

Phase-averaged wave models

By Erick Rogers, NRL 7322

Acknowledgments

Yalin Fan (NRL 7322) and Silvia Gremes (U. New Orleans) contributed to Figure 5.

1. Introduction

In simplest terms, a phase-averaged wave model is one that does not treat waves individually but instead uses the wave *spectrum* as the prognostic variable. These models are sometimes called “spectral models” for this reason, though in truth, a phase-averaged model can be monochromatic, and a phase-resolving model can be used to ultimately predict spectra. We do not intend to cover the topic of phase-averaged wave modeling comprehensively herein. This is already treated excellently and in greater detail by longer works such as Komen et al. (1994), Young (1999), Cavaleri et al. (2007) and Holthuijsen (2010). Our goal here is to offer alternative perspective and include topics and frank discussion which are not typically included. Emphasis is on the wave models SWAN (Booij et al. 1991) and WAVEWATCH III® (“WW3”, Tolman (1991) WW3DG (2016)), since this author has more firsthand experience with these two models than any other model. These two are good examples of phase-averaged models since they are open source, relatively modern (e.g. using message passing interface (MPI) parallelization and including modern physics parameterizations) and enjoy widespread, international use. Both models originally borrowed heavily from features of a common ancestor, the WAM model (WAMDIG 1988), which was a widely shared community model. Today, development of WAM is fractured, e.g. the European Centre for Medium-Range Weather Forecasts continue to develop a proprietary, fully modernized version: ECWAM. Herein, we use the term “3GWAM” to refer to the “third generation” of wave models, which includes SWAN, WW3, WAM(s) and a handful of other models. In its formal definition, the “third generation” label implies that it is a model which explicitly includes relevant physical processes and does not make a priori assumptions about spectral shape (Komen et al. 1994).

The governing equation of phase-averaged models is the wave energy balance equation (most versions of WAM) or wave action balance equation (SWAN, WW3). The models operate by integrating this equation. There are five independent variables: wave frequency (or alternatively, wavenumber), wave direction (or alternatively, the Cartesian form of wavenumber), geographic position (meters, or degrees latitude/longitude), and time. Thus, some possible forms of the prognostic variable (spectral density) are: $E(f, \theta, x, y, t)$, $N(f, \theta, \phi, \lambda, t)$, $N(k, \theta, \phi, \lambda, t)$, and $N(k_x, k_y, \phi, \lambda, t)$.

We introduce phase-averaged models by contrasting with phase-resolving models:

Advantages

1. There is no requirement of having a minimum number of computational “points per wavelength”, which is particularly important if shorter waves are being modeled, e.g. the high frequency tail. It is challenging to efficiently represent many length scales simultaneously in a phase-resolving model.

Multimedia Report

2. Geographic resolution is flexible, which implies that a model design can conform to the resolution of expected scale of variability, as appropriate. These scales tend to follow model input; we discuss this below.

Disadvantages

1. Phase-averaged models tend to include more parameterization. This is especially true for source terms. Though, phase resolving models typically do require their own parameterizations for anything that is not explicitly computed (breaking, capillary waves, atmosphere, turbulence, viscosity, compressibility, etc.).
2. Phase-averaged models tend to utilize a larger number of assumptions built into the governing equation, e.g. that the sea surface is Gaussian. For kinematics, phase-averaged models usually rely on linear wave theory, while some phase-resolving models do not.
3. Though today's phase-averaged models offer an enormous number of output variables computed from the spectra or the source terms, the phase-resolving model can, in principle, provide even more, e.g. some explicitly model wave shape and wave grouping. Freak wave prediction is possible using a phase-averaged wave model, but this relies on associations and correlations rather than explicit modeling (e.g. Janssen 2002).

A careful reader will have noticed that computation time was not mentioned above. Though some phase-resolving models are exceedingly expensive to run, there are sufficient counter examples (fast phase-resolving models and slow phase-averaged models), that blanket statements are impossible. Similarly, though phase-resolving modeling tends to be more academic, and phase-averaged modeling more suitable for operational use, one can find counter-examples.

3GWAM output has many uses and forms. One common use is in operational nowcasting and forecasting of wave conditions, and in that context, perhaps the most common product is the map of significant waveheight H_s . An example is given in Figure 1.

2. Kinematics

Here, we use “kinematics” to refer to processes described using the advection term of the governing equation, $\nabla \cdot \vec{c}N$. A brief list is given here:

1. Refraction by bathymetry and currents. Obviously, this creates a change in wave direction, and perhaps less obviously, can change the energy level locally via focusing and defocusing. This occurs when the phase velocity is non-uniform along wave crests (perpendicular to the axis of propagation).
2. Shoaling by bathymetry, and analogous behavior in cases of horizontally sheared currents. This occurs with the group velocity (plus currents) is non-uniform along the axis of propagation.
3. Anything that causes the dispersion relation to deviate from its conventional form will necessarily produce behavior analogous to (1) and (2) above. Two examples are the effect of a viscous mud layer at the seafloor (e.g. Dalrymple and Liu (1978)) and the effect of a viscous ice cover at the surface (e.g. Keller 1998). At time of writing, these kinematic effects are limited to academic applications.
4. Advection in frequency space as waves encounter horizontal shear in currents.
5. Diffraction has been implemented in SWAN via modifications to the advection terms (Holthuijsen et al. 2003). However, there are difficulties with this approach and it is our

opinion that this process is best predicted using a phase-resolving model, especially in the vicinity of engineering structures.

3. Source terms

We briefly list source terms available in 3GWAMs here, and in contrast to those prior descriptions, the following includes discussion of the compromises, approximations, and other artifice of numerical modeling.

Exponential wind input. Some methods used today are 1) the parameterization of Komen et al. (1984), loosely based on experimental work of Snyder et al. (1981), 2) the parameterization of Janssen (1991), loosely based on the quasi-laminar theory of Miles (1957), 3) models based on fitting to a boundary layer model, e.g. Tolman and Chalikov (1996), and 4) the parameterization of Donelan et al. (2006), largely empirical but drawing inspiration from the sheltering theory of Jeffreys (1924). There are large differences between these methods, even in fundamental features such as the directional distribution, and the integrated atmosphere-wave momentum flux.

Linear wind input. Exponential wind input requires energy to increase energy, so it does not act on a calm surface. In wave models, this problem can be addressed in one of two ways. One is to “seed” the model state with some very small initial energy. The other approach is to implement a linear wind input source term. The Phillips resonance mechanism is linear, so it is commonly used. Both SWAN and WW3 have linear wind input which is nominally based on this mechanism, following the work of Tolman (1992) which is, in turn, based on Cavaleri and Malanotte-Rizzoli (1981). In truth, the connection to the Phillips resonance is practically non-existent, particularly with respect to the directional resonance at $\theta_r = \pm \cos^{-1}(C/U)$. However, it can be argued that the accuracy of the source term is inconsequential, because it has little impact on the modeling outcome beyond the first stages of wave growth.

Steepness-limited breaking (whitecapping). Breaking is strongly phase-dependent, so a spectral source term is necessarily highly parameterized. Even so, this source term has seen remarkable improvement since the early, simple formulation by Komen et al. (1984). The primary shortcoming of the early formulation was non-physical features such as the breaking of swell, and the unintended non-physical influence of wind sea on swell and vice versa (see Rogers et al. 2003). The key breakthrough was the effective incorporation of the concept of a spectral threshold, below which no breaking occurs, and above which breaking increases nonlinearly (e.g. Babanin and van der Westhuysen 2008; Ardhuin et al. 2010). Also important was the recognition that multiple physical mechanisms can contribute to breaking. Modern formulations tend to include at least two of the following: 1) breaking caused by instability of waves at the frequency in question (independent of other frequencies), 2) dissipation caused by turbulence and instability associated with the breaking of waves that are longer than the frequency in question, and 3) breaking associated with the “concertina effect”, in which short waves are modulated by the orbital velocities of long waves.

Non-breaking dissipation. With the threshold mechanism for wave breaking in modern wave models, it becomes necessary to include an additional mechanism to produce the slow dissipation which is implied from observations of swell propagating across long distances (e.g. Ardhuin et al. 2009). One postulated cause for this is friction at the near-surface atmospheric boundary

layer, e.g. with energy lost to turbulent eddies in the air (Ardhuin et al. 2010). Another postulated cause is generation of turbulence in the water as wave orbital velocities interact with background turbulence (Babanin 2006). Though the causes are different, the equations are nearly equivalent (Babanin 2011). A third cause may be interpreted as a re-imagining of the first. Here, rather than treating the losses to the atmosphere using a friction model, they are computed by including a “negative wind input”. This wind input term is computed the same way as the corresponding positive wind input, but reduced by some factor, taking the paradigm used in the experimental study of Donelan (1999).

