

On ensemble prediction of ocean waves

By LEANDRO FARINA*, *Centro de Previsão de Tempo e Estudos Climáticos (CPTEC), Instituto Nacional de Pesquisas Espaciais (INPE), Cachoeira Paulista, SP 12630-000, Brazil*

(Manuscript received 28 March 2001; in final form 28 August 2001)

ABSTRACT

The numerical ensemble prediction is a well accepted method for improving the performance of atmospheric models. In the context of ocean wave modeling little has been researched or documented about this technique. An essential study of the method of ensemble prediction applied to deep water waves has been carried out. A framework is defined for obtaining perturbations of the directional wave spectra and for employing an ensemble of wind fields generated by an atmospheric model. The third-generation global wave model WAM is used with real atmospheric conditions to investigate the effect on wave predictions of perturbed initial conditions and atmospheric forcing. Due to spectral shape stabilisation, perturbing wave initial conditions has limited utility in ensemble prediction. However, the members could be used in wave data assimilation schemes in an interactive way. Using ensembles of the atmospheric condition can generate diverging solutions, justifying the ensemble procedure by itself. In the cases studied, it is observed that the ensemble mean outperformed the other members. The solution behaviour suggests using a lower-order approximation of the model to generate ensemble members with less computational cost.

1. Introduction

Ensemble prediction looks at minimising the effect of the uncertainty, or error, in the initial condition of physical models. The nonlinearity of atmospheric models, for instance, can produce medium-range results that are qualitatively different if small errors are present in the initial condition. Taking the fact that there are always errors in observations and in the analysis, and therefore in the initial condition, the concept of a single and deterministic solution of the model's governing equations becomes fragile. On the other hand, an ensemble of initial conditions, perturbed with errors representing the uncertainty present in the analysis, and generating other ensemble of solutions, or predictions, has the property of providing more information on the long-term behaviour of

the solution. This procedure characterises the ensemble prediction system (EPS). Its main advantages are a greater reliability in the solution, and the generation of several possible events and the probabilities associated with them.

In the context of modeling and prediction of ocean waves, little has been done and documented on ensemble forecasting. We remark upon the work developed at the ECMWF (European Centre for Medium-Range Weather Forecasts) by Hoffschmidt et al. (2000) and Janssen (2000). There, the optimisation of ship routing is studied by employing ensemble prediction of waves, where only the atmospheric data are perturbed. In the center above, daily wave ensemble forecasts became an operational product in June 1998.

Differences between the evolution of atmospheric and ocean wave model solutions will become evident in the following sections. Essentially, in contrast to the case of atmospheric models, solutions of wave models are not very

* Corresponding author.
e-mail: farina@cptec.inpe.br

sensitive to perturbations in the initial condition. Nevertheless, perturbations in the wind fields can produce divergence between solutions of wave models.

2. The third-generation wave model

A third-generation wave model is governed by the action balance equation. This equation describes the evolution of the wave energy spectrum F forced by the source terms and this can be written as

$$P(F) = \frac{DF}{Dt} = S_{in} + S_{nl} + S_{ds}, \quad (1)$$

Here,

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{c}_g \cdot \nabla_{\mathbf{c}_g},$$

where $\mathbf{c}_g$ is the propagation velocity of a wave group in the four-dimensional space referent to the spatial coordinates $\mathbf{x} = (x_1, x_2)$ and the wave vector (k_1, k_2) . The right-hand side of (1) represents terms of source (S_{in}) describing the surface wind input, sink (S_{ds}) describing dissipation due to whitecapping and bottom friction, and the term representing the nonlinear wave-wave interactions (S_{nl}). For further details about these terms and this equation, see Komen et al. (1994). To solve eq. (1), one has to prescribe the spectrum at a certain time and the surface winds for all time integration intervals.

The significant wave height, $H_s = H_{1/3}$, defined as the average of the third highest observed or measured waves, is evaluated using the wave spectrum, in the following form (Ochi, 1998):

$$H_s = 4\sqrt{E}, \quad (2)$$

where E , the wave energy, is given by

$$E = \int_0^{2\pi} \int_0^\infty F(f, \theta) df d\theta.$$

3. Model implementation

The wave model WAM, cycle 4, was implemented at a global domain using the wind stress field from the CPTEC/COLA T62L28 atmospheric global model with a resolution of 1.875

degrees. Numerical integration of global models such as WAM requires some attention when including high latitudes in the mesh. Incorporating these latitudes in the computational domain can cause numerical instabilities due to possible violations of the Courant–Friedrichs–Levy (CFL) criterion. Thus, in order to keep a time step greater or equal to 15 min, and in this way satisfy the CFL criterion, a domain between latitudes 77°S and 77°N was selected. In this region, the ocean ice cover has some importance in the modeling. Simulations with and without ice cover were carried out.

It can be remarked that because of the fact that the wave model is integrated on a global domain, no information at boundaries is lost for the wave spectrum. This loss occurs in limited area applications, where the generation of swells and the energy flux through the boundaries is disregarded.

The implementation described above is used in all experiments in this work.

4. Method

Analogously to ensemble forecasting techniques used in atmospheric models, a method for ensemble prediction of ocean waves can be formulated.

Employing the breeding method (Toth and Kalnay, 1997) for the initial conditions of the wave field and an arbitrary method for perturbing and generating the members of the atmospheric conditions, we have the following.

Suppose that the wave spectra $F^0(f, \theta, x_1, x_2, t_0)$ and $N + 1$ wind fields

$$W^0(x_1, x_2, t), W^1(x_1, x_2, t), \dots, W^N(x_1, x_2, t),$$

for the time t varying in the interval $[t_0, t_f]$ and assuming the value t_1 , with $t_1 < t_f$ are given. The value t_f is the forecasting final time and t_1 is the integration restart time. We will use also the notation F_i^k for $F^k(f, \theta, x_1, x_2, t_i)$ and W_i^j for $W^j(x_1, x_2, t_i)$.

