The Global Atmosphere and Ocean System, Vol. 8, No. 1, pp. 67–83

Taylor & Francis
Taylor & Francis Group

LIMITERS IN THIRD-GENERATION WIND WAVE MODELS[1]

HENDRIK L. TOLMAN

SAIC-GSO at Ocean Modeling Branch, Environmental Modeling Center
NOAA/NCEP, 5200 Auth Road Room 209, Camp Springs, MD 20746

(Accepted November 2001)

Third-generation ocean wave models include a so-called limiter in the integration of the source terms to guarantee numerical stability at economical numerical time steps. The original limiter has previously been associated with the sensitivity of model results to the numerical time step. More recent limiters appear to remove this sensitivity by eliminating the numerical convergence from the resulting integration scheme. This is contrary to rudimentary numerical principles as well as the underlying philosophy of third-generation wave models. The present study investigates the effects of limiters and large model time steps using time-limited wave growth test. It is shown that the conventional limiter results in stable model results even if the numerical time step violates the time scales of wave growth. Contrary to common belief, its impact is not necessarily limited to the equilibrium range of the spectrum, and the limiter systematically *enhances* growth rates in the intermediate stages of wave growth. Particularly initial growth errors increase significantly with increasing maximum discrete spectral frequency f_{max}. Relaxation of the limiter is shown to reduce initial growth errors, but does so at the expense of notable errors in the spectral shape. In the present paper the limiter was relaxed by introducing a new asymmetric limiter that retains full convergence. Initial results obtained with this limiter are similar to those of the advocated nonconvergent limiters. Although this limiter still needs rigorous testing and further development, its initial results suggest that there is no justification for using nonconvergent limiters.

Keywords: third-generation wave models; limiter; nonconvergence; parameterization

1. INTRODUCTION

Third-generation ocean wave models solve some form of the spectral energy or action balance equation, for instance

$$\frac{DF}{Dt} = S, \tag{1}$$

where F is the wind wave spectrum, and S represents source terms for spectral wave energy due the influence of wind, wave breaking ('whitecapping') nonlinear interactions and additional (mostly shallow-water) processes. In third-generation wave models, all sources on the right side of this equation are explicitly parameterized and accounted

[1]OMB contribution Nr. 210.

ISSN 0095-8972 print: ISSN 0000-000X online/01/010000-00 © 2002 Taylor & Francis Ltd
DOI: 10.1080/10236730290003392

for in the integration of the equation. The first operational third-generation wave model was the WAM model (WAMDIG 1998, Komen *et al.*, 1994), which solves Eq. (1) for the energy density spectrum $F(f,\theta)$ as a function of the spectral frequency f and the spectral direction θ. To achieve stable results at reasonable time steps, this model integrates the source terms in time using a semi-implicit scheme and a so-called 'limiter' that restricts the maximum change for each spectral bin per time step. Only recently this limiter and the argumentation behind it have been discussed in detail (Hersbach and Janssen 1999, henceforth as HJ99). In WAM cycles 1 through 3, the limiter ($\mathcal{L}_0$) is given as

$$\mathcal{L}_0 = 0.62 \times 10^{-6} g^2 f^{-5}, \tag{2}$$

$$-\mathcal{L}_0 < \Delta F_{\text{lim}} < \mathcal{L}_0, \tag{3}$$

where ΔF_{lim} is the discrete change of spectral energy density per time step after application of the limiter. The maximum change allowed ($\mathcal{L}_0$) corresponds to about 10% of the Phillips spectrum with the energy level according to Pierson and Moskowitz (1964) for fully developed seas[1]. Furthermore, the integration is only performed up to a cut-off frequency f_c, above which a parametric tail corresponding to the Phillips spectrum is applied. This cut-off also helps to stabilize the integration, although arguments for applying such a cut-off appear rooted in physics rather than numerics. The cut-off frequency is usually set to 2.5 times the peak or mean frequency (see Komen *et al.*, 1994).

With the above integration method and limiter, WAM cycles 1 through 3 give stable results, but the results proved to be sensitive to the discrete time step Δt of the model (e.g., Tolman 1992, Fig. 1; HJ99 Section 1).

Tolman (1992, henceforth denoted as T92) removed the time step dependence from the integration by using the limiter to calculate the maximum allowed time step, and by dynamically adjusting the time step. When the time step is calculated separately for each spatial grid point, this method proves to be efficient for large-scale models[2], where conditions of active wave generation cover only a small part of the domain. For the active generation areas the time step then becomes significantly smaller, but for the remaining spatial grid points the time step can safely be chosen much larger than the customary 1200 s. In practice, this implies that the average source term integration time step for large-scale models significantly increases, making such models more economical to operate. This method is also advocated by Hargreaves and Annan (2001). Reservations voiced by Hersbach and Janssen (2001) are based on an erroneous representation of T92 (see Appendix).

The above solution is not applicable to the WAM model, because the design of this model requires that a single overall time step be used throughout the model. In WAM cycle 4, along with new physics parameterizations, a new formulation for the limiter was introduced to reduce the time-step dependence of the solution (see HJ99). This limiter is given as

$$\mathcal{L}_4 = \mathcal{L}_0 \frac{\Delta t}{\tau} \tag{4}$$

[1]That is, in the mean wave direction, assuming a $\cos^2$ type directional distribution.
[2]See Appendix for relevant details of implementation.

$$-\mathcal{L}_4 < \Delta F_{\text{lim}} < \mathcal{L}_4, \tag{5}$$

where $\tau = 1200$ s is a normalization time step. Like this limiter, the discrete change of spectral density per time step ΔF approximately scales with the time step Δt. Thus the relative impact of this limiter is independent of the time step and will not disappear for $\Delta t \to 0$. This implies that the limiter becomes an integral part of the solution, and that a numerical scheme that includes this limiter no longer converges to the solution for the physics parameterizations of the model. In fact, a model based on this limiter may be expected to closely reproduce the previous solution for a time step $\Delta t - 1200$ s, independent of the actual time step Δt. Convergence of the integration of source terms and hence an independence of the numerical solution from the limiter requires that the impact of the limiter reduce with reduced time step of the model, or

$$\frac{\partial \mathcal{L}}{\partial(\Delta t)} < 1. \tag{6}$$

A weaker than linear dependence of $\mathcal{L}$ on Δt thus retains full convergence, albeit with slower speed than for a limiter that is independent of Δt.