Nonlinear interactions. This refers to energy- and momentum-conserving redistribution of energy within the spectrum. In shallow water, the fast near-resonant triad interactions are the most important, but these have historically been challenging to represent in phase-averaged models (Cavaleri et al. 2007). In the open ocean, the resonant four-wave interactions become important. These interactions are weak, requiring many wavelengths to effect significant change to the spectrum, but over the large scales of the open ocean, they become very important. Solution of these interactions has been a focus of much research for decades, and solution methods range from the fast and approximate DIA (“Discrete Interaction Approximation”, Hasselmann et al. 1985) to exact methods which can be computationally slower by two orders of magnitude (see Tolman 2011). One exact method is Hasselmann and Hasselmann (1985), and there are a number of intermediate methods, e.g. van Vledder (2012), Tolman (2013). The implications of the reduced accuracy of DIA is discussed below.

Surf breaking. This source term is sometimes referred to as depth-limited breaking. In principle breaking in deep water and shallow water could be represented with a single unifying source function (e.g. Filipot and Ardhuin 2012), but in practice, this is seldom done, and they are treated separately. The deep-water breaking formulations typically do not act fast enough to produce realistic profiles of waveheight across a beach profile, and so the surf breaking (e.g. Battjes and Janssen 1978) is included as an engineering solution. The source term acts as a “safety valve” to prevent extremely large wave heights in the nearshore, keeping in mind that in the case without dissipation, energy goes to infinity as depth approaches zero, due to shoaling.

Bottom friction. In finite depths, wave orbital motion interacts with rigid seafloor, and that near-bottom boundary layer produces dissipation of the wave energy, e.g. Madsen et al. (1988). These source terms tend to be highly parameterized, and typically suffer from insufficient operational knowledge of the seafloor roughness, and variability of the same. However, the models at least capture the appropriate qualitative behavior, since the spectral calculation is based on orbital motion, which is readily calculated.

Fluid mud (non-rigid seafloor). When waves propagate in shallow and intermediate depths with a muddy seafloor, they exert time-varying pressure, which creates a two-layer wave system, and energy lost to viscosity in the mud layer implies a damping of the surface waves. This damping can be very strong, especially in shallower depths. This can be estimated using mathematical models, e.g. Dalrymple and Liu (1978), Ng (2000). Models have been implemented in SWAN and WW3 (Winterwerp et al. (2007); Rogers and Holland (2009); Rogers and Orzech (2013)). Unfortunately, inputs are notoriously difficult to define, so uses of these parameterizations tend to be academic. Viscosity is difficult to estimate. Mud thickness can be obtained from core

Multimedia Report

samples, but since it is always inhomogeneous and often nonstationary, sparse sampling has limited value to the modeler. Also, the thickness of the mud is not in fact a useful variable. Rather, the model needs an estimate of the fluidized mud layer thickness; the immobile fraction of a mud layer does not contribute to dissipation.

Dissipation by vegetation. Marine vegetation can also dissipate wave energy. Wave motion in sea grass has been modeled as flow in and around narrow cylinders. The method of Dalrymple et al. (1984) has been implemented in SWAN (Suzuki et al. 2011).

Dissipation by sea ice. Traditionally, the effect of sea ice on waves has been represented in wave models by deactivated grid points, i.e. treating them as land. This approach has been used with WAM, SWAN, and WW3. More recently, dissipation by sea ice has been implemented in at least two models, ECWAM and WW3, using physics parameterizations (Doble and Bidlot 2013; Rogers and Orzech 2013; WW3DG 2014, 2016). There are three major groups of methods: 1) mathematical models which consider the dissipation that occurs within the ice layer, 2) mathematical models which consider the dissipation that occurs below the ice layer, and 3) fully empirical parameterizations. In the first case, the losses are associated with frictions. For the case of ice in a continuous sheet, there can be internal friction (hysteresis). For the case of suspensions of ice, e.g. brash and frazil ice, there can be losses from pumping of liquid water through the gaps between ice. Below the ice, the losses are associated with the boundary layer friction, and the boundary layer may be laminar or turbulent, just as with friction at the seafloor. There are further parallels: like seafloor bottom friction dissipation, all three methods of predicting dissipation by sea ice require information that is usually difficult to obtain: ice type and ice characteristics (rheology). Parameterizations and settings which provide good skill for one ice type cannot be expected to have skill with other ice types. However, there are causes for optimism. First, there has been significant progress recently to produce new, more accurate observational datasets for use in model development, e.g. Thomson et al. (2018), Rogers et al. (2018). Second, as with bottom friction, qualitative behavior can already be adequately represented in models, as there is a robust and intuitive increase in dissipation rate with wave frequency, e.g. Figure 2. This feature did not exist in 3GWAMs prior to 2010. Selection of the power dependency of this dissipation rate vs. frequency relation is not fully resolved, but is coming into better focus with new research (Meylan et al. 2018).

Scattering by sea ice. When wavelength is comparable to the floe size, wave-ice interaction enters a “scattering regime” (Bennetts and Squire 2015), in which the process should not be neglected. Wave energy may also be scattered in the case of propagation of swell through continuous ice, as the waves encounter sharp variations in ice thickness (Squire et al. 2009). This is a non-dissipative (conservative) process which is represented in WW3 using redistribution of energy within the spectrum (WW3DG 2016; Ardhuin et al. 2016).

Scattering by irregular seafloor. Bragg resonance with topographical features can result in scattering. This can be either forward or backward scattering. Ardhuin and Herbers (2002) found that for a case on the North Carolina shelf, forward scattering predominates, and leads to significant broadening of the spectrum. This process has been included in WW3, but like some other source terms in this list, application tends to be for academic purposes: it requires a spectrum of the seafloor and can be expensive to compute.

Reflection. Reflection from steep coastlines (Ardhuin and Roland 2012) and icebergs (Ardhuin et al. 2011) has been implemented in WW3. In the former study, the authors find that errors in directional spread are improved by including this effect. Reflected waves have a somewhat unique role whereby they produce partial standing waves, which penetrate to extreme depths and produce seismic noise (Longuet-Higgins 1950).

Deep-water physics packages

Collectively, the wind input source function(s), steepness-limited breaking, and non-breaking dissipation are known as “source term (ST) packages”. In WW3, these include: ST1 (Komen et al. 1984; Booij et al. 1999), ST2 (Tolman and Chalikov 1996), ST3 (Bidlot et al. 2005), ST4 (Ardhuin et al. 2010), and ST6 (Rogers et al. 2012; Zieger et al. 2015; Rapizo et al. 2017; Liu et al., submitted). SWAN includes several other source term packages, such as ST1, ST6, and one by van der Westhuysen et al. (2007). Other source terms are selected independently from the ST packages.

These three source terms are grouped together because their calibration is interdependent, and together they largely control the rate of growth and decay of wave energy in deep, open water (i.e. where depth-dependent processes and ice are absent). The four-wave nonlinear interactions are of course also very important in this situation, especially for spectral shape (e.g. Resio et al. 2016), but as mentioned above, it is selected independently from the ST packages, so for example, ST6 can be applied in conjunction with any solver for the four-wave nonlinear interaction source term in SWAN and WW3. This leads to the obvious question: does a change in nonlinear solver necessitate a retuning of the source term package? The answer is not simple. Tolman (2011) asserts that any change to the nonlinear solver does require a recalibration of ST2. Results of Perrie et al. (2013) imply that this is also true in ST4 if DIA is replaced with a more accurate solver. Rogers and van Vledder (2013) find that change to the nonlinear solver does *not* require recalibration of ST6 in SWAN; however, it should be kept in mind that there was little or no swell in that study. Liu et al. (subm.) find that a change to the nonlinear solver again does not require recalibration of wind input and breaking terms of ST6 in WW3 but they do find it necessary to recalibrate the swell dissipation in a global application of ST6. This outcome is unsurprising: DIA is known to push too much energy to the low-frequency side of the spectral peak (e.g. Rogers and van Vledder 2013) and in an global application, this has implications for the calibration of the swell dissipation.

