THE AIR-SEA INTERFACE
Radio and Acoustic Sensing, Turbulence and Wave Dynamics

Edited by

M. A. Donelan
Rosenstiel School of Marine and Atmospheric Science
University of Miami
Miami, Florida, USA

W. H. Hui
Department of Mathematics
Hong Kong University of Science and Technology
Kowloon, Hong Kong

W. J. Plant
Applied Physics Laboratory
University of Washington
Seattle, Washington, USA

Rosenstiel School of Marine and Atmospheric Science
University of Miami, Miami, Florida

WAVE MODEL FITTING USING THE ADJOINT TECHNIQUE

G. Barzel, R. B. Long, S. Hasselmann and K. Hasselmann
Max-Planck Institut für Meteorologie, Hamburg, Germany

Abstract. The global wave model WAM solves the transport equation for wind generated waves. The model performance can be improved by adjusting free parameters in the source terms against data. The adjoint technique was used to assimilate data obtained from growth rates into a one-dimensional WAM model. Two sets of data have been used: One from the JONSWAP experiment, another from a reanalysis of fetch-limited growth data by Kahma and Calkoen.

The growth rates computed with the new values for the free parameters show, in contrast to the WAM model (cycle 3 and 4), an almost linear energy growth for small fetches which gives a better agreement with data.

1 Introduction

The purpose of this paper is to demonstrate the use of data assimilation by the adjoint method to improve the performance of a numerical wave model, specifically, a one-dimensional version of the WAM cycle 4 model currently being run operationally at ECMWF. The model is fitted to observations consisting of experimentally determined growth curves using the adjoint procedure, and improved parameter sets for the source terms of the wave model are derived.

In Part 1 below, we outline the operation of the adjoint method and present some technical remarks on how we implemented it in this application. In Part 2, the basic equations of the wave model are presented, the cost function quantifying the model/data misfit is defined, and the data briefly described. Part 3 discusses results obtained for different sets of free parameters in the model source terms for wind input and wave dissipation. Finally, Part 4 offers some closing remarks.

2 What is an 'Adjoint Model'?

2.1 The Basic Idea

Data assimilation by the adjoint method operates by minimizing a cost function $\mathcal{J}$ which measures the mismatch between available observations (data) and their model counterparts. Any quantity that is measurable and can also be computed from model results may be used as data, e.g., in the case of wave models, total energy, peak frequency, one or two dimensional wave spectra, etc. The cost function is generally defined as $\mathcal{J} = (\boldsymbol{d}^{\mathrm{obs}} - \boldsymbol{d}^{\mathrm{mod}})^T \boldsymbol{D}^{-1} (\boldsymbol{d}^{\mathrm{obs}} - \boldsymbol{d}^{\mathrm{mod}})$ where $\boldsymbol{d}^{\mathrm{obs}}$, $\boldsymbol{d}^{\mathrm{mod}}$ are vectors of observed data and their model equivalents, respectively, and $\boldsymbol{D}$ is the observation error covariance matrix (or some appropriate approximation thereof).

For the purpose of exposition, we can regard the model as a numerical algorithm that starts with a set (for notational convenience, a vector) of control variables $\boldsymbol{\Phi} = \{\phi_j\}_{j=1,\cdots,M}$, then evaluates a set of intermediate variables

$\Psi = \{\psi_i\}_{i=1,\cdots,N}$, then terminates in a value for the cost function $\mathcal{J}$. The ϕ_j may be initial conditions, the wind field, or any kind of input parameters, such as, e.g., free parameters in the source terms. The intermediate variables ψ_i are taken to be indexed by their subscripts i in the order in which they are evaluated in the numerical algorithm.