Extensive experience with WAM cycle 4 has shown that this model behaves poorly at short fetches (e.g., Hersbach 1996). This behavior can be attributed to improper scaling behavior of limiter (4). To correct this, several alternative limiters have been proposed for WAM cycle 4. These limiters generally replace the Phillips type spectral shape of limiters (2) and (4) with a Toba type spectral shape, and replace $\Delta t/\tau$ with a properly scaling linear dependence on Δt (e.g., HJ99, Luo and Sclavo, 1997, Hargreaves and Annan 1998, Monbaliu *et al.*, 2000). A good example of such a limiter is the HJ99 limiter, which is given as

$$\mathcal{L}_4 = 3.0 \times 10^{-7} g u_* f^{-4} f_c \Delta t, \tag{7}$$

where u_* is the friction velocity, limited by its Pierson Moskowitz value, and f_c is the dynamically calculated cut-off frequency as described above. The test cases presented in the above papers suggest that such limiters indeed result in a small dependence on the numerical time step Δt, and show proper scaling behavior for short fetches. HJ99 furthermore justify the use of a nonconvergent limiter by stressing the importance of proper scaling behavior of the model while adhering to large time steps, and by mentioning our general lack of understanding of the physics in the spectral range where the limiter is usually activated. It therefore seems that at least for engineering purposes, where the bottom line is that the model performs properly, reliably and economically, the most recent limiters like (7) are a major step forward for WAM, and potentially for other third-generation wave models.

From a scientific point of view, however, the nonconvergence of all new limiters violates rudimentary principles of numerical modeling. It is also contrary to the philosophy of third-generation wave models. This philosophy implies that the evolution of the spectrum should be determined by the parameterization of the source terms alone. This philosophy was adopted in order to get the best possible wave model, that could also be used as a tool to investigate alternative formulations for source terms. If the limiter becomes a systematic part of the solution, it becomes difficult if

not impossible to use the model for the latter purpose. The argument of HJ99 that their limiter is used only sporadically might be considered to counter this reservation. This argument, however, is at odds with the additional claim of HJ99 that the time step dependence of the model can only be removed adequately with a nonconvergent limiter[3].

From the above it appears clear that nonconvergent limiters like (7) may be acceptable for engineering applications, but reduce the usefulness of the model as a scientific tool. Moreover, if these limiters are used in engineering applications, it is of crucial importance to know what the limitations of the applicability of these limiters are. Such limitations have not been addressed by previous authors.

Considering this, the present paper seeks to investigate the impact of limiters in third generation wave models. To this end a test model and test case are selected in Section 2. Effects of this limiter in this test case are illustrated and discussed in Section 3. This test will be used to support or disprove common wisdom regarding limiters, and furthermore gives an indication of the limits of applicability of limiters in combination with large time steps. To further access effects of relaxed limiters, and to address the potential of alternative limiters, a new asymmetric but convergent limiter is presented and tested in Section 4. The results of this study are discussed in Section 5. This section also presents a more detailed comparison of the present results with HJ99. Finally, conclusions are summarized in Section 6.

2. TESTING LIMITERS

To test the effects of a limiter, a wave model and limiter need to be selected. In the present study, WAM cycle 3 physics as described in WAMDIG (1988) and the corresponding limiter (2) are used.

This limiter is chosen as it is only published and widely used convergent limiter. Its effects can easily and elegantly be addressed by successively reducing the time step of the model. The main argument against this limiter is that a Toba type frequency dependency ($\mathcal{L} \propto u_* f^{-4}$) might be more appropriate than the Phillips type dependency ($\mathcal{L} \propto f^{-5}$). This convergent limiter, however, will by definition not become an integral part of the model results. The only requirement of the limiter then is that it is a reasonable approximation of expected (changes of the) spectra. Over three decades of experience with Phillips type spectra more than qualify these spectra as such. Furthermore, the direct dependency of the Toba type spectrum on u_* implies that $\mathcal{L} \to 0$ for $u_* \to 0$, unless u_* is 'filtered' (HJ99). The inherently subjective nature of this filtering, in the opinion of the author, negates physical arguments for using a Toba spectral type in the limiter.

Two additional remarks need to be made regarding limiter (2). First, this limiter displays proper scaling behavior. Following HJ99 Section 2, it can be written in a non-dimensional form as

$$\mathcal{L}^* = 0.62 \times 10^{-7} f^{*-5}, \tag{8}$$

[3]Note that in a precursor to HJ99, Hersbach (1996) proposed a convergent limiter and mentions the importance of convergence. This limiter was replaced by (7) in HJ99.

where $\mathcal{L}^* = \mathcal{L}g^3u_*^{-5}$ is the nondimensional limiter, and $f^* = fu_*g^{-1}$ is the nondimensional frequency. Secondly, because the limiter displays proper scaling behavior, it is restrictive when applied at different scales.

In this context it should be noted that the non-scaling behavior of the WAM cycle 4 limiter (4) as identified by HJ99 is caused by the addition of the factor $\Delta t/\tau$. This factor is also a main reason for this limiter to be too restrictive for short fetches. In the 1 km grid spacing case in HJ99 Fig. 2, $\Delta t/\tau = 0.1$. This limits the allowed change per time step to only 1% of the Phillips equilibrium level, effectively prohibiting wave growth. HJ99, however, erroneously suggest that limiter $\mathcal{L}_0$ instead of the factor $\Delta t/\tau$ is the source of the suppressed growth at short fetches.

The WAM cycle 3 physics have been selected here for their robustness, ease of implementation, and reproducibility. An argument against using these physics would be that they are not 'state of the art'. Presently, at least three third generation wave models are widely distributed, WAM, SWAN (Booij *et al.*, 1996, 1999; Ris *et al.*, 1999) and WAVEWATCH (Tolman and Chalikov 1996; Tolman 1999). Even with different physics parameterizations, including the WAM cycle 3 physics, these models show fairly similar growth behavior. Therefore, general findings regarding limiters as obtained for one parameterization should at least be qualitatively applicable to other parameterizations. This can be tested qualitatively by comparing the present results with those of HJ99.

Effects of the limiter are easiest isolated in a simple time-limited homogeneous test case for deep water as used in T92 and HJ99. In this test case, the following equation is solved

$$\frac{\partial F(f,\theta)}{\partial t} = S(f,\theta). \tag{9}$$

Unlike in HJ99, fetch-limited test cases are not necessary, because limiter (2) shows proper scaling behavior. The spectrum has been discretized with 24 directions ($\Delta\theta = 15°$), and 40 frequencies from 0.042 to 1.72 Hz with a logarithmic distribution where $f_{i+1} = 1.1f_i$. In the test cases presented in the following section, the initial conditions consist of a JONSWAP spectrum with a peal frequency $f_p = 0.5$ Hz and $\gamma = 3.3$. This discrete frequency range was chosen to assure that the prognostic part of the spectrum (i.e., the spectrum up to f_c) always falls within this range. Note that the initial peak frequency f_p was deliberately chosen significantly larger than in previously published tests in order to illustrate specific limiter behavior (as will be discussed below).