Consider $M/2$ random perturbations $\varepsilon_0^k(f, \theta, x_1, x_2)$, $k = 1, 2, \dots, M/2$, such that $|\varepsilon_0^k| = \mathcal{O}(\delta)$, where δ is the typical mean square error in F_i .

Summing to the spectrum these perturbations and also their opposites, $-\varepsilon_0^k$, we obtain the

following ensemble of M initial conditions F_0^k for the spectrum.

$$F_0 + \varepsilon_0^1, \dots, F_0 + \varepsilon_0^{M/2}, F_0 - \varepsilon_0^1, \dots, F_0 - \varepsilon_0^{M/2},$$

plus the *control initial condition* F_0^0 .

The initial condition matrix $D_0^{(N+1) \times (M+1)}$ is defined by

$$D_0^{(N+1) \times (M+1)}$$

$$= \begin{bmatrix} (F_0^0, W_0^0) & (F_0^0, W_0^1) & \dots & (F_0^0, W_0^N) \\ (F_0^1, W_0^0) & (F_0^1, W_0^1) & \dots & (F_0^1, W_0^N) \\ \vdots & \vdots & & \vdots \\ (F_0^M, W_0^0) & (F_0^M, W_0^1) & \dots & (F_0^M, W_0^N) \end{bmatrix}$$

Applying the action balance equation,

$$P(F) = S(F, W), \quad t_0 \leq t \leq t_f,$$

with $(F, W)|_{t=t_0} = D_0^{(N+1) \times (M+1)}$, we get $(N+1) \times (M+1)$ predictions for the time t_f , equal to 6 days, for instance. In this process of integration, one gets for $t = t_1$, equal to 6 or 12 h, the spectra

$$F_1^0, \tilde{F}_1^1, \dots, \tilde{F}_1^M,$$

that will be used to generate the new perturbations in the following form. Define

$$\varepsilon_1^k = \mathcal{R}(F_1^0 - \tilde{F}_1^k),$$

where $\mathcal{R}$ is a linear function for rescaling the spectral error. To make $\mathcal{R}$ explicit, define α and β as the minimum and maximum of $F_1^0 - \tilde{F}_1^k$, respectively, and A as the perturbation amplitude. Then, $\mathcal{R}$ is defined as

$$\mathcal{R}(y) = -\frac{A}{\alpha - \beta} y + A \frac{\alpha + \beta}{\alpha - \beta}, \quad y \in [\alpha, \beta].$$

In this way, the new initial conditions of the spectrum, for $t = t_1$, will be given by

$$F_1^0 \text{ and } F_1^k = F_1^0 + \varepsilon_1^k, \quad k = 1, \dots, M.$$

Then, the new initial condition matrix is expressed by $D_1^{(N+1) \times (M+1)}$. This process is then repeated indefinitely.

The initial conditions $W^i(x_1, x_2, t)$ are generated by ensemble methods for atmospheric models, such as the breeding method (Toth and Kalnay, 1997; Lorenz, 1965) singular vectors (Molteni et al., 1996) or the perturbations based on empiric orthogonal functions (EOF) (Zhang and Krishnamurti, 1997; Coutinho, 1999).

5. Spectrum perturbation

We will analyse in this section the effect of wave spectrum perturbations in the evolution of the wave model solution.

Take $N = 0$, i.e., only the control wind field will be considered. The wave spectrum F admits perturbation with respect to its arguments f , θ or x . Initially, in order to simplify the analysis, we impose perturbations only in the frequency range $\mathcal{J} = [0.37 \text{ Hz}, 0.41 \text{ Hz}]$. Thus, two perturbations of the spectrum, one positive, F^p , and other negative, F^n , form the members with the following properties.

$$F^p|_{f \in \mathcal{J}} = F^0|_{f \in \mathcal{J}} + 6,$$

$$F^n|_{f \in \mathcal{J}} = F^0|_{f \in \mathcal{J}} - 12 \quad \forall x, \theta. \quad (3)$$

The values 6 and 12 correspond approximately to 10 and 20 times, respectively, the spectrum maximum. Therefore, these are large-amplitude perturbations and represent a jump in the high-frequency spectrum.

Figure 1 shows the evolution of the directional spectrum shape of the control and of the two members above, as a function of the direction θ and of the frequency f , for $x = (300^\circ, 30.7^\circ)$. In this simulation, the initial time is 12.00 UTC of 12 December 2000 and wind stresses from the CPTEC/COLA global atmospheric model were used as forcing of the wave model that had a cold start, i.e., a JONSWAP spectrum (Komen et al., 1994) is assumed as the initial wave field. One can observe that even with the large-amplitude perturbation imposed, the two members and the control converge rapidly to a common spectral shape, in a time not greater than 48 h. This result confirms the observations of Young and Van Vledder (1993) on the phenomenon they denominate spectral *shape stabilisation*, caused mainly by the contribution of the nonlinear interaction term S_{nl} (Komen et al., 1994) of the balance eq. (1). In that work, constant wind fields were used mainly in the study of the one-dimensional spectrum, defined as

$$\mathcal{F}(f) = \int_0^\infty F(f, \theta) d\theta.$$

Consider now a negative perturbation with the absolute value of the positive perturbation ($A = \pm 6$) and at $x = (90^\circ, -30.7^\circ)$. In this case,

