\{ (1 \pm \alpha_3) \exp\left[-\frac{1}{2}\left(\frac{\theta - \bar{\theta} \pm \bar{\theta}_o(f) - 360n}{\sigma(f)} \right)^2 \right] \right\}, \tag{A.5}$$

$$\bar{\theta}_o(f) = \alpha_4 \begin{cases} 14.93 & \text{for } f < f_p, \\ \exp\left[5.453 - 2.750\left(\frac{f}{f_p}\right)^{-1}\right] & \text{for } f \geqslant f_p, \end{cases} \tag{A.6}$$

$$\sigma(f) = \alpha_5 \begin{cases} 11.38 + 5.357\left(\frac{f}{f_p}\right)^{-7.929} & \text{for } f < f_p, \\ 32.13 - 15.39\left(\frac{f}{f_p}\right)^{-2} & \text{for } f \geqslant f_p, \end{cases} \tag{A.7}$$

where $\bar{\theta}_o(f)$ is an offset direction, with both the positive and negative values added in the summation. The factors α_3–α_5 are added again to allow for the generation of an envelope of possible distributions. α_3 allows for unequal branches of the bimodal distribution, α_4 controls the separation of the branches of the bimodal distribution, and α_5 controls the general width of the directional distribution. Note that in Eqs. (A.5)–(A.7) all directions are expressed in degrees, whereas Ewans writes (A.5) in radians and the others in degrees. The final directional distribution becomes a linear combination of the two:

$$D(f, \theta) = \alpha_6 D_H(f, \theta) + (1 - \alpha_6) D_E(f, \theta). \tag{A.8}$$

The directional shear is defined as the change of mean direction $\bar{\theta}$ in the directional distribution D as a function of the frequency f. For wind seas, the spectrum at high frequencies tends to line

up with the wind direction. For lower frequencies, it may deviate progressively. A realistic shear can, therefore, be obtained by using a mean direction that is a function of $f^{-\beta}$, where β is a positive, tunable power. Defining the shear as 0 at f_p, and as θ_s at some high frequency f_c, the sheared mean direction becomes

$$\bar{\theta}(f) = \frac{\theta_s\left(f^{-\beta} - f_p^{-\beta}\right)}{|f_p^{-\beta} - f_c^{-\beta}|}. \tag{A.9}$$

Taking f_c relative to f_p, say $f_c = 3f_p$, this can be rewritten as

$$\bar{\theta}(f) = \frac{\theta_s}{1 - 3^{-\beta}}\left[\left(\frac{f}{f_p}\right)^{-\beta} - 1\right], \tag{A.10}$$

where β and θ_s determine the magnitude of the shear. Finally, random modulations of the spectrum are applied with a given normalized relative change σ_r, and with correlation length scales σ_f and σ_θ in frequency and direction space, respectively.

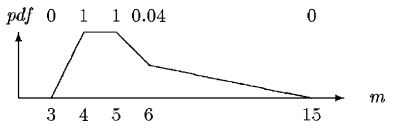
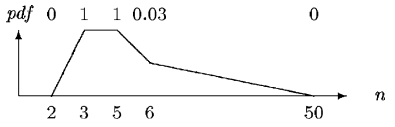
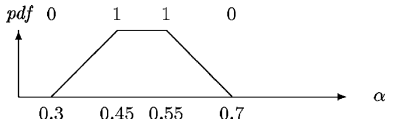
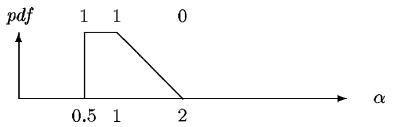
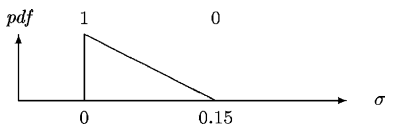
A set of test spectra is generated using a spectral discretization with 36 directions ($\Delta\theta = 10°$), and 31 frequencies ranging from 0.475 through 3.62 (normalized). Individual spectra are obtained by generating Monte Carlo realizations of all relevant spectral parameters.