3. EFFECTS OF THE LIMITER $\mathcal{L}_0$

Figure 1 shows growth curves for the above test case as obtained with various time steps with a wind speed $U_{10} = 20\,\mathrm{m\,s^{-1}}$. The largest time step $\Delta t = 1200$ s (dotted line) is often used in large-scale third generation wave models. Successive reduction of the time step shows a large impact of the time step on the initial significant wave heights $H_s (= 4\sqrt{E}$, E is the total energy in the spectrum F, panel a), with convergent behavior near full-grown conditions ($t > 15$ h). With successive reduction of the time step the impact of the time step on the solution becomes negligible, and the convergent solution

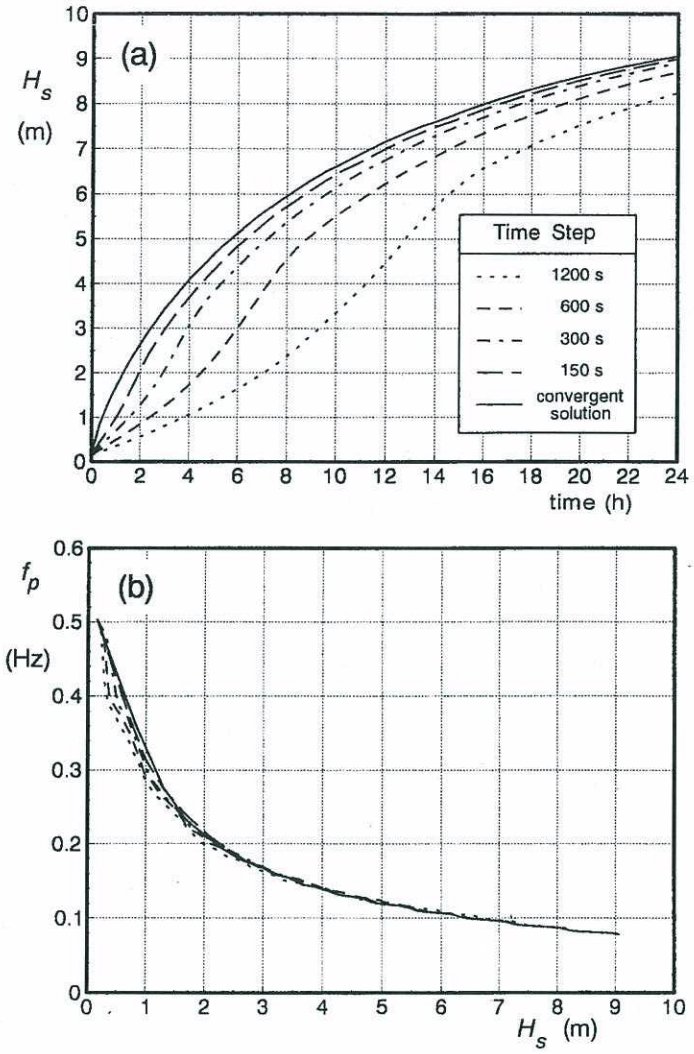

FIGURE 1 Time limited wave growth for a wind speed $U_{10} = 20$ m/s for several time steps for WAM cycles 1–3 physics and numerics (WAMDIG 1988). twenty-four discrete directions and 40 discrete frequencies ranging from 0.042 to 1.72 Hz. Initial conditions consist of a JONSWAP spectrum with peak frequency 0.5 Hz and peak enhancement factor $\gamma = 3.3$. (a) Wave height H_s as a function of time. (b) Peak frequency f_p as a function of the wave height H_s.

is reached. This solution is in principle free of time discretization errors. The convergent solution as presented in Fig. 1 (solid line) has been obtained with $\Delta t = 5$ s. For practical purposes, this solution is reached for $\Delta t = 60$ s (not presented in figure).

The corresponding evolution of the peak frequency f_p [estimated using a parabolic fit to the frequency spectrum $F(f)$] is presented in Fig. 1b. The frequency is plotted as a function of the wave length H_s to provide a simple estimate of errors in the spectral shape. In the initial growth stages, errors in $f_p(H_s)$ can be up to 20%. Although such errors are notable, they are small compared to the corresponding errors in the wave height H_s (Fig. 1a).

Before continuing with analyzing specific effects of the limiter, it should be noticed that Fig. 1 shows a much larger impact of the time step than Fig. 1 of T92, which was obtained with a virtually identical model and test. The only differences between the test cases are the discrete spectral range and initialization of the models. In the present test, initial growth is forced to occur at much higher frequencies, where the much shorter time scales of growth are expected to be impacted much more for identical limiters and time steps.

Effects of a limiter can be assessed by considering the ratio between the discrete spectral change after the limiter is applied (ΔF_{lim}) and the change before the limiter is applied (ΔF).

$$\epsilon(f,\theta) = \frac{\Delta F_{\text{lim}}(f,\theta)}{\Delta F(f,\theta)}. \tag{10}$$

Of particular interest in such an analysis are patterns that can be identified with instability. Tentatively, these are oscillations at the scale of the grid. Visual inspection of the spectral distribution of this ratio suggest that such oscillations mostly occur in f space, with consistent behavior along directions (figures not presented here). Furthermore, the predominant importance of nonlinear interactions, which are thought to be the main source for the instabilities, lies in its redistribution of energy over frequencies. Therefore, it appears useful to consider the ratio of limited to unlimited changes as a function of the frequency only

$$\epsilon(f) = \frac{\int_0^{2\pi} \Delta F_{\text{lim}}(f,\theta)\, d\theta}{\int_0^{2\pi} \Delta F(f,\theta)\, d\theta}. \tag{11}$$

Finally, an impression of the overall impact of the limiter can be obtained by addressing this ratio for the entire spectrum,

$$\epsilon_t = \frac{\int_0^{fc}\int_0^{2\pi} \Delta F_{\text{lim}}(f,\theta)\, d\theta\, df}{\int_0^{fc}\int_0^{2\pi} \Delta F(f,\theta)\, d\theta\, df}. \tag{12}$$

Because this ratio has no meaning in the diagnostic part of the spectrum, the integration is performed up to f_c only.