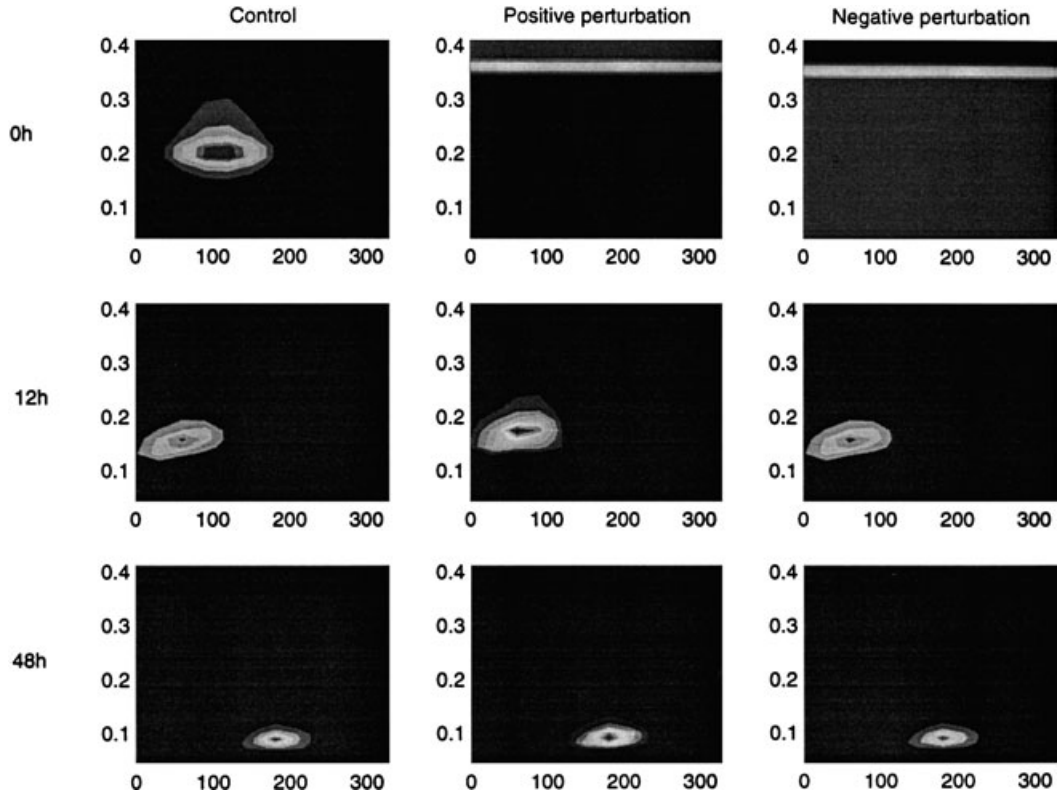


Fig. 1. Evolution of the directional spectra of control and of two perturbations, for $x = (300^\circ, 30.7^\circ)$, starting on 12.00 UTC 12 December 2000. The perturbations are defined by the expressions (3).

shown in Fig. 2, the convergence is somewhat slower, although still well pronounced. It is possible to observe the process of shape stabilisation with the perturbed spectra converging to the peak around $f = 0.2$ and $\theta = 150^\circ$. In the negative perturbation, it is noticed the presence of a secondary peak, for $t \in [6 \text{ h}, 12 \text{ h}]$ during the convergence process.

In this case, as in the previous one, we notice a faster convergence of the negative member. This behaviour will also be observed in the next case we analyse.

With the same atmospheric conditions of the previous examples, consider perturbations for all frequencies and directions, for the points located in the latitude range $\mathcal{J} = (-38^\circ, -48^\circ)$, in the following form

$$F^p|_{x_2 \in \mathcal{J}} = F^0|_{x_2 \in \mathcal{J}} + 3, \quad (4)$$

$$F^n|_{x_2 \in \mathcal{J}} = F^0|_{x_2 \in \mathcal{J}} - 3, \quad \forall f, \theta, x_1. \quad (5)$$

In this case it is interesting to investigate the significant wave height field, in the forecast time of 96 h, equal to 12.00 UTC 16 December 2000, when the occurrence of intense wave activity was observed in parts of Brazil's south, southeast and northeast coasts. In Fig. 3, one can see the effect of the positive perturbation 12, 48 and 96 h after its introduction in the model integration. At 12 h, the perturbation is easily detectable, reaching a stage of quasi-stabilisation at 48 h. Notice that at this time, differences above 2 m in the wave fields at the Atlantic and Pacific Oceans, between the control and the positive member predictions, can be hardly seen. The presence of variations in the 96 h forecast in the Atlantic, near the Brazilian region where the intense wave activity occurred, is therefore remarkable. Thus, like the variations in the 48 h forecasts in the Indian Ocean, to the west of Australia, the differences in the Atlantic at 96 h can be in part attributed to the nonlinear