Through the model equations, all intermediate variables and the cost function are completely determined by the values assumed by the control variables ϕ_j. We seek the minimum of $\mathcal{J}$ in the space spanned by these variables. This occurs where the gradient $\nabla_k \mathcal{J} = \frac{\partial}{\partial \phi_k} \mathcal{J}(\Phi) = 0$. Typically, the dependence of $\mathcal{J}$ on the ϕ_k is expressed in the model through the hierarchy of intermediate variables ψ_i, making it difficult to compute the components of $\nabla_k \mathcal{J}$ directly. This difficulty is avoided by elevating the intermediate variables to independent status, then seeking a constrained minimum in this much larger space. The constraints are imposed by constructing a *Lagrange* function in which the model equations $\psi_i = E(\Phi, \{\psi_j\}_{j<i})$ appear as constraints, each multiplied by a Lagrange undetermined multiplier λ_i:

$$\mathcal{L}(\lambda, \Psi, \Phi) = \mathcal{J} - \lambda^T (\Psi - \mathrm{E}) = \mathcal{J} - \sum_{i=1}^{N} \lambda_i \left(\psi_i - E_i \left(\Phi, \{\psi_j\}_{j<i} \right) \right) \quad . \quad (1)$$

The saddle point of the Lagrange function in the space spanned by the control variables, intermediate variables, and Lagrange multipliers can be shown to be the same as the minimum of the cost function in the space of the control variables alone. We find the saddle point by setting partial derivatives of $\mathcal{L}(\lambda, \Psi, \Phi)$ with respect to each of its indicated arguments to zero, then solving the resulting equations simultaneously. The vanishing of partial derivatives with respect to the λ_i yields the model itself as the first set of equations to be satisfied. The vanishing of partial derivatives with respect to the intermediate variables ψ_i yields the so-called adjoint equations,

$$\frac{\partial \mathcal{L}}{\partial \psi_N} = 0 \rightarrow \lambda_N = \frac{\partial \mathcal{J}}{\partial \psi_N} \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial \psi_k} = 0 \rightarrow \lambda_k = \frac{\partial \mathcal{J}}{\partial \psi_k} + \sum_{i=k+1}^{N} \lambda_i \frac{\partial E_i}{\partial \psi_i} \quad \text{for } k < N \quad (2)$$

Finally, taking partial derivatives with respect to the control variables yields the gradient equations,

$$\frac{\partial \mathcal{L}}{\partial \phi_k} = \nabla_k \mathcal{J} = \frac{\partial \mathcal{J}}{\partial \phi_k} + \sum_{i=1}^{N} \lambda_i \frac{\partial E_i}{\partial \phi_k} \quad , \quad (3)$$

(the first equality holding under the condition that the model and adjoint equations are satisfied). These must also vanish at the minimum of the cost. In practice, values of the control variables are estimated and the model and adjoint equations solved; the results are then used in the gradient equations to compute the gradient of the cost at the assumed values of the controls. This is then used in a gradient-descent algorithm to improve the estimate of the controls and to iteratively seek out the cost minimum.

The collection of adjoint equations is sometimes referred to as the 'adjoint model' because it superficially resembles a linearized version of the model itself. Note that the 'adjoint model' in a sense runs backward, starting with the evaluation of the Lagrange multiplier λ_N, related to the last calculated model variable ψ_N, and then calculating the $\lambda_{i<N}$ recursively. Because of the role played by the derivatives of the cost function in the adjoint equations, the adjoint model appears to be 'driven' by the mismatch between observations and their model counterparts.

2.2 Implementation of an Adjoint Model

Constructing the adjoint of a complicated numerical model can be a tedious task. In the past, the usual approach has been to identify the intermediate variables ψ_i with the state variables of the numerical model. The partial derivatives appearing in the adjoint and gradient equations were then computed analytically, using the chain rule as necessary, and the resulting expressions converted into computer code.

In contrast, our approach has been to construct the adjoint code directly from the model code, treating every assignment statement in the model code as a constraint equation defining a model variable (excluding statements defining constants, flags, input/output parameters, and other quantities which are not variables of the model). Thus, every variable has its multiplier and can be treated as independent in taking the required partial derivatives. In outline, the procedure is as follows:

Knowing that the multipliers must be evaluated in the reverse order to that of the model variables, we erect the framework of the adjoint code from beginning to end by tracing backward through the model code from end to beginning, generating an adjoint subroutine, branch, and loop for every such structure in the model. The adjoint subroutines will be called and the branches and loops traversed in reverse order to that in the model. Within these structures, individual lines of adjoint code are generated from the corresponding line of model code by regarding each assignment statement in the model code as a constraint to be multiplied by a Lagrange multiplier identified with the left-hand-side variable. The right-hand-side is then differentiated with respect to each of the intermediate and control variables appearing there, treating all others as independent in the process. Each such derivative represents one of the terms in the summation appearing in the adjoint or gradient equations derived above, whence is generated a line of adjoint code which accumulates this contribution into the appropriate sum. Care must be taken whenever the assignment statement is not strictly a mathematical equation (e.g., when it has the form $x = x + \ldots$), and it is necessary to initialize the variables in which the contributions are to be accumulated at the appropriate places in the code. But there are no fundamental difficulties in carrying out this program, and the resulting code is clean, mirrors the structure of the model, and exhibits a one-to-few correspondence between individual lines of code in the model and in the adjoint that makes maintenance and modification of the code easy and uncomplicated. The procedure is so straightforward that it is possible to automate it, so that adjoint code could

be generated by the computer. An 'Adjoint Model Compiler' which operates on similar principles has been developed at MPI (Giering and Maier-Reimer, 1996).

For the present application, once the model and adjoint routines were written they were linked into an umbrella subroutine (the model), which returns a value of the cost function when called, and another (the adjoint), which returns a value for the cost gradient when called. These were then linked with NAG library routine E04DGE, which uses the cost function and gradient evaluations in a sophisticated, iterative gradient-descent algorithm to find the values of the control variables that minimize the cost.

3 The Model Physics and the Data

3.1 Basic Equations of the Model

The global third generation wave model WAM is described in WAMDI Group (1988), Komen *et al.* (1994) and Günther *et al.* (1992). Thus we give only the equations relevant to understanding the optimization we performed. The basic equation to be solved by the wave model is the energy transport equation:

$$\frac{D}{Dt}F = \frac{\partial F}{\partial t} + \vec{c}_g \vec{\nabla} F = S_{\text{in}} + S_{\text{ds}} + S_{\text{nl}} + \cdots \quad \text{with} \quad F = F(\vec{k}; \vec{x}, t) \qquad (4)$$

where F is the 2-D wave spectrum and $\vec{c}_g$ is the group velocity. We use a so-called 'one-dimensional' version which considers only the fetch-limited case, i.e. a constant wind blowing offshore perpendicular to a straight coastline. The water is assumed to be deep and the wind constant both in direction and magnitude.

For the integration of eq.(4) we use an implicit finite-difference scheme (WAMDI Group, 1988). The source terms on the right hand side are: the input by the wind, S_{in}, the dissipation due to whitecapping, S_{ds}, and the nonlinear transfer between resonant waves, S_{nl}. In the implementation of WAM cycle 4 the following expressions are used (note that the wind input (Janssen, 1991) takes into account the effect of the waves on the wind, and the dissipation (original form: Komen *et al.* (1984)) has been adjusted to balance this wind input):

$$S_{\text{in}}(f,\theta) = \max[0, \gamma_{\text{in}}(u_*,\theta)]\, \omega\, F(f,\theta)\,; \quad S_{\text{ds}}(f,\theta) = -\gamma_{\text{ds}}(\mathcal{E},\bar{\omega})\, \omega\, F(f,\theta) \qquad (5)$$

with

$$\gamma_{\text{in}}(u_*,\theta) = W_M \frac{\beta}{\kappa^2} \frac{\rho_{\text{a}}}{\rho_{\text{w}}} \mu\, (\ln(\mu))^4 \max\left[0, \left(\frac{u_*}{v_{\text{ph}}} + W_D u_c\right) \cos(\theta - \theta_{\text{w}})\right]^2 \qquad (6)$$

$$\gamma_{\text{ds}}(\mathcal{E},\bar{\omega}) = D_M\, c_{\text{ds}} \left(\frac{\alpha}{\alpha_{\text{PM}}}\right)^{2D_S} \left(\left(1 - \frac{D_F}{2}\right)\left(\frac{\omega}{\bar{\omega}}\right)^2 + \frac{D_F}{2}\left(\frac{\omega}{\bar{\omega}}\right)^4\right) \qquad (7)$$