Figure 2 shows ϵ_t, the ratio of limited to unlimited change of total prognostic energy, for computations with $\Delta t = 1200$ s using the limiter of Eqs. (2) and (3). Presented are the overall ratio (solid line), and ratios based on positive or negative $\Delta F(f,\theta)$ only (dashed and dotted lines, respectively). The suppressed growth rates in the initial hours integration as evident in Fig. 1 (compare solid and dotted lines) can be explained by the fact the limiter allows only a fraction of the positive and total energy changes during the first hours of integration ($\epsilon_t \ll 1$, solid and dashed lines in Fig. 2), whereas a larger part of negative changes is allowed by the limiter (dotted line). Between 12 h and 16 h, the limiter has a bigger impact on negative changes ΔF than on positive ones (dotted and dashed lines). This implies that (contrary to common belief) limiters may artificially *increase* growth rates ($\epsilon_t > 1$).

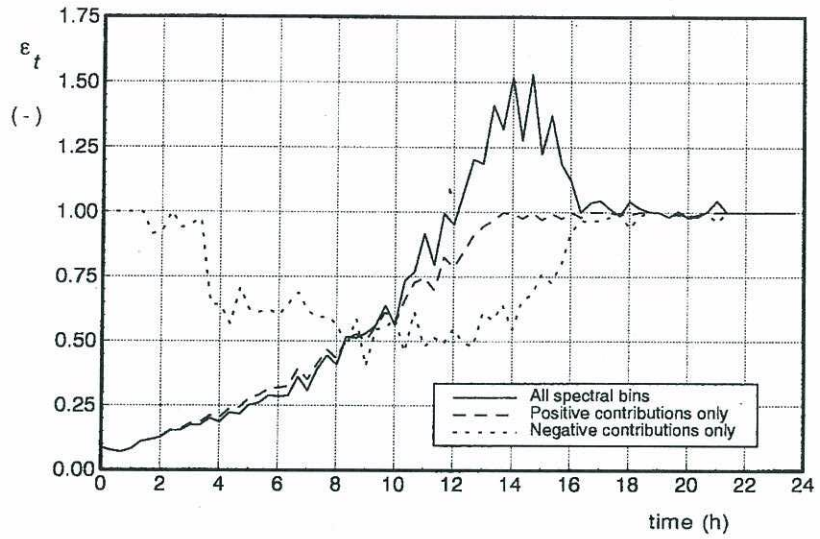

FIGURE 2 Ratio of total prognostic discrete energy change after and before application of the limiter [ε_t, Eq. (12)]. $\Delta t = 1200$ s, conventional limiter $\mathcal{L}$ according to Eqs. (2) and (3).

Figure 3 shows the distribution of effects of the limiter $\mathcal{L}_0$ over frequencies [$\epsilon(f)$] for several times, again making a distinction between all, positive and negative values of ΔF (○, dashed and dotted lines, respectively).

After 1 h of integration (upper left panel of Fig. 3), the limiter allows less than 10% of the positive and total expected changes in spectral energy density [$\epsilon(f) \leq 0.1$, ○ and dashed lines] for the spectral peak (●) and higher frequencies, whereas it leaves negative changes ΔF unchanged [$\epsilon(f) = 1$, dotted line]. This implies that the model attempts to significantly increase the energy levels throughout the spectrum. For high frequencies this is expected as WAM cycles 1 through 3 are known to overestimate the energy level at high frequencies compared to the fairly realistic initial conditions used here. This behavior is therefore at least partially caused by poorly chosen initial conditions. It is nevertheless representative for arbitrary source term formulations, as a wave model will have to regenerate the equilibrium range of a spectrum if preexisting swell comes under the direct influence of rapidly increasing winds.

After 3–6 h of integration, the equilibrium level is reached for frequencies above f_p (●), and the limiter influences positive and negative contributions at similar rates. The limiter is activated for alternatively negative and positive contributions, suggesting that it actively suppresses instabilities. At the same time, the limiter still systematically suppresses wave energy growth at the spectral peak frequency f_p by as much as 80% or more [●, $\epsilon(f) < 0.2$]. A similar suppression of about 35% is still present after 12 h of model integration (lower right panel).

Figures 2 and 3 identify some of the pathology of the limiter; less than 25% of the total change of energy is allowed in the first 5 h of integration and less than 50% in the first 8 h (solid line in Fig. 2, $\epsilon_t < 0.25$ and 0.50, respectively). Whereas the limiter suppresses growth rates at the spectral peak (see above description of Fig. 3). Thus, contrary to common belief, effects of the limiter are not necessarily confined to the equilibrium range of the spectrum. Furthermore, values of ϵ that are systematically

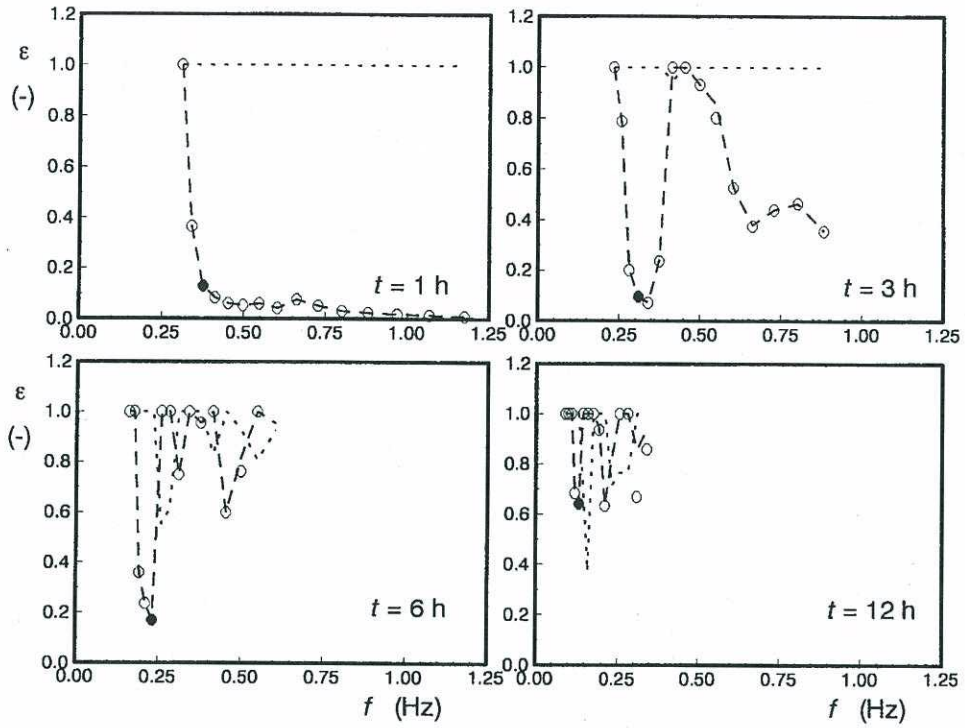

FIGURE 3 Ratio of discrete spectral change after and before application of the limiter corresponding to Fig. 2 for several times [$\varepsilon_i(f)$, Eq. (11)]. Prognostic part of the spectrum only ($f<f_c$). ○: total spectral change. ●: total change at discrete peak frequency. Dashed line: positive contribution of ΔF only. Dotted line: negative contributions only. Note only occasionally ○ will be out of the range of the figures (and even be negative), when the total change of energy before applying the limiter is close to 0.

as low as 0.1 indicate that the physical parameterizations in the model should result in a local equilibrium in a single time step or less. This implies that the model time step is simply incompatible with the time scales of development of both the total energy and the spectral energy at f_p.