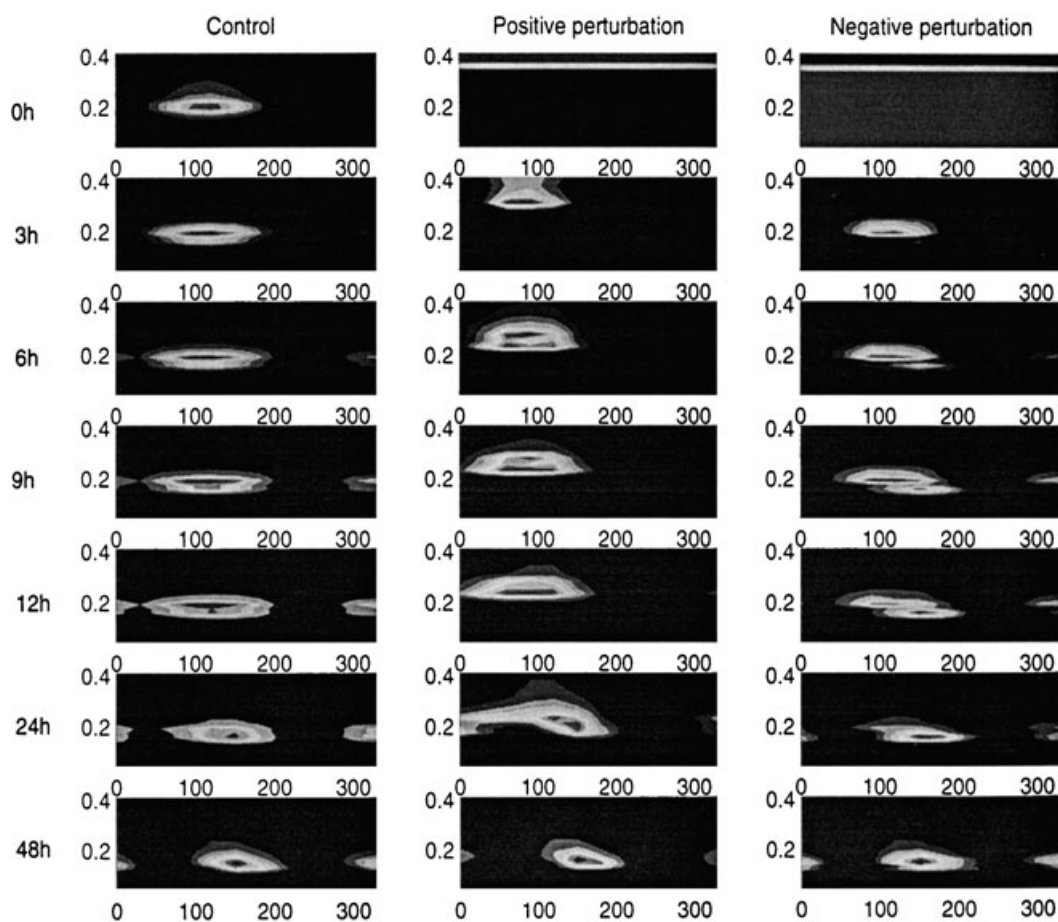


Fig. 2. Evolution of the directional spectra of control and of two perturbations, for $x = (90^\circ, -30.7^\circ)$, starting on 12.00 UTC 12 December 2000. The perturbations have amplitude equal to ± 6 . (5:7)

interactions between waves reminiscent of the positive perturbations introduced near waves of strong energy generated by wind. The negative member convergence is again faster, and already at the 48 h forecasts a stabilisation is noticeable that remains at $t = 96$ h, as shown in Fig. 4.

As observed above, third-generation wave models are essentially insensitive to the spectral initial condition. In some cases variations after apparent stabilisation were verified. However, these are still within the limit of short-range forecasts. The situation is that all first-generation models omit the representation of the term S_{nl} and all second-generation models represent it in an oversimplified way.

Thus, apparently, the wave spectral perturba-

tion is not so relevant in the context of ensemble forecasting methods. The incorporation of ensemble perturbations could be carried out in wave data assimilation procedures and used as valuable data in the calculation of the cost function. Komen et al. (1994, chapter 6) observed that with improved analysis, the relaxation time for the wave field correction can exceed 5 days, in regions dominated by swell.

6. Perturbation in the atmospheric conditions

In contrast to the spectrum initial condition, the atmospheric condition in a wave model has a characteristic that prevents the occurrence of