Thorough comparison of the ST packages is beyond the scope of this manuscript. This can be found in Stopa et al. (2016), van Vledder et al. (2016), Liu et al. (subm.) and other publications. However, we present here two idealized comparisons. Figure 3 shows the fetch and duration waveheight growth curves for ST2, ST3, ST4, and ST6 with wind speed $U_{10} = 15$ m/s. These simulations are initialized from rest. The fetch and duration values reported by Moskowitz (1964) for fully developed conditions are marked on each plot, though it may be assumed that observed cases did *not* start from rest. The significant waveheight predicted by the Pierson-Moskowitz (1964) (“PM”) parametric model for fully developed conditions is 5.5 m at this wind speed. Vertical contours indicate duration-limited conditions. Horizontal contours indicate fetch-limited conditions. Such growth curves are a useful indicator of model behavior. For example, ST2 tends to be slower than WAM4 during initial growth while overpredicting energy as

windsea transitions to swell (Tolman, personal communication), and this is reflected in the figures. The plots indicate strong similarity in ST3, ST4, and ST6; especially the latter two. Notably, the models are unanimous that the fetch/duration values reported by Moskowitz (1964) *do not* correspond to fully developed conditions. In real ocean cases of very large fetch/duration, that fetch/duration is defined by the storm itself, and we postulate that the apparent cessation of growth observed by Moskowitz was associated with the inevitable slackening of winds, rather than being associated with the approach towards an asymptotic limit under constant winds. The latter is, of course, the traditional interpretation of the PM spectrum, e.g. Booij et al. (1999).

Of course, even when waveheight is identical between two models, the spectrum may be very different. Figure 4 compares frequency distributions of energy and the three primary deep-water source functions (S_{in} , S_{ds} , S_{nl4}). These are for the same wind speed, with unlimited fetch, and 15 hours duration. In terms of spectral shape, the more modern models (ST4, ST6) are in good agreement, while the older ST2 is an outlier, with a lower dominant wave period. Wind input and whitecapping in ST6 is much stronger in the high frequencies than ST2 and ST4. In context of whitecapping, this implies that ST6 is closest to consistency with Pushkarev and Zakharov (2016), who argue that there is no dissipation near the spectral peak. The strong differences in frequency distribution of wind input has implications for coupling to other models, since conversion from wind input to wave-supported stress (i.e. from energy flux to momentum flux) involves a factor $1/C$ where C is the phase velocity.

4. Grids

3GWAMs are solved on grids in both spectral and geographic space. The general objective with model resolution is to sufficiently resolve the variability that would be exhibited by the model if it was solved without discretization (the latter being generally impossible, of course). This required resolution is primarily dictated by the model forcing, e.g. the scales of variation of the driving winds, the currents, the ice, or the bathymetry/coastline.

Frequency space is defined on a logarithmic spacing, implying smaller steps for lower frequencies. When the DIA is used for four-wave nonlinear interactions, the spacing is usually according to $f_i = 1.1f_{i-1}$, where f_i indicates a bin in the frequency grid. The increment factor 1.1 is preferred because it was used in the original development and calibration of the DIA. For general cases, e.g. a global wave model, the starting frequency has traditionally been around 0.04 Hz, but it can be set significantly higher in restricted basins (e.g. lakes and enclosed seas) with defined climatology for expected spectral range. More recently, lower starting frequencies have been adopted in global wave models, e.g. 0.035 Hz, recognizing that extreme extratropical storms can produce energy at these frequencies. Even when the low frequency bands (0.035-0.40 Hz) contribute little to the waveheight within the storm, the swell energy that radiates outward at these frequencies can be important at remote locations, especially when energy at higher frequencies is absent or suppressed (e.g. Ardhuin et al. 2016). The high frequency limit is typically between 0.5 Hz and 1.0 Hz with modern physics packages, e.g. 0.6 to 0.7 Hz is used in WW3 by Ardhuin et al. (2010). Beyond this limit of the prognostic frequency grid, a diagnostic tail is added; this can be either fixed (e.g. 1.0 Hz in Booij et al. (1999) and generally recommended for SWAN) or a flexible limit based on local wind speed and/or mean period (e.g. Tolman 1991). Of course, if a 3GWAM is being used to model only swell in restricted domains, forced at lateral boundaries, one may safely use a narrower frequency range.

Multimedia Report

For the directional grid, uniform spacing is used, typically around 10° , which implies 36 directional bins to represent waves from any direction. Though again, for modeling of only swell, with restricted coastal domains and forcing at the boundaries, one may elect to omit some directions (e.g. ones directed offshore) to reduce computation time.

Geographic grids exist in three main categories: structured regular, structured irregular, and unstructured. Further, they can be in a Cartesian system (x, y) or a spherical (latitude/longitude) system (ϕ, λ) . “Structured” is used here to indicate that the grid is logically rectangular. A structured irregular grid in Cartesian system would have all grid points defined on $x(i = 1 \dots n, j = 1 \dots m)$, $y(i = 1 \dots n, j = 1 \dots m)$ where n and m are the number of rows and columns (thus “logically rectangular”). “Regular” is used here to indicate uniform spacing, e.g. $(\Delta x, \Delta y)$. A Cartesian structured regular grid also has points defined on $x(i, j)$, $y(i, j)$, but because of uniform spacing, this simplifies to $x(i)$, $y(j)$. Unstructured grids are not logically rectangular; here grid points are defined in a list, $x(1 \dots n)$, $y(1 \dots n)$, where n is the number of nodes. Unstructured grids can be composed of triangular grid cells, but this is just one type. Another example of an unstructured grid is one which is made by taking a regular grid and deleting some grid points, e.g. to have larger spacing offshore or larger longitude spacing $\Delta \lambda$ at extreme latitudes.

Selection of grid type often involves weighing the cost of grid creation against the cost of running an inefficient model. Regular grids are the easiest to create, and in the context of SWAN, can be changed by simply modifying a few numbers in an input file. However, they can be inefficient. For example, high resolution may be applied to describe a coastline and bathymetry, but this same high resolution is then applied in offshore portions of the same grid, where it is not needed. A global grid with regular latitude/longitude spacing implies highly non-uniform spacing in terms of real distances, which in case of conditionally stable propagation, forces use of an inefficiently small time step (a small fraction of grid cells dictate a small time step for the entire grid). Irregular and unstructured grids can address these inefficiencies. For example, in the global context, we can use regular spacing a low latitudes, Lambert conformal projections at mid-latitudes, and polar stereographic projections at high latitudes to create modeling systems which permit large time steps for propagation. For modeling coastlines, the unstructured grids are most efficient, since high resolution can be applied where it is needed, and coarse resolution applied offshore, where wave conditions are more homogeneous. The cost, of course, is that irregular grids and unstructured grids—especially the latter—can require more effort to create and manage. Creation of unstructured grids usually requires specialized software, and the creation of boundary forcing for grids with irregular boundaries can be troublesome. Also, there exists a caveat for the case of unstructured grids which apply high resolution where needed: these grids usually require an unconditionally stable propagation scheme, since otherwise, the smallest grid cell will dictate the propagation time step.

Tolman (2008) created an ingenious approach for WW3 which combines the simplicity of regular grids with the efficiency of irregular and unstructured grids. This uses a regular grid, but with offshore portions masked out, meaning that those locations are computed in a host grid which provides boundary forcing. This idea unfortunately creates a new challenge of defining irregular, internal boundaries not oriented with the four cardinal directions. Tolman (2008)

implemented this idea alongside a “mosaic approach” of two-way nesting, which automates the boundary management, though not the grid creation.

The strategy of “applying high resolution only where it is needed” is problematic in cases where the determining factors are non-stationary: if it is desirable to resolve high-resolution details of the dynamic model inputs (e.g. surface currents, winds), then adaptive (non-stationary) grid methods should be used, e.g. Popinet et al. (2010). Such technology is still relatively new.

5. Numerics

Discretization in numerical models results in numerical error, and coarser discretization generally implies larger error. However, it is not the resolution itself that increases error, but rather the resolution relative to the curvature (and other spatial derivatives) of the thing being propagated: in our case, spectral density. Also important: the numerical scheme used and the Courant number, $\mu = C\Delta t/\Delta x$, where C is the speed at which energy is advected, usually taken as the sum of the group velocity and the mean current. Different schemes have different dependency on resolution and Courant number. For example, the first order scheme of SWAN has less error with smaller Courant numbers. The first and third order schemes of WAM and WW3 have opposite behavior: they have less error with larger Courant number, though they become unstable beyond a value near 1.0.