It could be argued that the limiter $\mathcal{L}_0$ is a victim of its own success; it allows stable integration for time steps which are grossly incompatible with the physics at the frequencies that carry the bulk of the wave energy (i.e., close to f_p). Due to this incompatibility, it is unrealistic to expect an integration with any limiter to be both stable and accurate in the present test case. This does not imply, however, that the improvement of the limiter $\mathcal{L}_0$ cannot be achieved. Therefore, alternative limiters are considered in the following section.

4. AN ALTERNATIVE LIMITER

The results presented in the previous section show that the limiter $\mathcal{L}_0$ has two distinct effects on the integration of the source terms; it suppresses instabilities, but in doing so it also suppresses wave growth in general. Furthermore, instabilities occur mostly above the spectral peak frequency, and have an expected signature related to the spectral grid

resolution. Considering that the instabilities can be regarded as a superposition on the physically realistic spectral changes, and that in principle only the instabilities need to be limited, it appears logical to replace the symmetric limiter (3) by an asymmetric limiter

$$(\zeta - 1)\mathcal{L}_0 < \Delta F_{\text{lim}} < (\zeta + 1)\mathcal{L}_0, \tag{13}$$

where ζ represents the (nondimensional) asymmetry. A first guess for ζ (denoted as ζ_1) can be obtained from a local (in the spectrum) average of the discrete spectral change normalized with $\mathcal{L}_0$

$$\zeta_1 = \overline{\Delta F(f,\theta)/\mathcal{L}_0}, \tag{14}$$

where the area of averaging has to be large enough to average out local oscillations in $\Delta F(f,\theta)$ related to instabilities, yet small enough to be close to the physically realistic part of $\Delta F(f,\theta)$. Note that such an approach in principle corresponds to a relaxation of the limiter for the physically realistic part of ΔF, without relaxing the limiter for the unstable contributions to ΔF. Somewhat arbitrarily, ζ_1 will be calculated here using five grid points in f and θ spaces (25 in total), with relative weights of one for the central three points in both spaces, and 0.5 otherwise. As is obvious from Fig. 2, ζ_1 may be big as 10 in the initial stages of the integration ($\zeta_1 \propto \epsilon^{-1}$). Such large values of ζ would allow the model to reach the Phillips energy level in a single step, and would clearly lead to accuracy and probably to stability problems. Initial experiments with the asymmetric limiter indeed showed that $|\zeta| < 1$ is required near f_c to suppress instabilities, but that ζ may be larger near the spectral peak frequency. Therefore, ζ in Eq. (13) is estimated here from ζ_1 and a maximum allowed value ζ_{max}

$$\zeta = \pm \min(|\zeta_1|, \zeta_{\text{lim}}), \tag{15}$$

where the sign of ζ equals the sign of ζ_1, and where the limiting value ζ_{lim} is estimated as

$$\zeta_{\text{lim}} = \begin{cases} \zeta_{\text{max}} & \text{for} \quad f \le 0.4 f_c \\ \ldots & \text{for} \quad 0.4 f_c < f \le 0.7 f_c \\ \min(1, \zeta_{\text{max}}) & \text{for} \quad f > 0.7 f_c \end{cases}, \tag{16}$$

where ... identifies a linear combination of the adjacent constant values. Figure 4 shows the results of calculations with $\Delta t = 1200$ s and with several values of the maximum nondimensional asymmetry ζ_{max}.

Figure 4 shows a positive impact of a moderate asymmetry of the limiter ($\zeta_{\text{max}} \le 3$) on the model results. With increasing asymmetry, the wave heights $H_s(t)$ converge to the convergent solution (compare dashed and solid lines in Fig. 4a). For larger allowed asymmetries ($\zeta_{\text{max}} = 4$) the model starts to display an unrealistic 'overshoot' behavior (upper dashed line in Fig. 4a). This behavior is not related to numerical instability, as spectra remain well behaved (figures not presented here). Instead, it appears to be related to numerical accuracy. Limited accuracy might be expected, because this limiter allows for 'equilibrium' in the spectrum to be reached in as little as two discrete time steps.

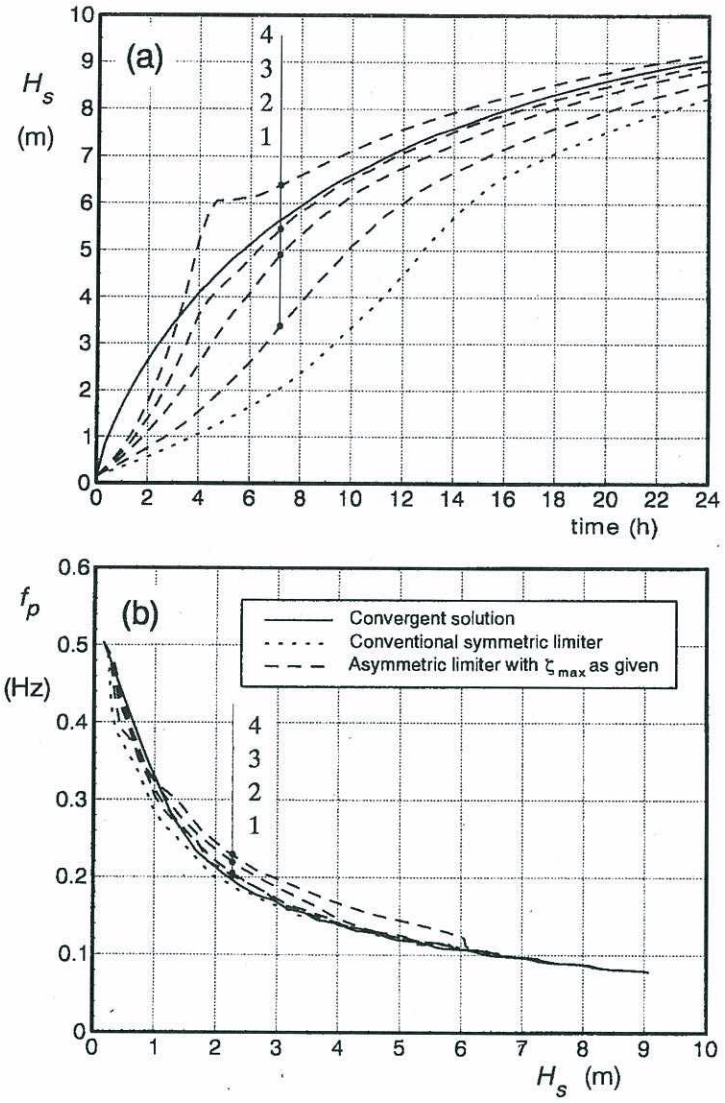

FIGURE 4 Like Fig. 1 for asymmetric limiter and $\Delta t = 1200$ s.