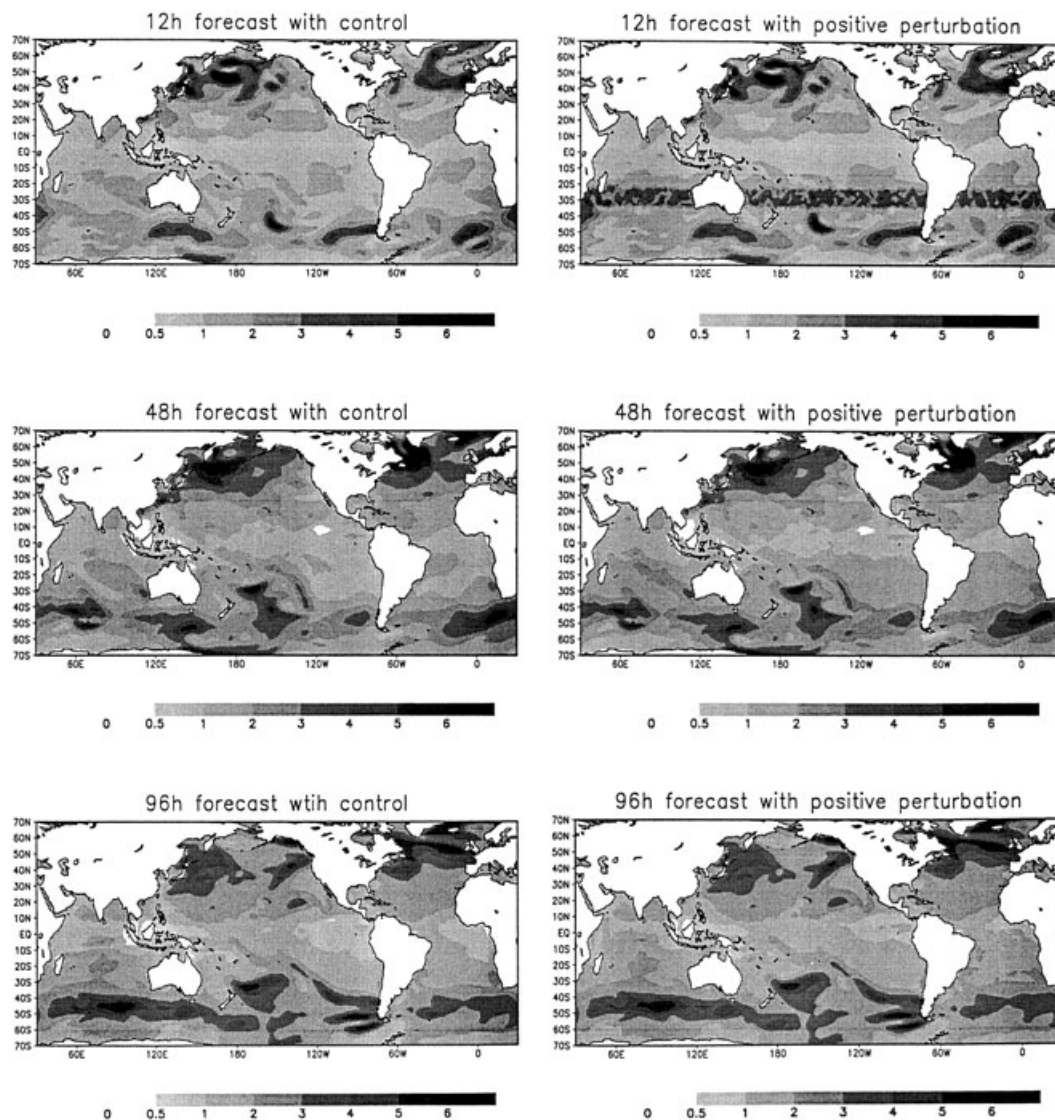


Fig. 3. Significant wave height (m) for control and the negative perturbation, according with (4), starting on 12.00 UTC 12 December 2000. (5:10)

shape stabilisation, observed in the previous section. This differing aspect is the fact that the wind data are strictly not an *initial* condition. The wind fields are prescribed for all values of time, until the end of integration. Then, given M members $W^j(x, t)$, $j = 1, \dots, M$ of an ensemble of atmospheric conditions, one can expect non-converging solutions of the model, once is assured that W^j differ among them. The shape stabilisation is

basically dictated by the term S_{nl} that responds and behaves according to the wind field input. By introducing M atmospheric conditions, one can think of M distinct shapes stabilisations.

In the case that we will analyse, $M = 20$ members form the ensemble of perturbed atmospheric initial conditions and the initial time is 12.00 UTC 10 July 2000. Along with the control condition and the ensemble mean, we will have a total of 22

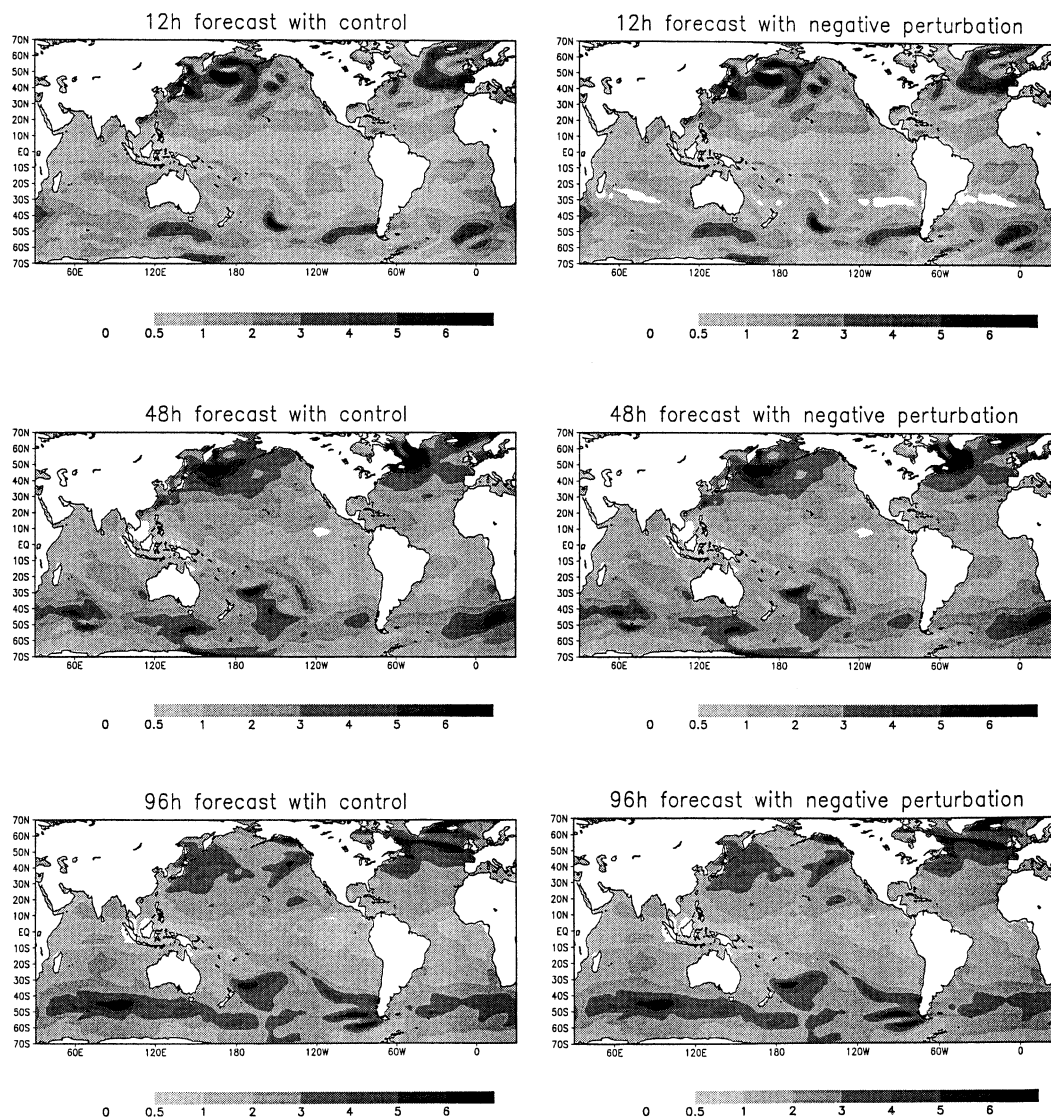


Fig. 4. Significant wave height (m) for control and the negative perturbation, according with (5), starting on 12.00 UTC 12 December 2000. (5.11)