and $\omega = 2\pi f$, ρ_{a} (ρ_{w}) is the density of air (water), v_{ph} is the phase velocity, f, θ denote frequency and direction in k-space, θ_{w} is the wind direction. In the wind input term $\kappa = 0.41$ is von Karman's constant, μ is a function depending

on the roughness length, z_0, and $\beta = 1.2$ and $u_c = 0.011$ have been obtained by tuning (Janssen, 1991). The mean frequency, $\bar{\omega}$, is calculated as inverse of the mean period, and $\alpha = \omega^4 \mathcal{E}/g^2$ is an integral wave steepness parameter. $\alpha_{\rm PM}$ is the theoretical value of α for a Pierson-Moskowitz spectrum, $\mathcal{E}$ is the total energy and g the gravitational constant. $c_{\rm ds}$ is a constant obtained by tuning.

We have introduced five free parameters into the source terms for wind input and dissipation. Two of them, W_M and D_M, serve as overall coefficients that balance these two source terms. In addition, W_D influences the directional term of the wind input while D_S and D_F determine the functional form of the dissipation function. Setting all parameters to 1.0 yields the original WAM cycle 4 forms.

The nonlinear transfer is represented in the model by the so-called 'discrete interaction approximation' which provides a good estimate of the exact Boltzmann integral expression (Hasselmann *et al.*, 1985).

3.2 Data and Cost Function

As data we use dimensionless energy, $\tilde{\mathcal{E}} = (g^2/u_{10}^4)\mathcal{E}$, peak frequency $\tilde{f}_p = (u_{10}/g)f_p$, and Phillips' parameter α. For nondimensionalizing we use the wind at 10m height, u_{10}, since the friction velocity, u_*, is a model output, hence no longer suitable as a scaling parameter. In the growth region the data are obtained from JONSWAP fetch laws (Hasselmann *et al.*, 1973) and reanalysis of several fetch limited experiments by Kahma and Calkoen (1992). For the fully developed sea we use the data given by Pierson and Moskowitz (1964) in the nondimensional form obtained by scaling with u_* (Komen *et al.*, 1984). These values are made dimensional again with u_*, calculated from the Wu formula, and then rescaled by u_{10}. The cost function is defined as:

$$\mathcal{J}(\tilde{\mathcal{E}}, \tilde{f}_p, \alpha) = \frac{1}{W} \sum_{i=1}^{N_{\rm obs}} w_i^E \left(\frac{\tilde{\mathcal{E}} - \tilde{\mathcal{E}}_{\rm obs}}{\tilde{\mathcal{E}}_{\rm obs}} \right)^2 + w_i^f \left(\frac{\tilde{f}_p - \tilde{f}_{p,\rm obs}}{\tilde{f}_{p,\rm obs}} \right)^2 + w_i^\alpha \left(\frac{\alpha - \alpha_{\rm obs}}{\alpha_{\rm obs}} \right)^2 . \tag{8}$$

where $W = \sum_i [w_i^E + w_i^f + w_i^\alpha]$.

4 Results of the Optimization

Table I shows the results of several assimilations in which different sets of parameters were varied and different growth curves used as data. Benchmark values of the cost were computed by running the model with all parameters at their WAM cycle 4 values (1.0) and computing the cost using growth curves from JONSWAP and from Kahma and Calkoen (1992). The cost is less for the JONSWAP data than for Kahma and Calkoen because the model was originally tuned to the JONSWAP observations. In the first experiment, we used only the coefficients D_M and W_M in eqs.(6), (7) as free parameters, obtaining only a slight reduction of the cost. This was also true when only the wind parameters W_M and W_D were varied. Varying all three dissipation parameters or all five wind and dissipation parameters together gave cost reductions of about 30% and, in the case of the data summarized by Kahma and Calkoen (1992), param-