The improved representation of the wave heights for increased asymmetries ζ_{max} comes to some degree at the expense of an increased error in the spectral shape. This is illustrated in Fig. 4b, which shows a systematic over-estimation of $f_p(H_s)$ which increases with increasing ζ_{max}. Qualitatively, this might be expected as wave growth occurs due to an interplay of two processes; the wind increases spectral energies for a given frequency and directions, whereas the nonlinear interactions shift energy to lower frequencies. Integration with large time steps violates physical time scale of the first process, as has been discussed in Section 3. In the convergent solution, f_p shifts by 20–30% per 20 min in the initial stages of wave growth. This implies that a new spectral peak is generated at frequencies where 20 min earlier virtually no wave energy was present. Because nonlinear interaction will obviously not be able to achieve this in a

single time step, the time scales of shifting the peak frequency are also grossly violated. Relaxing the limiter by allowing a larger asymmetry or by other means (as in HJ99), will have a direct impact on the local wave growth due to wind, and hence increase H_s significantly. It will increase the nonlinear interactions only indirectly when local spectral densities become larger. It will therefore only indirectly influence the shift of f_p. It is therefore not surprising that the shifting of f_p to lower frequencies starts to lag with an increasingly relaxed limiter.

5. DISCUSSION

The present study analyzes the effects of the so-called limiter on the source term integration in third-generation wind wave models. This limiter is intended to suppress instabilities in the spectrum that occur if economically large time steps are used. The present paper seeks to investigate the effects of this limiter starting with its original formulation in WAM cycle 3. As has been discussed in Section 2, the findings thus obtained are expected to be fairly representative for other limiters and physics parameterizations. To support this further, present findings will be compared to the results of HJ99 (obtained with WAM cycle 4 physics and their new limiter) where possible.

Limiters have been very successful in assuring numerical stability in the source term integration for third-generation wave models with large time steps, but have long been known to do so at the expense of suppressed initial wave growth rates (T92, HJ99, present Figs. 1 through 3). Not well known is that the use of a limiter may artificially increase growth rates by limiting negative spectral changes more severely than positive changes (Fig. 2). This behavior could nevertheless be observed in previous studies, as the growth rate $\partial H_s/\partial t$ for a given wave height typically is larger for a model based on a limiter than for the convergent solution in the intermediate growth stages (T92 Fig. 1, HJ99 Figs. 4 and 5).

The limiters might furthermore be considered very successful as stable integration can be obtained, even when the numerical time step grossly violates spectral time scales of wave growth. The present test shows that the impact of the limiter can be dramatic at the peak of the spectrum (Fig. 3), and need not be confined to the high-frequency equilibrium range of the spectrum as is often assumed. When the limiter is relaxed to allow the model to more closely follow its physical time scales, predicted wave heights H_s become more realistic. Ultimately, however, this will lead to a systematic overestimation of the peak frequency for a given wave height (Fig. 4, last paragraph of previous section). This behavior can also be observed in HJ99 Fig. 5. For the largest time step and hence least restrictive limiter, the wave heights are overestimated while peak frequencies show no noticeable errors for $t > 5$ h (dotted lines in left panels). Furthermore, friction velocities (u_*, upper right panel) show much larger errors than H_s. Because u_* in WAM cycle 4 is known to be sensitive to the high-frequency part of the spectrum, this is an additional indication of notable errors in the spectral shape.

Considering the above, there are two types of errors associated with limiters and the large model time steps they allow. The first is the systematically suppressed initial growth rate. Even if the limiter is relaxed as in the present study or as in HJ99, this behavior remains notable (Fig. 4, HJ99 Fig. 5). The relative importance of this error depends on two factors; the frequency at which initial growth starts, and the wind speed considered. The frequency range at which initial growth occurs is limited by

the maximum discrete model frequency f_{max}. If the inverse of this frequency is considered as a representative time scale for wave growth (as suggested in HJ99), a nondimensional time step $\Delta\tilde{t}$ can be defined as

$$\Delta\tilde{t} = f_{max}\Delta t. \tag{17}$$

In deep ocean wave model applications, and in the test case of HJ99, f_{max} is typically 0.4 Hz. For this relatively low maximum frequency, the initial growth error is moderate but notable as discussed above. If f_{max} is increased, the nondimensional time step $\Delta\tilde{t}_{max}$ will increase, and initial growth errors can be expected to increase accordingly. This may become a serious problem if the maximum frequency is increased to properly describe mixed wind-sea and swell systems on coastal scales (SWAN for this reason has a default $f_{max} = 1$ Hz), while an attempt is made to keep Δt similar to those of deep water models (e.g., Luo and Sclavo 1997; Monbaliu *et al.*, 2000). In terms of the wind speed of friction velocity, a nondimensional frequency f^* and time step Δt^* become

$$f^* = \frac{fu_*}{g}, \quad \Delta t^* = \frac{g\Delta t}{u_*} \tag{18}$$

The common test cases for effects of limiters are performed with fairly high wind speeds. For lower wind speeds, the nondimensional maximum frequency f^*_{max} will become lower, suggesting a more accurate description of initial growth. The nondimensional time step Δt^*, on the other hand, will become larger. This implies that 'fully grown' conditions will be reached in less time steps, which suggests that the overall accuracy of the growth curve calculation might be less than for higher wind speeds. Such behavior for lower wind speeds with more accurate initial growth, but with remaining errors in later growth stages, can easily be verified with calculations with reduced wind speeds (figures not presented here).

The second error introduced by the large time step is that of the spectral shape, where the shifting of f_p to lower frequencies with increasing H_s appears to lag systematically. In the present test, this error appears to cancel other shape errors in the initial growth stages (Fig. 4, $H_s < 1$ m), but becomes prominent for the intermediate growth stages ($1 < H_s < 6$ m). This error increases with the relaxation of the limiter. Because this error occurs at intermediate growth stages, it is expected to be potentially important at arbitrary wave model scales.