wave forecasts. The atmospheric conditions were generated through the method of EOF-based perturbations, developed originally by Zhang (1997). The perturbations are separated in two groups of 10, denoted by 1P, 2P, ..., 10P and 1N, 2N, ..., 10N, where P and N denote positive and negative, respectively. This nomenclature has its origin in the fact that the perturbations are generated by summing and/or subtracting components from a

certain field. Furthermore, a perturbation iN is related, by symmetry, to its *sister perturbation* iP . In our case some adaptations to the original method took place; for instance, the domain of perturbation was that between latitudes 45°S and 30°N . In this way, for $t = 0$, outside this domain, $W^j(x, t)$ do not differ. For further particulars of this procedure, see Coutinho (1999).

For the time values of 3 and 6 days, the root

mean-square error (RMSE) between the 22 forecasts and a reference solution will be evaluated. The reference solutions are defined as a control forecast of 12 h (*first guess*), verifying at 72 h and 144 h after the initial time. Thus, the reference solutions are integrations started at 60 h and at 132 h, with updated wind fields and spectrum. In this experiment, no atmospheric information prior to $t=0$ was considered; i.e., there was a cold initialization of the model at that date.

Figure 5 shows the significant wave height field of the reference solution and of certain forecasts for 12.00 UTC 13 July 2000 (therefore for 3 days of forecasts). There, we observe a control forecast, the ensemble mean (calculated by averaging the significant wave heights at each point) and two forecasts selected from the ensemble: those with the smallest and greatest RMSEs. These members are the forecasts of the perturbations 9N and 9P, respectively.

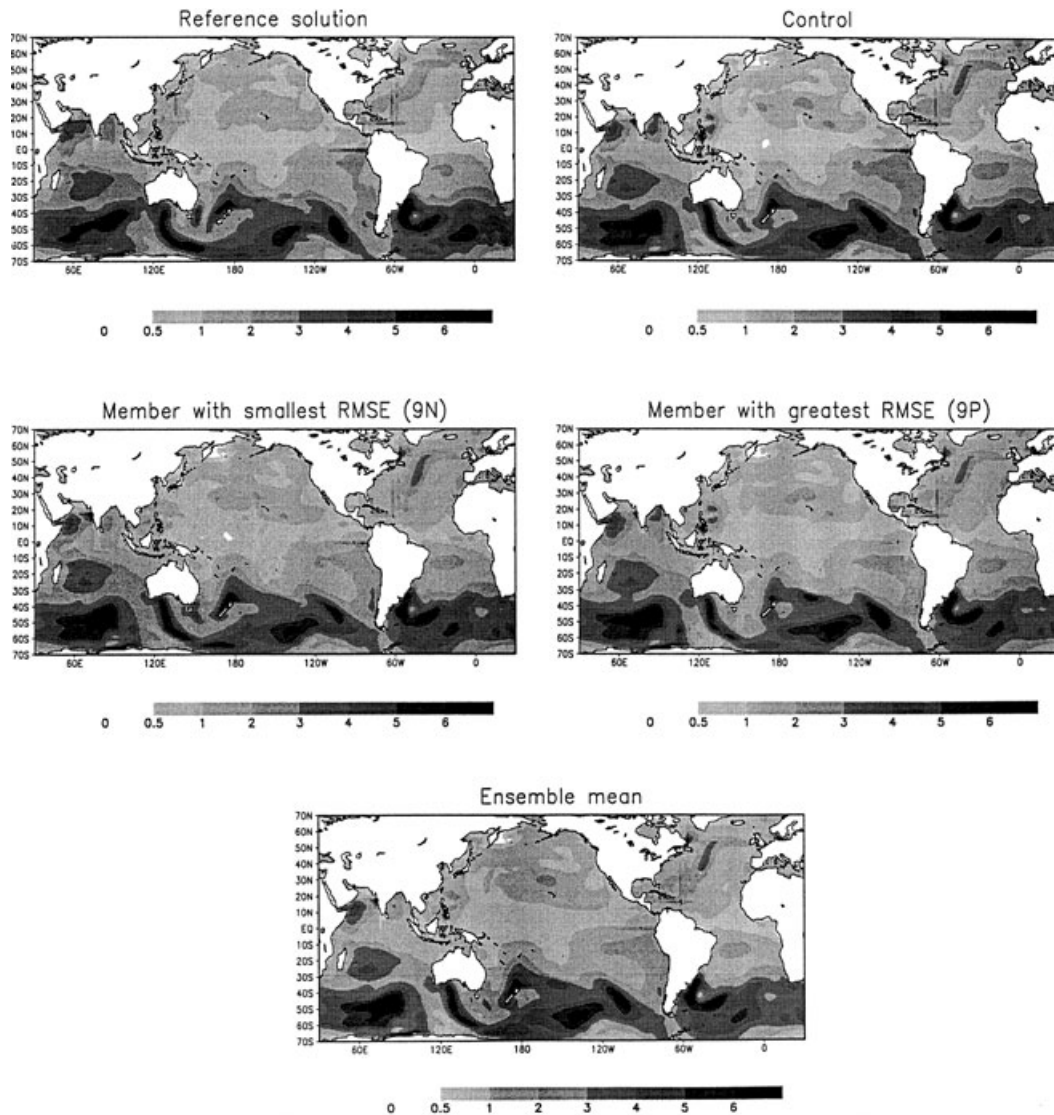


Fig. 5. Significant wave height (m). Forecasts of 72 h, initialized on 12.00 UTC 10 July 2000. (5:12)