Numerical error falls into two broad categories: diffusion error and phase error. Again, different numerical schemes include these in different proportions. Diffusion tends to smooth a swell field as it propagates, so that maxima are reduced and minima increased, and also it tends to bleed energy into regions which it would not propagate into using an exact method (e.g. ray tracing). One example of the latter is excessive propagation of wave energy into sheltered regions behind islands, and this error can become so large that it increases RMS error in comparisons against observations (Rogers et al. 2002). Phase error is the deviation of the numerical propagation speed from the true (intended) propagation speed for individual Fourier components of the numerical solution (Petit 2001). When these components separate during propagation, they can manifest as oscillations in the solution if they are not damped by diffusion, and these are known as “wiggles”. These wiggles tend to be most severe in higher order implicit schemes such as the one used in SWAN. Thus, a user is faced with the dilemma of choosing between larger numerical error with no wiggles (the diffusive first order implicit scheme) vs. numerical error that is significantly smaller but also more visibly non-physical. Further, unfortunately, the selection is not case-dependent, since the same thing that creates diffusion in the first order scheme (curvature of the wave field) also creates wiggles in the higher order scheme. Thus the selection is largely based on user preference.

The above dilemma is an outcome of the use of implicit schemes in SWAN which is in turn a result of the requirement of unconditional stability, which is in turn a result of SWAN’s primary design criterion: efficient modeling at high resolution. WAM and WW3 take a different approach. These models are originally designed for larger-scale modeling, which implies coarser resolution (e.g. greater than 2 km), so conditional stability is not objectionable. This allows use of explicit schemes which have higher overall accuracy and speed than their counterparts in SWAN.

Multimedia Report

WW3 is not without problems, however. It separates x- and y-propagation into separate sequential steps, in contrast to SWAN which solves for propagation in both dimensions with a single operation. In extreme canonical cases, this “splitting error” results in a distortion of features: initially round features become more square in appearance after propagation.

The higher order schemes of both SWAN and WW3 have numerical diffusion that is so small that it produces yet another problem: the “Garden Sprinkler Effect”, whereby the spectral discretization causes distinct features to become visible in the geographic distribution of swell fields after propagation. This is most severe when spectral discretization is coarse. With first order schemes, the effect is mostly masked by diffusion. In both SWAN and WW3, methods exist to counteract the effect (Booij and Holthuijsen 1987; Tolman 2002).

The problems mentioned above are not necessarily apparent in skill scores, particularly in the open ocean where the curvature of the wave field tends to be smaller. Rather they are first evident as aesthetic problems, e.g. in spatial distribution of swell energy. In a time series comparison against observations, diffusion will manifest as a reduction in the variance of the model time series, but again, this does not necessarily affect traditional skill scores such as RMS error when larger errors (e.g. from wind forcing) dominate.

SWAN has an additional method for improving the efficiency of high resolution modeling: stationary computations. Here, the assumption is made that the wave are in steady state, having achieved balance with the local winds and boundary conditions. This is more appropriate for small regions, e.g. 25 km by 25 km, and error becomes significant for larger areas, e.g. with waves reacting too quickly to changes in wind, or swells arriving too early (Rogers et al. 2007).

The time step also plays a role in numerical error. Similar to the x/y propagation issue above, WW3 separates the computations of propagation from computation of source terms, where SWAN solves this simultaneously. The impact of the split time stepping in WW3 is largely in the control of the user. By using infrequent reconciliation of source term integration and propagation, computations can be accelerated, but with a loss of accuracy.

SWAN has a time step problem which is arguably more severe. It applies a “limiter” such that large changes to the spectrum (from one time step to the next) are prevented. This improves model stability. However, it means that the model solution can become dependent on the time step size (Tolman 2002). This implies that in cases of rapid wave growth, e.g. a sudden squall following relative calm, the growth will be underpredicted by SWAN during the initial hours. However, like with the case of propagation error, the problem is most evident in idealized cases. In hindcasts of real scenarios, this author has not observed significant sensitivity to time step size. WW3 solves the limiter problem by using dynamically reducing the time step size according to strength of the net source terms (Tolman 1992).

6. Inputs

The non-uniform input permitted for 3GWAMs include: wind at 10 meters, wind stress, bathymetry, surface currents, air-sea temperature difference, water levels, bottom friction parameter(s), sea ice concentration, sea ice thickness, other sea ice parameters, iceberg parameters, mud parameters, and vegetation parameters. Of course, there is significant variation

Multimedia Report

between models and no single model includes every item in this list. Moreover, there are differences between models with respect to which fields are permitted to be nonstationary. Just as wave models can be created for many different scales (global models, surf zone models, etc.) and different conditions (swell propagation through a tropical archipelago, windsea in an ice-infested sea, waves under a tropical cyclone, swells arriving at a muddy coastline, etc.), the crucial input variables will be different. At global scale and larger regional scales, the accuracy of the wind input is often the primary factor determining wave model skill, even more so than the model physics parameterizations (Cardone et al. 1996, Bidlot et al. 2002, Rogers et al. 2005). This implies that, in those situations, the wave modeler can only influence the second-order errors. Further, the wind problem is not unique. Bathymetry error, coupled with shoreline- and depth-dependent physics, can cause severe difficulties. Some coastlines, especially in the military context, may be inaccessible for traditional bathymetric surveys. At an exposed sandy coastline, the beach profile is seasonal and determined by storm events (CERC 1973). A third example is the situation in and near the ice edge, where even modest errors in the ice edge position can make accurate wave forecasting impossible (e.g. Rogers et al. 2018), and that position can change much even during the short (4 to 24 hour) window between satellite overpasses.

Notably, there is no explicit method in these models for providing them with gridded information about bottom type or ice type. For example, a model of the northern Gulf of Mexico is not explicitly instructed regarding where the southern Louisiana coastline is muddy or the Florida panhandle is sandy. However, extending SWAN or WW3 to allow multiple sediment types in a single grid would be only a modest technical challenge, since corresponding dissipation mechanisms already exist in the models. The larger technical challenge is to provide the model with nonstationary and non-uniform information required by these mechanisms, such the depth of fluidization of the mud layer and the state of a sandy seafloor as it transitions between a flat surface where the roughness is associated with the sand grains, and a rippled surface, where the roughness is associated with the bedforms. In both the mud and sand cases, it is of course a coupled problem, with the wave orbital motion impacting the state of the seafloor (Winterwerp et al. 2012, Smith et al. 2011), but application to deterministic modeling is still relatively academic. Similarly, with respect to ice type: many parameterizations for dissipation by sea ice exist in WW3 (WW3DG 2016; Rogers et al. 2018). Some are purely empirical, while others are solutions to boundary value problems or otherwise based on theoretical calculations. Regardless, parameterizations or settings with parameterizations are sometimes associated with specific ice types, e.g. grease ice, thick frazil ice, pancake ice, sheet ice, or floes from broken sheet ice. The intuitive conclusion is that if we know the ice type, we can more accurately estimate the dissipation of wave energy. Unfortunately, the analogy to seafloor interaction still holds: estimation of sea ice type is still not a reality for operational use. We can imagine that this situation will improve soon, as satellite technology and process models improve. Again, like the seafloor, it is a coupled problem, both with ice breakup by waves (e.g. Williams et al. 2013; Collins et al. 2015; Boutin et al. 2018) and the role of waves in ice formation (Shen et al. 2001, 2004; Doble 2009; Thomson et al. 2017).

Two issues with wind forcing are worth further discussion. One is the modification of the effective wind speed by surface currents, such that the wind speed used for source term calculation is reduced when the wind and currents are aligned, which is of course not

uncommon. However, a wave modeler must be judicious when including this effect. In the real ocean, such a scenario results in lower drag on the atmosphere and thus an increase in wind speed. As such, it may be problematic to apply in cases where the atmospheric model does not also include the currents. The second issue is air-sea stability. Wave modelers often assume stable atmosphere, since it is inconvenient to provide the air-sea temperature differences to the model, the effect on wave growth is usually small, and the available stability corrections are not routinely evaluated for effectiveness. However, as our models steadily improve, small errors, previously unnoticeable, may now be otherwise. Figure 5 compares model waveheight errors against air-sea temperature differences for a hindcast with the NRL global model, run without this stability information. The correlation is especially remarkable if one considers that over much of the ocean, swell dominates, implying that waveheight is not determined by the local winds. The result suggests that a well-crafted stability correction would provide significant benefit.

For wave model grids that are not global and not in enclosed seas, boundary forcing must be provided. This is most often in the form of directional spectra. Ideally, the geographic resolution of the boundary forcing should approximately match the resolution of the model providing the forcing, so that computed spatial variability of the wave field is not lost during the nesting process. The transition from open water to coastal regions can reveal errors that are not as apparent offshore. Error in spectral shape offshore can become error in significant waveheight nearshore. For example, accurate prediction of the blocking of swell by islands demands high accuracy of directional distribution (Rogers et al. 2007). Also, strongly frequency-dependent reactions to bathymetry enhance the impact of spectral error, even for validations that focus primarily on significant waveheight.