Previous authors have done an excellent job in reducing the time step dependency of third generation wave models as was discussed in the introduction. Nevertheless, the potential for the above errors remains in their models as discussed here. The main drawback of newly proposed limiters such as (7) is their nonconvergent behavior, which can only be justified for engineering applications (see Section 1). Even in such applications, the nonconvergence has one major drawback; it makes it virtually impossible to assess the magnitude of errors. In a convergent model, the magnitude of numerical errors can be assessed by rerunning selected cases with systematically reduced time steps. For a nonconvergent model the effects of reducing the time step are unpredictable, and there are no simple tools to estimate the magnitude of its errors.

The most elegant way to eliminate the above errors is to locally reduce the time step if necessary (T92, Hargreaves and Annan 2001), thus eliminating the need for a limiter.

The model structures of WAM and SWAN, however, do not allow such an approach. It therefore remains useful to investigate convergent limiters.

Upon casual inspection, the present asymmetric convergent limiter of Eqs. (13) through (16) may not look like a good solution compared to the new limiter of HJ99 (compare Fig. 4 and HJ99 Fig. 5). As discussed before, however, this is at least partially due to the more taxing initial conditions used in the present study. To obtain a more equal comparison, the present test has been rerun with a wind speed $u_{10}=25\,\mathrm{m\,s^{-1}}$ and an initial peak frequency $f_p=0.25\,\mathrm{Hz}$, as in HJ99. A single maximum asymmetry $\zeta_{\max}=3$ has been used, based on the results of Fig. 4. The results of this test are presented in Fig. 5.

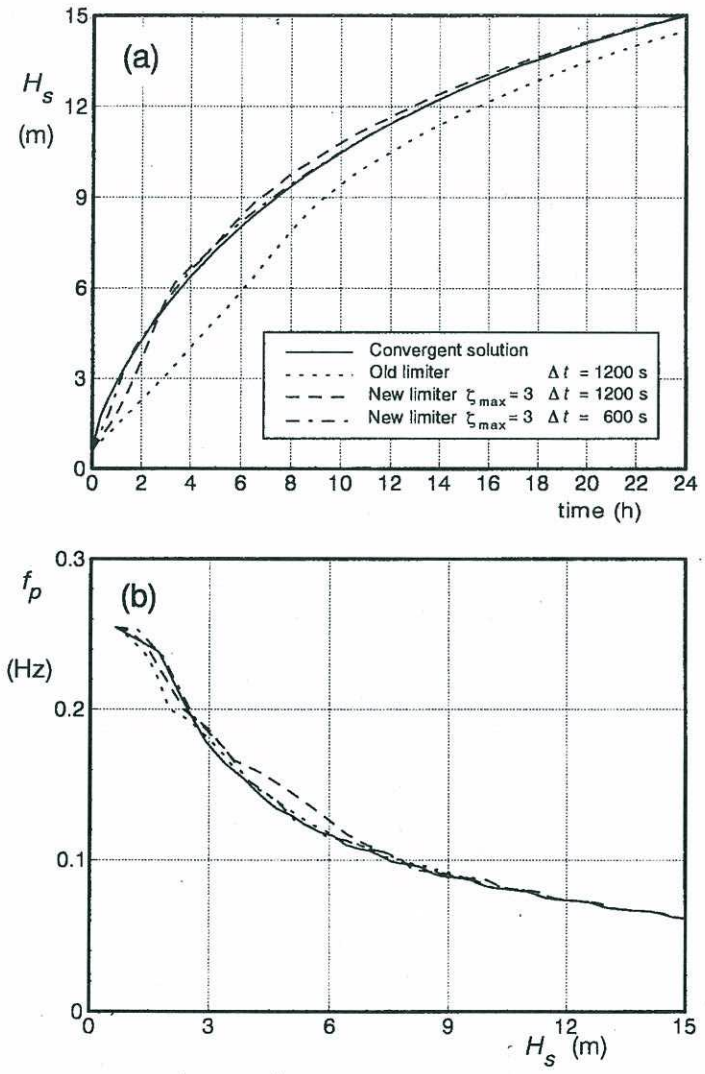

FIGURE 5 Like Fig. 1 for the test case of Hersbach and Janssen (1999). $U_{10}=25\,\mathrm{ms^{-1}}$, and initial conditions with $f_p=0.25\,\mathrm{Hz}$. Note that the solid and chain lines practically coincide.

A comparison of the present Fig. 5 with Fig. 4 of HJ99 shows qualitatively similar impacts of the original limiter in spite of the different physics parameterizations used (compare dotted and solid lines in both figures, note that the limiters are identical for $\Delta t = 1200$ s). It does appear, however, that the WAM cycle 4 physics as used in HJ99 is slightly more sensitive to the limiter than the WAM cycle 3 physics used here. The errors for the present convergent limiter (compare dashed and solid lines in Fig. 5) are both qualitatively and quantitatively similar to those of the nonconvergent limiter (7) of HJ99 (their Fig. 5, compare dotted and solid lines). As expected, effects of reducing the time step to $\Delta t = 600$ s are very different for both limiters. For the present limiter, the reduction of the time step virtually eliminates time discretization errors in the model (compare chain and solid lines in Fig. 5). In HJ99 the reduction of the time step has much less impact, particularly regarding initial growth errors. Considering this, the new asymmetric limiter appears to behave comparably to the HJ99 limiter, but does so without sacrificing the convergence of the numerical scheme.

It should be noted that the asymmetric limiter was developed for the present study mainly to illustrate the effects of a progressively relaxed yet convergent limiter. At best, the present study can be considered as a proof of concept for such a limiter. If such a limiter is to be applied to operational models, its behavior needs to be tested much more rigorously than in the present paper. Furthermore, the averaging scheme in Eq. (14) as well as the limitation in Eq. (16) are chosen in an *ad hoc* manner and require additional attention. Finally, an important aspect of the limiter is the stabilization of the source term integration for high frequencies. Other improvements can also improve the stability of the source term integration in this spectral regime, and should therefore be considered simultaneously. Such improvements are for instance (i) Non-central time integration as first suggested by Hargreaves and Annan (oral presentation at 1996 WISE meeting, 2000, 2001), and used by HJ99. (ii) Application of a smooth transition to the parametric tail as suggested by Tolman and Chalikov (1996). (iii) Development of improved parameterizations of the nonlinear interactions without the spuriously large interactions at high frequencies as is common in the presently used Discrete Interaction Approximation (see Hasselmann *et al.*, 1985 Fig. 7).