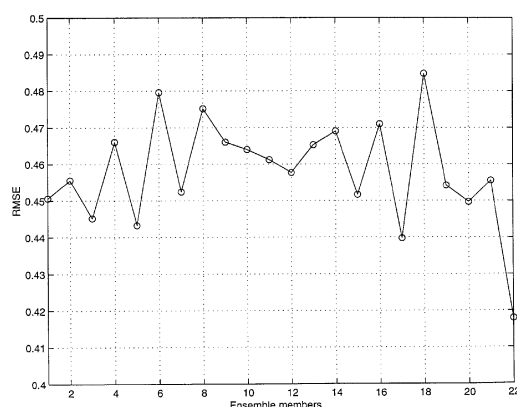


Fig. 6. Root mean square error for members of the ensemble for the forecasts of Fig. 5. The members are denoted by the numbers 1 to 20, in the order 1N, 1P, 2N, 2P, ..., 10N, 10P. The control is represented by abscissa 21 and the ensemble mean, by 22. (5:13)

Although is difficult, from this figure, to determine the forecast closest to the reference, we notice some variations in the patterns, mainly in the Pacific.

A more pragmatic analysis of the forecast ensemble can be inferred through the results in Fig. 6, which show the RMSEs. The members are denoted by the numbers 1 to 20, in the order 1N, 1P, 2N, 2P, ..., 10N, 10P. The control is represented by the abscissa 21, and the ensemble mean by 22. The superior performance of the ensemble mean is remarkable. Another noticeable point is the smaller error presented by the negative members.

Analogously, we show in Figs. 7 and 8 the significant wave height fields and RMSEs for the 6 days forecast. Here, we observe patterns more diverse, and again the ensemble mean is best, although this time one of the members (namely 7N) has comparable quality.

It is suggested that the ensemble mean has the best performance. This seems to confirm the results of Janssen (2000) in his work on the optimisation of ship routes using wave ensemble prediction. In the case where there is a clear occurrence of distinct clusters of solutions, the mean will tend to be erroneous. In this case, a complementary study to determine other members with superior properties is desirable.

7. Discussion

Ensemble wave prediction is in its infancy and the results of this work are expected to serve as motivation for the continuity of research, in order to obtain more sophisticated techniques for the generation and analysis of results. Next, we describe the conclusions of the present paper.

The method of ensemble forecasting was investigated in the context of deep-water wave modeling. A shape stabilisation for directional spectra with real atmospheric conditions was shown, through the cases studied. It was concluded that the initial wave spectral perturbation in the ensemble method has limited utility, if this is looked at separately. Nevertheless, perturbations of this nature are closely related to wave data assimilation techniques by the presence of wave field modifications in the cost function.

Data assimilation and ensemble wave modelling can interact in the following manner. The perturbations from the ensemble method could be used and compared with the modifications from the observations improving the assessment of the assimilated data quality. In fact, several *first guesses* (equal to the number of members plus the ensemble means) could be used in such a quality control. On the other hand, the observational spectrum and significant wave height data can provide information for determining a suitable mechanism for the generation of the ensemble initial conditions.

Perturbations of the surface wind fields can supply important information in short- to medium-range predictions. This additional information can have impacts on probability forecasts, by providing a means of estimating extreme wave events more reliably. In the case studied, the atmospheric conditions were generated through the method of EOF-based perturbations incorporated in the CPTEC/COLA global model. The sensitivity of the WAM model to this type of wind field perturbations was shown quantitatively by means of root mean-square errors. The performance of the ensemble mean, superior to the control and all other members, is remarkable. This situation suggests a very simple way of improving the reliability of wave predictions when local extreme events are not relevant.

The benefits brought by an ensemble prediction system to forecasting are shown. Perhaps