7. Validation

Validation establishes the degree to which output from a model can be trusted. It quantifies the skill of the wave model, and (if done carefully) provides attribution or evidence for the cause of error, so that future predictions can be improved. The most conventional wave parameters included in validation are significant waveheight H_s and dominant wave period, for which peak period T_p is sometimes used. However, experienced wave modelers avoid T_p because the model frequency resolution is often coarse relative to the natural variability of peak period during a wave event, resulting in step-wise time series. They instead use a metric for dominant wave period that is calculated via integration of the spectrum (discussed below). Experienced modelers include the 10-meter wind speed U_{10} in even the most basic validation, since atmospheric forcing is always the first suspect when windsea is poorly predicted. Also, it is always a good idea to compute bulk parameters using integration over a common frequency range. For example, if the buoy spectra stop at 0.4 Hz, the model integration should also stop at 0.4 Hz. This requires additional work organizing and managing spectra, but it is most often worthwhile. We summarize our three recommendations for the most elementary validation: 1) include wind speed, 2) use a wave period metric based on spectral integration, and 3) use consistent limits on frequency integration.

The moments of the non-directional spectra are particularly useful in validation: $m_n = \int E(f) f^n df$. In the spectral context, significant waveheight H_s is defined using the zero-moment

waveheight: $H_{m0} = 4\sqrt{m_0}$. The third moment m_3 is proportional to surface Stokes drift in deep water (e.g. Ardhuin et al. 2009), and the fourth moment m_4 is proportional to mean square slope in deep water (e.g. Liu et al. 2000; Li et al. 2013). These higher moments are particularly useful when studying the impact of source terms on the spectral tail (e.g. dissipation by sea ice, Rogers et al. 2016). However, for the same reason, they are sensitive to measurement errors which are often larger in the spectral tail, so model-data misfit must be interpreted carefully. The m_2 moment is related to mean square of particle velocities (orbital motion). The m_{-1} moment, which places more emphasis on lower frequencies, is proportional to energy flux in deep water. Mean periods are defined using ratios of moments, e.g. $T_{m,-1,0}$ is calculated using m_{-1} and m_0 and is a superior alternative to T_p as a metric for dominant period. Additional emphasis can be placed on energy at the spectral peak in the period metric by using the moments on $E^4(f)$ rather than those of $E(f)$ (Collins et al. 2014). The period $T_{m,0,1}$ is perhaps the most commonly used definition of “mean period”. The mean period $T_{m,0,2}$ places more emphasis on the shorter waves and in deep water is close to the “zero up-crossing period” T_z calculated from sea surface time series, just as H_{m0} is close to $H_{1/3}$, being, respectively, the spectral and time series definitions of H_s .

A group of metrics for non-directional spectra is related to frequency width or its inverse (narrowness). The so-called “peakedness” parameters fall into this category. The metrics are often used for evaluation of simulations of idealized wave growth, establishing basic model behaviors (e.g. Hasselmann et al. 1985), and in fact such comparisons are simple enough that they are sometimes performed qualitatively, without defining a metric (e.g. Alves et al. 2002). Validation with width/narrowness parameters using scatter plots and skill scores is quite challenging to interpret, since in one part of a time series may be windsea only, so the width parameter will reflect the width of the windsea, and another part of a time series may be a mixed sea/swell scenario, and the same parameter indicates the separation and relative size of the two systems. This can be addressed by restricting usage to unimodal spectra, e.g. by looking at windsea-dominated cases, Rogers and van Vledder (2013), or by isolating the system of interest. The spectral narrowness is associated with groupiness and probability of freak waves (e.g. Janssen 2003, 2005). Many metrics have been proposed, e.g. Longuet Higgins (1957), Goda (1985), Babanin and Soloviev (1998). An excellent review can be found in Saulnier et al. (2011). We do not recommend the use of metrics which require fitting to an idealized spectrum, e.g. the JONSWAP peakedness parameter γ , since it is indirect and impedes reproducibility.

Bulk parameters described above are calculated from the entire wave spectrum. A highly effective alternative to this is calculation of energy within a handful of frequency bands, e.g. one for low frequencies, two for middle frequencies, and one for highest frequencies. Frequency bands can either be fixed in frequency (f) space, e.g. Rogers et al. (2005), or in normalized frequency (f/f_p) space, e.g. Rogers and Wang (2007).

Directional accuracy is sometimes included in validations. Most often, this comes in one of two forms: 1) qualitative side-by-side comparisons of directional spectra (e.g. WAMDIG 1988) or 2) quantitative evaluation of time series of a “mean direction” or “peak direction” metric (e.g. Moon et al. 2003). When the ground truth comes from a directional buoy, the best approach is to use the “directional moments” computed from the Fourier coefficients describing directional

Multimedia Report

distribution at each wave frequency, e.g. Kuik et al. (1988). It is not advisable to apply a directional estimator (e.g. Maximum Likelihood Estimator) to the buoy data when skill statistics are the ultimate goal; this adds an unnecessary step, and some estimators will even modify the directional moments. Mean wave direction and directional spread are the low order moments. Shape parameters skewness and kurtosis are higher order moments, though quantitative comparisons with these are rare, and with good reason. Thorough quantitative comparison with the lower order moments are already difficult: since they are frequency dependent, the modeler must adopt a strategy for condensing into manageable form. Skill scores are more meaningful if the parameters are not computed by integrating/averaging over all frequencies. One useful approach is to compute the parameters integrating frequencies near the spectral peak (Ardhuin et al. 2003). Another approach is to analyze the skill for several frequency bands (Rogers and Wang 2007). A third approach is to focus on a short time series and present analysis on the full frequency distribution (Romero and Melville 2010).

Any validation effort should be mindful of the agency of the model. For example, if observations indicate a major influence of a physical process on a wave parameter, and the model captures most of this variability in the parameter via its parameterization of the process, with, say, 15% error in the parameter, that may be a success. In another case, with very low spatial/temporal variability of the same parameter, 15% error may be intolerable. In another case, errors in wind forcing (or ice, bathymetry, etc.) may make accuracy of the model's parameterizations irrelevant. In a final example, if a swell prediction model is compared to observations only a short distance from the offshore boundary, then it is primarily a validation of the boundary forcing rather than the parameterizations of the model.

Validation should also be mindful of measurement error. This is especially important when looking at spectra, higher moments of spectra, or directional information. We should also be careful when dealing with new measurement types (e.g. new buoy designs). One approach is to attempt to estimate measurement uncertainty, e.g. Alves et al. (2002) include a quantification of statistical uncertainty due to sampling variability. However, there are other types of measurement error, more difficult to estimate. For example, Thomson et al. (2015) find a major impact of biofouling of a Datawell buoy on the higher moments of the non-directional spectrum. O'Reilly et al. (1996) argue that NDBC (National Data Buoy Center) buoys have problems with directional spreading estimates, relative to Datawell buoys. The strength (or lack thereof) of the noise filter used during processing may affect the spectrum. In the largest wave experiment during the ONR-funded Sea State campaign field experiment, there are major discrepancies between data from two drifting buoy types, deployed in the same area, and this had major impact on interpretation of comparisons against WW3 hindcasts (Rogers et al. 2018). In recent unpublished work, we have found that the response function used in 3-meter NDBC discus buoys is non-optimal in the higher frequencies, which can result in incorrect conclusions about the magnitude (and even the sign!) of model bias in low-pass mean square slope (proportional to m_4).

It is uncommon for validation exercises to address phase error in the time series. However, this is an issue of practical importance. For example, if the forecast swell arrival time is accurate, but is early or late, this will affect operational planning. This is especially true for the case of older swells, since there is a longer duration for phase errors to accumulate. It may be caused by

numerical error, incorrect frequency distribution, or physical processes that are omitted from the modeling. For a good example of this type of model evaluation, the reader is referred to Jiang and Babanin (2016).

Most wave validation studies evaluate the skill of model hindcasts or analyses. However, in an operational context, the forecast skill is obviously of interest. For forecasts that are further in the future (longer horizons), the error is larger, and this is primarily caused larger errors in the wind forcing, and in the case of the Marginal Ice Zone, the ice forcing. Good examples of evaluations of forecast skill are Bidlot and Holt (1999), Bidlot et al. (2002), Tolman et al. (2002), and Bernier et al. (2016).