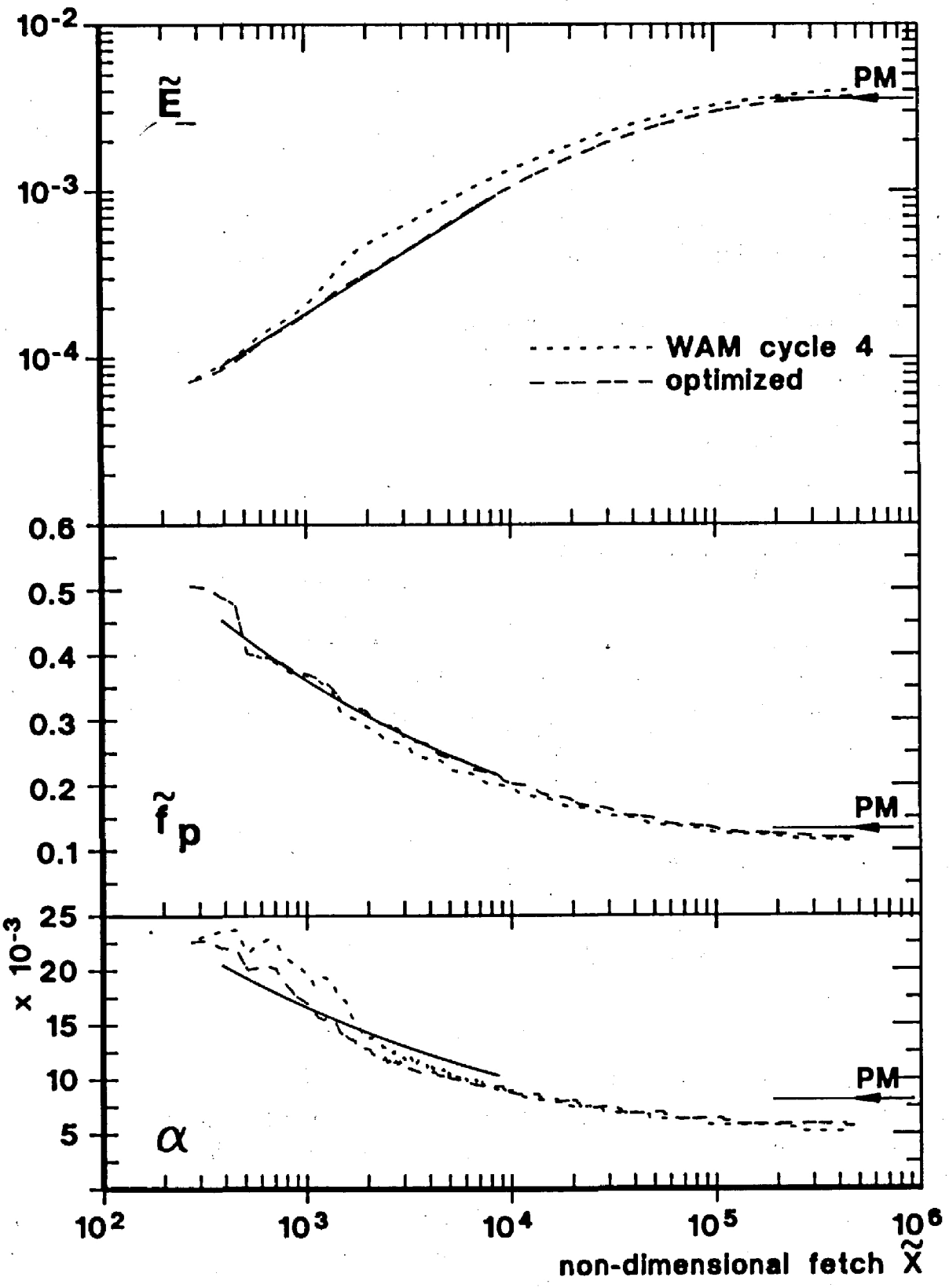

Fig. 1. Energy and peak frequency nondimensionalized with u_{10} and Phillips' parameter α as functions of the dimensionless fetch $\tilde{x}$. Solid lines indicate the 'observation' points obtained by using Kahma and Calkoen's fetch laws in the growth region and Pierson-Moskowitz data for the fully developed sea.

eter values that are significantly different from their nominal values of 1.0. Fig. 1 shows the improvement we obtained for the five-free-parameter assimilation using the Kahma and Calkoen formula in the growth region. Note that the large change in the energy growth is due to Kahma and Calkoen's data, the original parameters in the source terms having been tuned to JONSWAP fetch laws.