6. CONCLUSION

Most third-generation wind wave models use a so-called limiter to assure numerical stability when the model is integrated with economically feasible time steps. This limiter has previously been shown to result in a noticeable dependence of model results on the discrete time step of the numerical integration. In recent papers, this time step dependency has been removed by scaling the limiter with the time step. This approach removes the numerical convergence from the integration scheme. This is contrary to rudimentary numerical principles and to the philosophy behind third-generation wave models.

The present study investigates the effects of limiters starting from the original WAM cycle 3 approach. It is shown that a conventional convergent limiter allows for stable model results even if the discrete model time step grossly violates time scales of (initial) wave growth. Contrary to common belief, the effects of the limiter are not necessarily confined to the equilibrium range of the spectrum, and the limiter can result in artificially increased growth rates in the intermediate growth stages. In its original form,

errors induced by the limiter are dominated by suppressed initial growth rates. Such errors become more prominent if the maximum model frequency f_{max} is increased.

If the limiter is relaxed, errors in the initial growth rates can be significantly reduced. Ultimately, however, the relaxation of the limiter in combination with large time steps will result in errors in the spectral shape, where the evolution of the peak frequency f_p will become too slow for the given wave height evolution. Although these results have explicitly been obtained for the WAM cycle 3 physics, similar behavior appears evident in WAM cycle 4 when the limiter is relaxed (see HJ99).

The present results for relaxed limiters have been obtained with a new asymmetric limiter that retains full convergence. Because this limiter appears to result in similar model behavior than can be obtained with previously suggested nonconvergent limiters, there appears to be no good justification to use the latter. The present study, however, can only be seen as a proof of concept for the new limiter. Much work needs to be done before it can be safely implemented in operational models.

Acknowledgments

The author thanks Hargreaves and Annan, and Hersbach and Janssen for supplying preprints of their manuscripts during early stages of this study. The author thanks L.H. Holthuijsen, N. Booij, D.V. Chalikov and D.B. Rao for their comments on early drafts of this paper.

References

Booij, N., Holthuijsen, L.H. and Ris, R.C. (1996). The "SWAN" wave model for shallow water. *Proc. 25th Int. Conf. Coastal Engng.*, ASCE, Orlando. pp. 668–676.

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

Hargreaves, J.C. and Annan, J.D. (1998). Integration of source terms in WAM. *Proceedings of the 5th International Workshop on Wave Forecasting and Hindcasting*. pp. 128–133.

Hargreaves, J.C. and Annan, J.D. (2001). Comments on "Improvement of the short fetch behavior in the WAM model". *J. Atmos. Oceanic Techn.*, **18**, 711–715.

Hasselmann, S., Hasselmann, K., Allender, J.H. and 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. *J. Phys. Oceanogr.*, **15**, 1378–1391.

Hersbach, H. (1996). On the improvement of the WAM cycle 4 model for small grid spacings, KNMI memorandum OO-96-01. 4 pp.

Hersbach, H. and Janssen, P.A.E.M. (1999). Improvement of the short fetch behavior in the WAM model. *J. Atmos. Oceanic Techn.*, **16**, 884–892.

Hersbach, H. and Janssen, P.A.E.M. (2001). Reply to comments on "Improvement of the short fetch behavior in the WAM model". *J. Atmos. Oceanic Techn.*, **18**, 716–721.

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

Luo, W. and Sclavo, M. (1997). Improvement of the third generation WAM model (cycle 4) for application in nearshore regions. *POL Internal Document*, **116**.

Monbaliu, J., Padilla-Hernandez, R., Hargreaves, J.C., Carretero Albiach, J.C., Luo, W., Sclavo, M. and Günter, H. (2000). The spectral wave model, WAM, adapted for applications with high spatial resolution. *Coastal Eng.*, **41**, 41–62.

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.

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

Tolman, H.L. (1992). Effects of numerics on the physics in a third-generation wind-wave model. *J. Phys. Oceanogr.*, **22**, 1095–1111.
Tolman, H.L. and Chalikov, D.V. (1996). Source terms in a third-generation wind-wave model. *J. Phys. Oceanogr.*, **26**, 2497–2518.
Tolman, H.L. (1999). User manual and system documentation of WAVE WATCH III version 1.18. NOAA/NWS/NCEP/OMB technical note 166. 110 pp. Available from http://polar.wwb.noaa.gov/waves/wavewatch.
WAMDIG (1998). The WAM model – a third-generation ocean wave prediction model. *J. Phys. Oceanogr.*, **18**, 1775–1810.

APPENDIX

Dynamically Adjusted Time Steps

In T92, the limiter $\mathcal{L}_0$ is used to calculate the time step as

$$\Delta t \propto \min_{\forall\theta,f}\left(\frac{\mathcal{L}_0(f)}{|S(f,\theta)|}\right), \tag{A.1}$$

for an explicit Euler scheme, or according to his Eq. (16) for the semi-implicit scheme of WAM. Hersbach and Janssen (2001) label this approach as not suitable for operational application because their analysis shows that $\Delta t \downarrow 0$ for $S \uparrow 0$ while $F \downarrow 0$. From Eq. (A.1), however, it is obvious that $\Delta t \to \infty$ for $S \to 0$ for all F. The misunderstanding of Hersbach and Janssen (2001) is based on their definition of a limiter as proportional to the instantaneous spectrum [their Eq. (4)], which would lead to a time step

$$\Delta t \propto \min_{\forall\theta,f}\left(\frac{\beta F(f,\theta)}{|S(f,\theta)|}\right), \tag{A.2}$$

where β is relative limitation level. If such a relative limiter is used to adjust the time step, it may lead to the above pathological behavior (if the equations used illustrate this indeed are representative for source terms). More importantly, it will by definition not allow for wave growth starting from $F=0$.

In the released version of WAVEWATCH (Tolman 1999), a time step based on a relative limitation (A.2) was added to a time step based on the parametric limitation (A.1) to assure accurate swell dissipation due to various processes (see Tolman 1999 pages 37–38 for details). To avoid pathological behavior, the latter time step is filtered as

$$\Delta t \propto \min_{\forall\theta,f}\left(\frac{\max[\beta F(f,\theta),\ \mathcal{L}_0(f_{\max})]}{|S(f,\theta)|}\right), \tag{A.3}$$

which is practice has proven to be accurate and economical for a wide range of model applications.

As a final economical safeguard, the distributed version of WAVEWATCH allows the user to set a minimum allowed time step. If the dynamically calculated time step is smaller than this minimum, the integration scheme reverts to the original convergent limiter. Typically this minimum time step is set to 120–300 s, at which time step level the impact of a convergent limiter is generally negligible. This minimum time step as well as limiter levels used in the calculation of the time step are user-defined in this model. Their impact is therefore easily checked for any application.