Finally, a validation exercise require selection of statistical measures (e.g. bias, RMS error, correlation, scatter index) for each parameter. The measures used by Cardone et al. (1996) are a good starting point. In cases where only one to three bulk parameters (e.g. waveheight, wind speed, and dominant period) are included, it is not difficult to include several statistical parameters. In other cases, there is risk of overwhelming the reader with numbers. In recent years, the Taylor diagram (e.g. Tolman et al. (2013); Zieger et al. (2015); Stopa et al. (2016)) has become a popular method of presenting several statistical measures visually, rather than in a table. Correlation is a popular choice as a statistic, but it can be highly misleading, especially when used out of context. For example, a model with severe bias may have perfect correlation. A second example: when standard deviation of a parameter is small (e.g. common in a short time series of dominant wave period), a model may exhibit high skill but with very poor correlation (failure to follow the minor changes in observational time series). Similarly, measures which involve a normalization by the observed values can favor models with negative bias and result in misleading statistics; Mentaschi et al. (2013) recommend an alternative method of normalization.

8. Other challenges

A wave modeler is faced with a number of difficulties. We have already discussed what is arguably the primary challenge: accurate input fields. This difficulty is heightened in cases where the model solution is required near rapid spatial and/or temporal changes of the fields, or if the variable is inherently difficult to determine (e.g. ice type from satellite).

Another challenge is in the spectral tail. Traditionally, e.g. with ECWAM with Janssen (1991) physics and WW3 with ST2 physics, the challenge is obviated by attaching the parametric tail (Section 4) at a relatively low frequency, e.g. 0.30-0.45 Hz. However, recently there has been an increase in efforts to explicitly model a larger frequency range, and with greater attention to accuracy in this portion of the spectrum (ST4 and ST6). An argument can be made that this is not the energy-containing region of a typical ocean spectrum, and for most applications, the benefit of modeling “more of the spectrum” is intangible. However, one can imagine certain benefits. If we compute source terms in the tail more accurately than a parametric extension, this is useful for coupling to other models, e.g. wave-to-ocean momentum flux from breaking. Also, if the prognostic portion of the tail (e.g. 0.4 to 1.0 Hz) is more accurate than a parametric tail in the same range, then it is reasonable to assume that where the parametric tail *is used* (beyond 1.0 Hz in this example), there will also be a benefit to accuracy. This will facilitate comparison to some satellite products, e.g. radiometers that estimate a mean square slope including very high frequencies.

The broad objective just mentioned is more accurate tail level. A more specific objective is the accurate prediction of the point of transition of the non-directional energy spectrum from $E(f) \propto f^{-4}$ to $E(f) \propto f^{-5}$. Authors have proposed dependencies of this transition point on inverse wave age U/C_p (Forristal 1981, Kahma and Calkoen 1992) or a combination of wave age and wind speed (Babanin 2011), where U is wind speed and C_p is the phase velocity of the peak frequency. Progress in prediction of the tail level has been substantial in recent years, as evidenced by the skill of m_4 prediction in Figure 6. Progress in prediction of transition frequency has been slower, but in WW3, new results are promising (Liu et al., subm.)

The efficient coupling between multiple models (atmosphere, ocean, ice, wave, sediment, hydrology, aerosol) presents a number of major challenges which are primarily technical in nature. The first problem is regridding. It is not necessarily a good idea to run the wave model on the same grid that is used by another model. Computational costs are dissimilar between models on a per-grid-cell basis, making it necessary to run the wave model at a coarser resolution than the ocean model. Running the wave model on a decimated version of the ocean model is not necessarily wise either. For example, we have found that WW3 is inefficient when run on the “tripole” grids used by the Navy’s ocean models because these grids (e.g. Barron et al. 2007) have highly non-uniform resolution. In the case of coastal coupling, the motivations for high resolution are similar (resolving bathymetry, coastlines, coastal currents, etc.) but not identical. Thus, different grids are often used, which necessitates rapid re-gridding of fields during runtime. Export (or import) of fields from (or to) WW3 is especially difficult in cases where the multi-grid feature of WW3 is used. At time of writing, this challenge has not been fully addressed. Another challenge is purely technical: building the code to allow all of these models to compile together and run as a single executable, so that the fields can be exchanged through system memory. Lastly (also purely technical), in an operational environment, there must be careful coordination and scheduling of required tasks for each model, e.g. field exchange, data assimilation, pre- and post-processing.

Various forms of observational data have been assimilated into wave models, either operationally or as technical demonstrations: Synthetic Aperture Radar (SAR, Breivik et al. 1998, Abdalla et al. 2006, Aouf et al. 2006); altimetry (Lionello et al. 1992, 1994, Wittmann and Cummings 2004); buoy data (Voorrips et al. 1997, Portilla 2009, Veeramony et al. 2010, Orzech et al. 2013); and coastal radar (Pantaleev et al. 2015). Some use optimal interpolation to modify the model initialization and others use one of several approaches to three- or four-dimensional variational (3D-Var or 4D-Var) assimilation, in which the latter improves the simulation dynamically. One often-cited problem with ocean-scale data assimilation is the tendency of windsea to quickly “forget” the data assimilation, relaxing to a state that is in balance with the local winds (Komen et al. 1994); this has been addressed in some cases through modification of the wind forcing (e.g. Bauer et al. 1996). Quality control is always a challenge, especially when the observational methods are less mature or robust (e.g. SAR and WERA radar, in our experience). Altimetry, on the other hand, is very robust, but is impaired by the fact that it only provides the total energy of the wave spectrum. These observations can be used to adjust the entire spectrum up or down, but in cases of both wind sea and swell components, this is problematic, and one can easily imagine situations where one component is adjusted such that the bias of this component is actually made worse. When skill is quantified using only

waveheight, such assimilation may improve the skill of short term forecasts while yielding mixed outcomes for the longer forecast horizons.

9. Common mistakes

This section is specifically for readers that are new to phase-averaged wave modeling. We review some common mistakes. Of course, several of these mistakes are not specific to this type of modeling.

Familiarize with the manual. Though most users' manuals are over 100 pages long, it is always worthwhile to at least scan the contents before starting, and repeat the process periodically. It is not uncommon for a user to become aware too late of key instructions or advice.

Avoid excessive numbers of grid points. This problem is especially common with regular grids. There is a temptation to model a large region at high resolution. This is usually unwise. The user should instead consider irregular or unstructured grids, or nesting. If a model implementation runs too slowly, the user should first ask if the number of grid points is the primary problem, before looking for non-standard shortcuts in the model physics, spectral grid, or numerics.

Use correct bounds of integration. As noted in Section 7, during validation, the bulk parameters should be calculated using consistent frequency range, model vs. observations. Doing this incorrectly usually has only minor consequence for wave height and dominant period, but it is often severe for parameters computed from higher moments of the spectrum, e.g. $T_{m,02}$ and m_4 . Also, it can have severe consequences for *all* parameters in cases where the climatological dominant period is not much lower than the buoy's maximum frequency, e.g. a large buoy deployed in a lake.

Check for blunders. Of course, the possibilities of human blunder are virtually limitless, but one example is to perform visual spot checks of input fields (bathymetry, ice, wind, etc.) by *outputting* these fields from the wave model and plotting them, thus verifying that they were not read in upside down, missing a factor 100, offset by 180° longitude, etc.

Be suspicious. We have already discussed how one should consider measurement error, especially when dealing with higher spectral moments or new measurement platforms. Similarly, notwithstanding the review process, the published literature can contain errors in methods, reasoning, and interpretation. A good example is the interpretation of idealized model results. Such experiments are useful, but one should avoid naïve extrapolation from these results to realistic applications. A thorough study is supported by results from both types of simulations (idealized and realistic).