TABLE I
Results of optimizations for different parameter sets.

	Dissipation			Wind input		Cost	
	D_M	D_F	D_S	W_M	W_D	JONSWAP	Kahma & Calkoen
Original	1.00	1.00	1.00	1.00	1.00	0.0311	0.0533
D_M, W_M	1.04	1.00	1.00	0.926	1.00	0.0272	—
W_M, W_D	1.00	1.00	1.00	0.932	0.923	0.0250	—
D_M, D_F, D_ε	1.05	0.995	1.14	1.00	1.00	0.0202	—
All (JONSWAP)	1.03	0.997	1.07	0.949	0.976	0.0228	—
All (Kahma & Calkoen)	1.06	0.721	1.13	0.878	1.05	—	0.0095

5 Conclusions and Outlook

While the adjoint method is not new, the procedure by which we have constructed the adjoint code is new. By assigning a Lagrange multiplier to every variable defined in the model code (which avoids the necessity for frequent and potentially painful application of the chain rule), it becomes possible to construct the adjoint code, line by line, directly from the model code. The procedure is straightforward and produces adjoint code that is as transparent as the model code, and is easy to maintain or modify. It also lends itself to automatic generation of adjoint code; a prototype 'adjoint model compiler' based on similar principles already exists at MPI and was used to verify some of our manually developed programs.

In this paper, we have demonstrated the application of the adjoint method for the fine-tuning of a complicated numerical wave prediction model. Improved parameterizations of the wind input and wave dissipation source terms were obtained. These new parameters will be tested in the global WAM model and, if results justify it, adopted in a future cycle.

Integrated quantities such as total wave energy and peak frequency contain a limited amount of information. Therefore, the next steps will include the use of more detailed data (2-D spectra) and the extension of the adjoint to two space dimensions plus time.

References

Günther, H., Hasselmann, S., and Janssen, P. A. E. M.: 1992, 'Wamodel Cycle 4', *Deutsches Klima Rechenzentrum: Technical Report* No.4.

Giering, R. and Maier-Reimer, E.: 1996, 'Data Assimilation into the Hamburg LSG OGCM with its adjoint model', *in preparation.*

Hasselmann, K. *et al.*, 1973: 'Measurements of Wind-Wave Growth and Swell Decay during the Joint North Sea Wave Project (JONSWAP)', *Dtsch. Hydrogr. Z.* **12**, 1-95.

Hasselmann, S., Hasselmann, K., Allender, J. H., and Barnett, T. P.: and T.P. Barnett: 1985, 'Computations and Parameterizations of the Nonlinear Energy Transfer in a Gravity-Wave Spectrum, Part II', *J. Phys. Oceanogr.* **15**, 1378-1391.

Janssen, P. A. E. M, 1991: 'Wave-induced stress and the drag of air flow over sea waves', *J. Phys. Oceanogr.* **21**, 1631-1642.

Kahma, K. and Calkoen, Ch. J.: 1992, 'Reconciling Discrepancies in the Observed Growth of Wind-generated Waves' *J. Phys. Oceanogr.* **22**, 1389-1405.

Komen, G. J., Hasselmann, S., and Hasselmann, K.: 1984, 'On the existence of a Fully Developed Wind-Sea-Spectrum', *J. Phys. Oceanogr.* **14**, 1271-1285.

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

Pierson, W. J. and Moskowitz, L.: 1964, 'A proposed spectral form for fully developed wind seas based on the similarity theory of S.A.Kitaigorodskii', *J. Geophys. Res.* **69**, 5181-5190.

The WAMDI Group, 1988: 'The WAM Model — A Third Generation Ocean Wave Prediction Model' *J. Phys. Oceanogr.* **18**, 1775-1810.