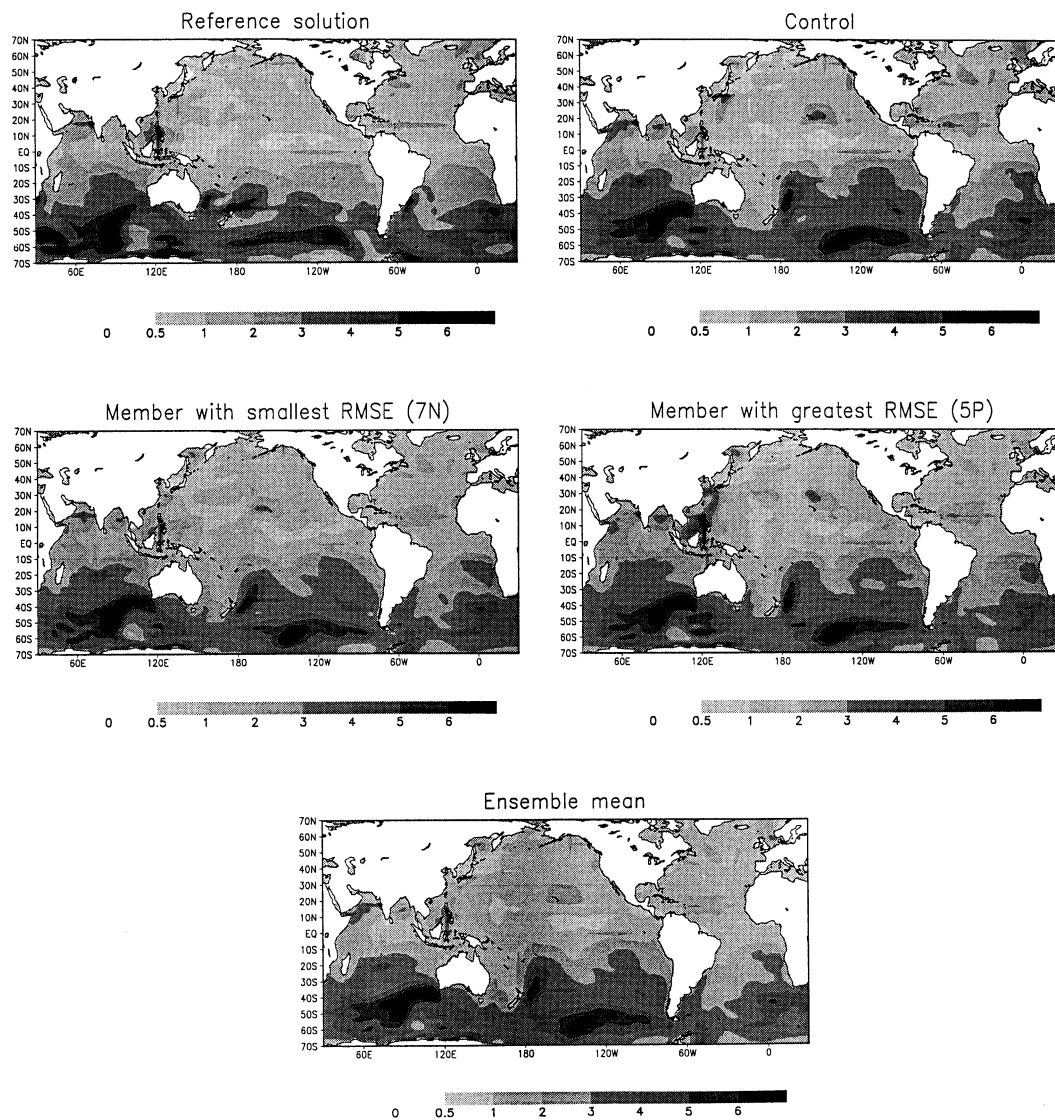


Fig. 7. Significant wave height (m). Forecasts of 144 h, initialized on 12.00 UTC 10 July 2000. (5:14)

this system's unique drawback is the high computational demand associated with it. In this work it was possible to measure how the wave fields differ among the ensemble solutions. The results suggest that it is acceptable to approximate the difference e , between two wave predictions of ensemble members with a weakly nonlinear, or even linear, relation between e and the difference in the corresponding atmospheric

conditions (members). Provided the computational cost for obtaining e is significantly less than the model integration itself, the ensemble solutions can be computed at a fraction of the original EPS computing time. The impact of this approximation on the cost of the EPS can be extensive. We are currently working on this topic and the results will be published in due course.

8. Acknowledgements

We are grateful to Antonio M. Mendonça of CPTEC for providing the wind field ensemble data, and to Carlos A. Nobre and José P. Bonatti for suggesting the research topic of this paper.

We acknowledge financial support by FAPESP (process no. 00/01284-2).

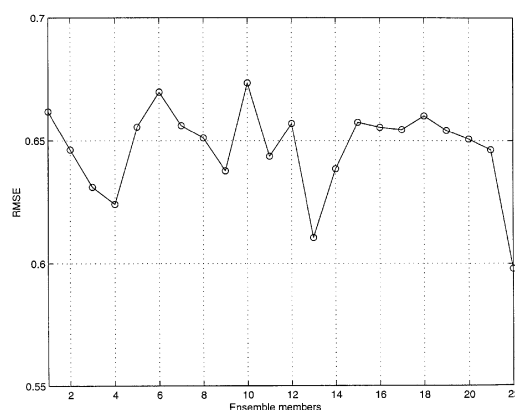


Fig. 8. Root mean square error for members of the ensemble for the forecasts of Fig. 7. The members are denoted by the numbers 1 to 20, in the order 1N, 1P, 2N, 2P, ..., 10N, 10P. The control is represented by abscissa 21 and the ensemble mean, by 22. (5:15)

REFERENCES

- Coutinho, M. M. 1999. Previsão por conjuntos utilizando perturbações baseadas em componentes principais. Master's thesis, INPE, São José dos Campos, Brazil.
- Hoffschmidt, M., Bidlot, J.-R., Hansen, B. and Janssen, P. A. E. M. 2000. Potential benefit of ensemble forecasts for ship routing. *ECMWF Technical Memorandum no. 287*, Reading, UK.
- Janssen, P. A. E. M. 2000. Potential benefits of ensemble prediction of waves. *ECMWF Newsletter* **86**, 3–6.
- Komen, G. J., Cavaleri, L., Donelan, M., Hasselmann, K., Hasselmann, S. and Janssen, P. A. E. M. 1994. *Dynamics and modelling of ocean waves*. Cambridge University Press, Cambridge, 532 pp.
- Lorenz, E. N. A. 1965. A study of the predictability of a 28-variable atmospheric model. *Tellus* **17**, 321–333.
- Molteni, F., Buizza, R., Palmer, T. N. and Petroliagis, T. 1996. The ECMWF Ensemble prediction system: methodology and validation, *Quart. J. R. Meteorol. Soc.* **122**, 73–119.
- Ochi, M. K. 1998. *Ocean waves, the stochastic approach*. Cambridge University Press, Cambridge, 319 pp.
- Toth, Z. and Kalnay, E. 1997. Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.* **125**, 3297–3319.
- Young, I. R. and Van Vledder, G. Ph. 1993. A review of the central role of nonlinear interactions in wind-wave evolution. *Philos. Trans. R. Soc. London, Ser. A* **342**, 505–524.
- Zhang, Z. 1997. Hurricane ensemble prediction using EOF-based perturbations. PhD Thesis. Florida State University, Florida.
- Zhang, Z. and Krishnamurti, T. N. A. 1999. A perturbation method of hurricane ensemble predictions. *Mon. Wea. Rev.* **127**, 447–469.