Multimedia Report

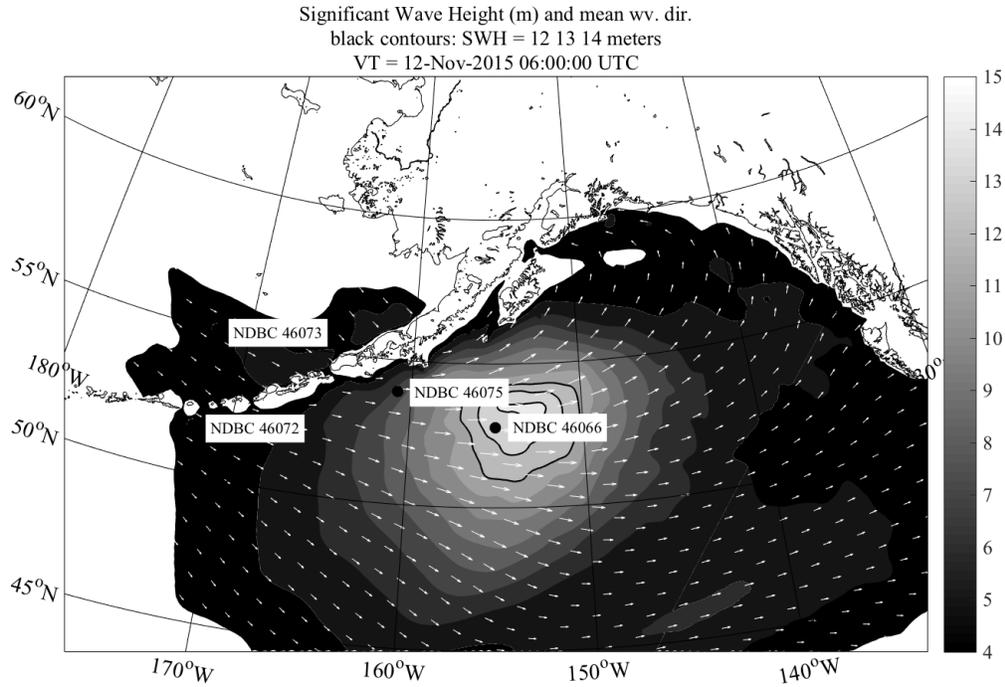


Figure 1. Significant waveheight (colors, in meters) and mean wave direction (arrows) from a WAVEWATCH III hindcast, south of Alaska. The three contours indicate waveheights of 12, 13, and 14 meters.

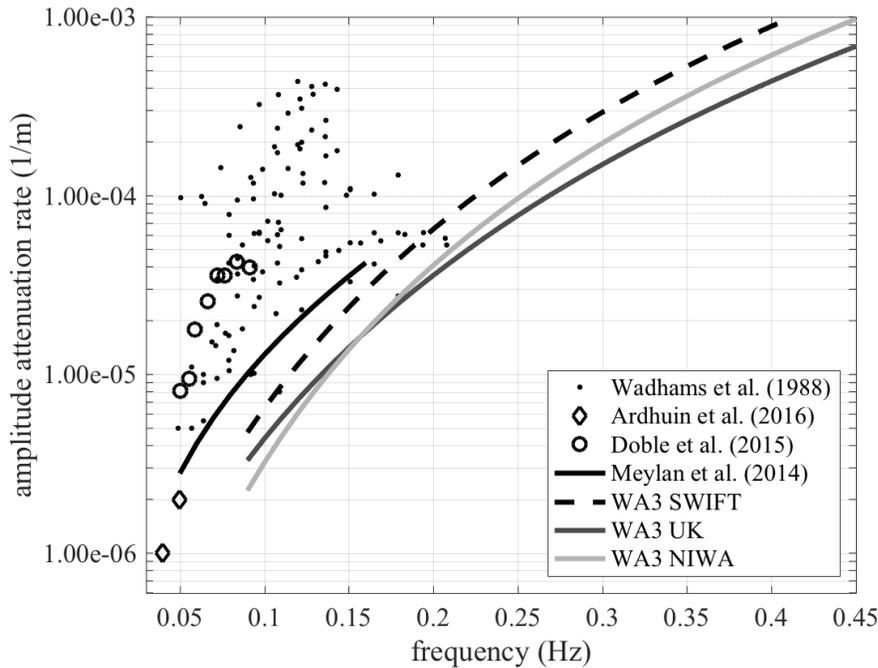


Figure 2. Comparison of the frequency distributions of dissipation of wave energy by sea ice. Wadhams et al. (1988), Arduin et al. (2016), Doble et al. (2015), and Meylan et al. (2014) are observational studies. The “WA3” lines are based on model-data inversion using method similar to that used by Rogers et al. (2016, 2018), from Wave Array 3 (Thomson et al. 2015) of the ONR-supported “Sea State” field experiment. “SWIFT”, “UK”, and “NIWA” are buoy types. Acknowledgments: SWIFT buoy data are from Jim Thomson and Madison Smith (UW/APL). UK buoy data are from Martin Doble (Polar Scientific) and Peter Wadhams (Cambridge U.). NIWA buoy data are from Alison Kohout (NIWA).

Multimedia Report

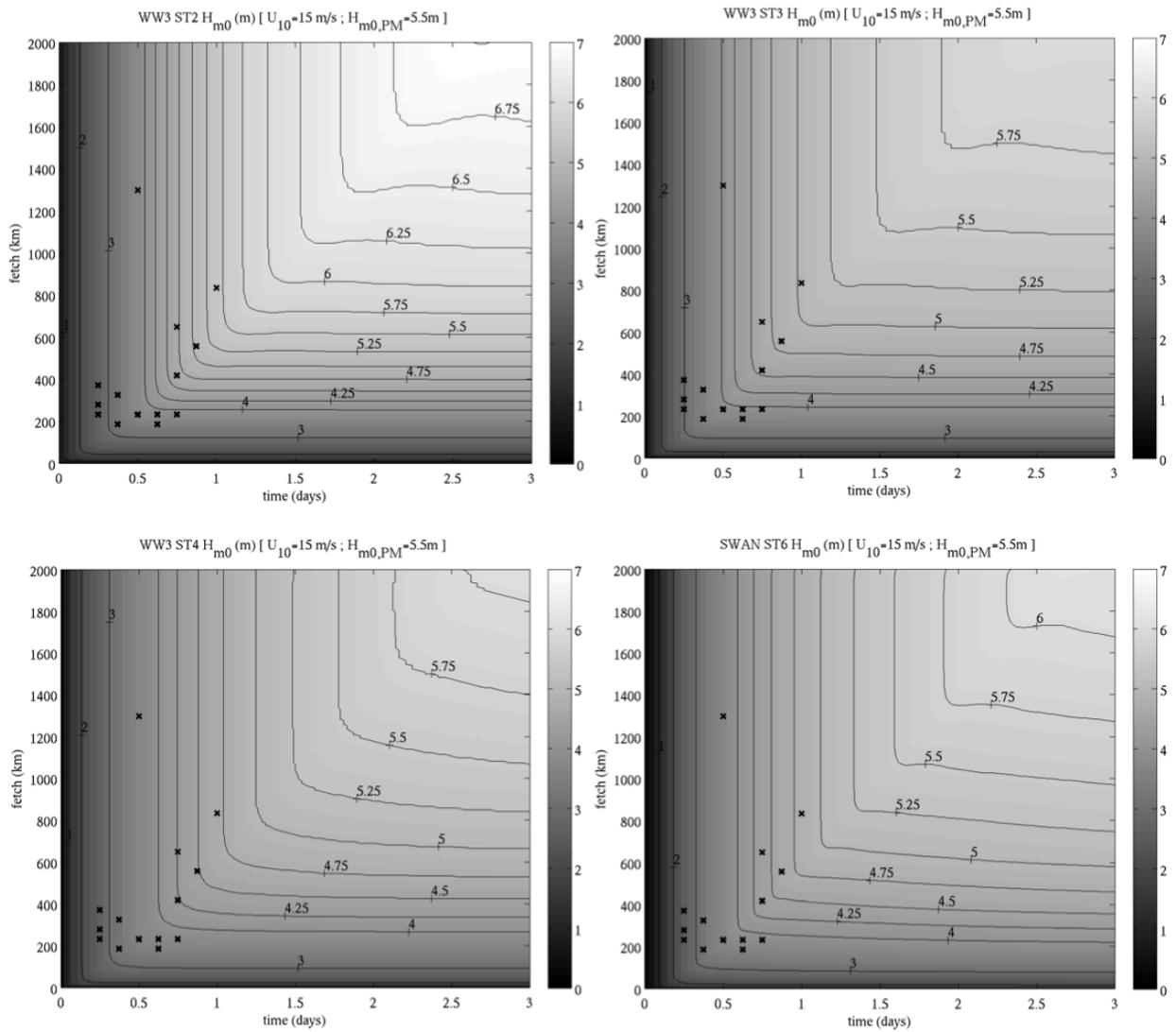


Figure 3. Fetch/duration comparisons of four model. Model source term package is indicated in title above each plot. See text for explanation.