The parameters f_p, a, γ, m, n and σ can be varied to obtain an envelope of possible spectra in Eq. (A.1). The peak frequency f_p and energy level a are set to $f_p = 0.1$ Hz, and $a = 0.0081$. Due to the normalization of spectra within the NNIA, these values, in essence, are irrelevant. The justification for the range of parameters used in the training data set is as follows:

- The peak frequency f_p is varied normally over an interval Δf to account for the fact that the peak is not always lined up with the discrete frequency grid.
- The high-frequency spectral slope parameter m is generally expected to be 4–5. Lower values are not likely to occur. Higher values are expected in conditions of diminishing winds where the spectrum remains coherent, but energy dissipation at high frequencies occurs much faster than at low frequencies.
- The low-frequency spectral slope parameter n is generally taken as approximately 4, although early observations of fully grown conditions quote values as low as 2. Particularly low values are of interest here to describe near full-grown conditions with strong directional shear, as occur regularly in conditions with rapid wind shifts. Occasional very high values are required as steep low-frequency parameterizations are used regularly in model (re-)initialization (for instance the 'seeding' algorithm used in WAVEWATCH III, Tolman, 2002).
- The peak-enhancement factor γ ranges from 1 for full-grown conditions to about 10 for young wind seas. To properly address conditions where spectra are still coherent but overdeveloped, more conditions with low γ are introduced.
- The directional spread parameter σ for the spectral peak enhancement has a small radius of influence, typically of the order of 0.05–0.07. The details of this parameter will literally be 'lost in the random noise'. Therefore, a constant value $\sigma = 0.07$ is used.

The tunable parameters α_1–α_6 in the directional distributions (A.3)–(A.8) are introduced to create an envelope around best-fit parameterizations as identified in Table 4 are subjective, but require no additional justification. Similarly, the directional shear parameters β and θ_s in

Table 4
Distributions of random parameters of the parametric spectral description as used in the generation of the training data set

Parameters	Descriptions	Eq.	Distribution
f_p	$F(f)$	(A.1)	Uniform over 0.1 Hz±0.5Δf
γ			25% $\gamma = 1$, 75% uniform over 1–10
m			
n			
α_3	$D(f,\theta)$	(A.5)	
α_6		(A.8)	Uniform over 0–1
Other α		(A.4)–(A.7)	
β	$\bar{\theta}(f)$	(A.10)	Like α_1 with $\beta = 0, 2, 4$
θ_s			Like α_3 with $\theta_s = -45°, -15°, 15°, 45°$
σ_f	Noise	–	Uniform over 0.01–0.1 Hz
σ_θ			Uniform over $\Delta\theta$–45°
σ_r			

Graphical depiction of piecewise linear probability density functions (pdf) qualitative only. Numbers across top of these

Eq. (A.10) are subjectively chosen. Note that fairly strong shears are included to represent conditions with rapidly changing wind directions, as often occur near fronts. Finally, the same is true for the parameters describing the random fluctuations. Note that triangular distribution was chosen to focus more on cases with moderate noise, while including some cases with significant noise.

References

Abdalla, S., Ozhan, E., 1993. Third-generation wind–wave model for use on personal computers. Journal of Waterway, Port, Coastal and Ocean Engineering 119, 1–14.

Abramowitz, M., Stegun, I.A., 1964. Handbook of Mathematical Functions. National Bureau of Standards. 1046 pp.

Allender, J.H., Barnett, T.P., Lybanon, M., 1985. SWAMP group (1985). In: The DNS model. An Improved Spectral Model for Ocean Wave Prediction. Plenum Press, New York, pp. 235–248 (Chapter 15).

Booij, N., Ris, R.C., Holthuijsen, L.H., 1999. A third-generation wave model for coastal regions, Part I, model description and validation. Journal of Geophysical Research 104, 7649–7666.