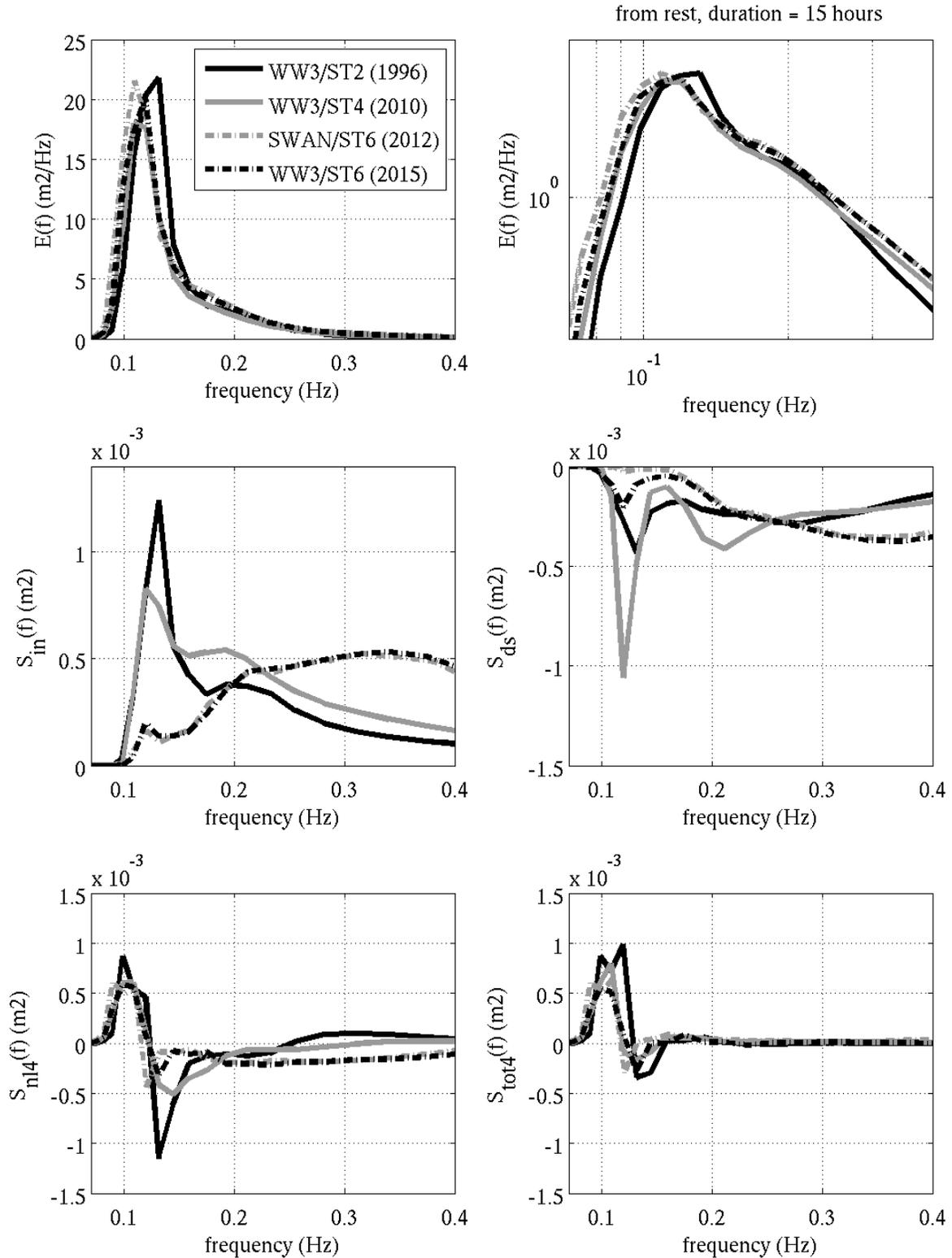


Figure 4. Results from four models, for a duration-limited case, with 10-meter wind speed of 15 m/s after 15 hours duration, starting from rest. $E(f)$ is energy density. $S_{in}(f)$, $S_{ds}(f)$, $S_{nl4}(f)$ are source terms for wind input, whitecapping, and four-wave nonlinear interactions, respectively. $S_{tot}(f)$ is the summation of the three source terms, and is the growth/decay rate, since advection terms are zero (unlimited fetch).

Multimedia Report

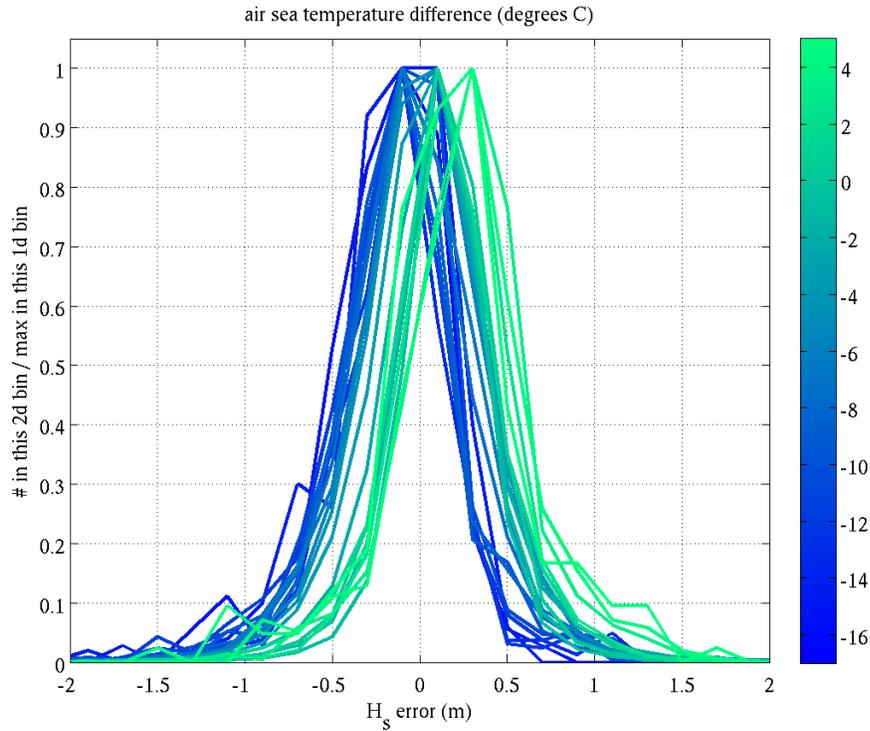


Figure 5. Histograms of waveheight error from a WAVEWATCH III hindcast. Observed waveheight is from buoys. Color scaling indicates air-sea temperature differences ($^{\circ}\text{C}$) taken from buoy observations. Acknowledgments: The WW3 hindcast was performed by Yalin Fan. Buoy co-locations were performed by Yalin Fan (NRL) and Silvia Gremes (U. New Orleans).

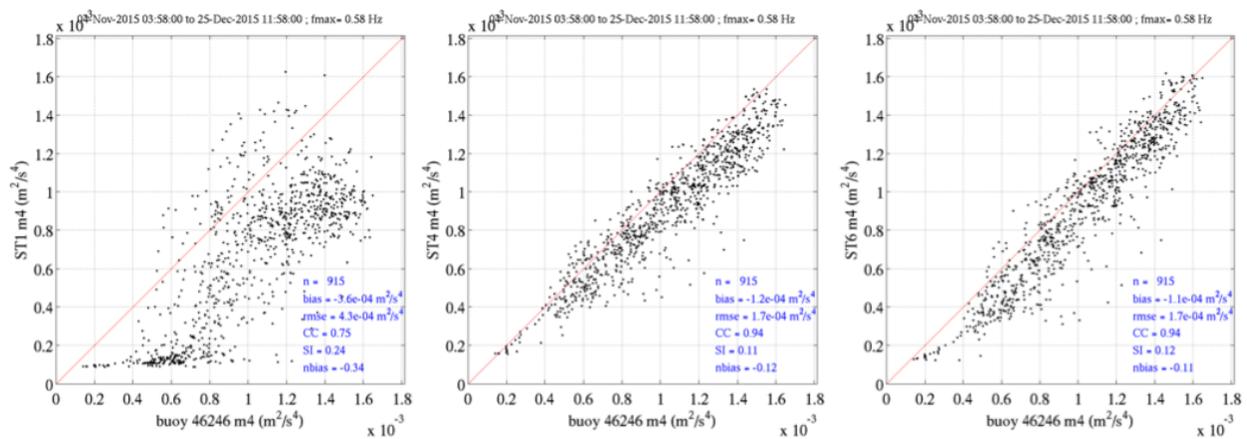


Figure 6. Validation of m_4 for three models: ST1/SWAN, ST4/WW3, and ST6/SWAN, using Waverider buoy CDIP 166, at Ocean Station Papa, owned and operated by APL/UW. Time period is 0400 UTC 4 Nov. to 1200 UTC 25 Dec. 2015. The m_4 parameter is proportional to the contribution to the mean square slope of the sea surface by frequency components up to 0.58 Hz. Additional information about this hindcast and comparison can be found in Rogers (2017).