Chen, D., Hagan, M., 1999. Optimal use of regularization and crossvalidation in neural network modeling. In: Proceedings International Joint Conference on Neural Networks. Paper No. 323.

Ewans, K.C., 1998. Observations of the directional spectrum of fetch-limited waves. Journal of Physical Oceanography 28, 495–512.

Hashimoto, N., Kawaguchi, K., 2001. Extension and modification of Discrete Interaction Approximation (DIA) for computing nonlinear energy transfer of gravity wave spectrum. In: Ocean Wave Measurement and Analysis. ASCE, New York, pp. 530–539.

Hasselmann, K., 1960. Grundgleichungen der seegangsvoraussage. Schiffstechnik 1, 191–195.

Hasselmann, K., 1962. On the non-linear transfer in a gravity wave spectrum, Part 1. General theory. Journal of Fluid Mechanics 12, 481–500.

Hasselmann, K., 1963a. On the non-linear transfer in a gravity wave spectrum, Part 2. Conservation theory, wave–particle correspondence, irreversibility. Journal of Fluid Mechanics 15, 273–281.

Hasselmann, K., 1963b. On the non-linear transfer in a gravity wave spectrum, Part 3. Evaluation of energy flux and sea-swell interactions for a Neuman spectrum. Journal of Fluid Mechanics 15, 385–398.

Hasselmann, S., Hasselmann, K., 1985. Computations and parameterizations of the nonlinear energy transfer in a gravity-wave spectrum, Part I: a new method for efficient computations of the exact nonlinear transfer integral. Journal of Physical Oceanography 15, 1369–1377.

Hasselmann, K., Barnett, T.P., Bouws, E., Carlson, H., Cartwright, D.E., Enke, K., Ewing, J.A., Gienapp, H., Hasselmann, D.E., Kruseman, P., Meerburg, A., Mueller, P., Olbers, D.J., Richter, K., Sell, W., Walden, H., 1973. Measurements of wind–wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP). Ergaenzungsheft zur Deutschen Hydrographischen Zeitschrift, Reihe A 12 (8), 95 pp.

Hasselmann, D.E., Dunckel, M., Ewing, J.A., 1980. Directional wave spectra observed during JONSWAP 1973. Journal of Physical Oceanography 10, 1264–1280.

Hasselmann, S., Hasselmann, K., Allender, J.H., Barnett, T.P., 1985. Computations and parameterizations of the nonlinear energy transfer in a gravity-wave spectrum, Part II: parameterizations of the nonlinear energy transfer for application in wave models. Journal of Physical Oceanography 15, 1378–1391.

Herterich, K., Hasselmann, K., 1980. A similarity relation for the nonlinear energy transfer in a finite-depth gravity-wave spectrum. Journal of Fluid Mechanics 97, 215–224.

Jolliffe, I., 1986. Principle Component Analysis. Springer Verlag, New York.

Komen, G.J., Cavaleri, L., Donelan, M., Hasselmann, K., Hasselmann, S., Janssen, P.E.A.M., 1994. Dynamics and Modelling of Ocean Waves. Cambridge University Press, Cambridge. 532 pp.

Krasnopolsky, V.M., Chalikov, D.V., Tolman, H.L., 2001. A neural network technique to improve computational efficiency of environmental numerical models. Tech. Note 199, NOAA/NWS/NCEP/OMB, 37 pp.

Krasnopolsky, V.M., Schiller, H., 2003. Some neural network applications in environmental sciences part I: forward and inverse problems in geophysical remote measurements. Neural Networks 16, 321–334.

Krasnopolsky, V.M., Chalikov, D.V., Tolman, H.L., 2002. A neural network technique to improve computational efficiency of numerical oceanic models. Ocean Modelling 4, 363–383.

Lorenz, E.N., 1956. Empirical orthogonal functions and statistical weather prediction. Sci. Rep. 1, MIT, Cambridge, Statistical Forecasting Project, 48 pp.

Nguyen, D., Widrow, B., 1990. Improving the learning speed of 2-layer Neural Networks by choosing initial values of the adaptive weights. In: Proceedings International Joint Conference on Neural Networks, vol. 3. pp. 21–26.

Phillips, O.M., 1960. On the dynamics of unsteady gravity waves of finite amplitude. Journal of Fluid Mechanics 9, 193–217.

Phillips, O.M., 1981. Wave interactions: the evolution of an idea. Journal of Fluid Mechanics 106, 215–227.

Pierson, W.J., Moskowitz, L., 1964. A proposed spectral form for fully developed wind seas based on the similarity theory of S.A. Kitaigorodskii. Journal of Geophysical Research 69, 5181–5190.

Resio, D.T., Perrie, W., 1991. A numerical study of nonlinear energy fluxes due to wave–wave interactions. Part 1: methodology and basic results. Journal of Fluid Mechanics 223, 609–629.

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning internal representations by error propagation. In: Rumelhart, D.E., McLelland, J.L. (Eds.), Parallel Data Processing, vol. I. MIT Press, Cambridge, MA, pp. 318–362.

Snodgrass, F.E., Groves, G.W., Hasselmann, K.F., Miller, G.R., Munk, W.H., Powers, W.H., 1966. Propagation of swell across the pacific. Transactions of the Royal Society of London A 259, 431–497.

SWAMP group, 1985. Ocean Wave Modelling. Plenum Press, New York. 256 pp.

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 2.22. Technical note 222, NOAA/NWS/NCEP/MMAB, 133 pp.

Tolman, H.L., 2003. Optimum Discrete Interaction Approximations for wind waves. Part 1: mapping using inverse modeling. Technical note 227, NOAA/NWS/NCEP/MMAB, 57 pp. + Appendices.

Tolman, H.L., 2004. Inverse modeling of Discrete Interaction Approximations for nonlinear interactions in wind waves. Ocean Modelling 6, 405–422.

Tolman, H.L., Chalikov, D.V., 1996. Source terms in a third-generation wind–wave model. Journal of Physical Oceanography 26, 2497–2518.

Tolman, H.L., Balasubramaniyan, 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 and Forecasting 17, 311–333.

Tracy, B., Resio, D.T., 1982. Theory and calculation of the nonlinear energy transfer between sea waves in deep water. WES Report 11, US Army Corps of Engineers.

Van Vledder, G.P., 2002a. Improved parameterizations of nonlinear four wave interactions for application in operational wave prediction models. Report 151a, Alkyon, The Netherlands.

Van Vledder, G.P., 2002b. A subroutine version of the Webb/Resio/Tracy method for the computation of nonlinear quadruplet interactions in a wind–wave spectrum. Report 151b, Alkyon, The Netherlands.

Van Vledder, G.P., Dee, D., 1994. Validation document of the third generation wave prediction model PHIDIAS. Report H1861, Delft Hydraulics.

Van Vledder, G.P., Herbers, T.H.C., Jenssen, R.E., Resio, D.T., Tracy, B., 2000. Modelling of non-linear quadruplet wave–wave interactions in operational wave models. In: Proceedings 27th International Conference on Coastal Engineering, Sydney, Australia. ASCE, New York, pp. 797–811.

Vapnik, V.N., 1995. The Nature of Statistical Learning Theory. Springer Verlag, New York. 288 pp.

WAMDIG, 1988. The WAM model—a third generation ocean wave prediction model. Journal of Physical Oceanography 18, 1775–1809.

Webb, D.J., 1978. Non-linear transfers between sea waves. Deep-Sea Research 25, 279–298.

Young, I.R., Van Vledder, G.P., 1993. A review of the central role of nonlinear interactions in wind–wave evolution. Transactions of the Royal Society of London 342, 505–524.