Naval Research Laboratory

Stennis Space Center, MS 39529-5004

NRL/MR/7320--18-9835

The IRI Grid System for Use with WAVEWATCH III

W. Erick Rogers
Ocean Dynamics and Prediction Branch
Oceanography Division

Robert S. Linzell
Perspecta, Inc.
Stennis Space Center, MS

December 21, 2018

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
21-12-2018	Memorandum Report	

4. TITLE AND SUBTITLE

The IRI Grid System for Use with WAVEWATCH III

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
0602435N

6. AUTHOR(S)

W. Erick Rogers and Robert S. Linzell*

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
73-6A50-08-5

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Research Laboratory
Oceanography Division
Stennis Space Centter, MS 39529-5004

8. PERFORMING ORGANIZATION REPORT NUMBER

NRL/MR/7320--18-9835

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
One Liberty Center
875 North Randolph Street, Suite 1425
Arlington, VA 22203-1995

10. SPONSOR / MONITOR'S ACRONYM(S)

ONR

11. SPONSOR / MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A: Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

*Perspecta, Inc., Stennis Space Center, MS

14. ABSTRACT

This report describes a grid system created for use with the wave model WAVEWATCH III®. The system uses three grids: irregular grids at high latitudes and a regular grid at low latitudes, and so it is referred to as "Irregular-Regular-Irregular" (IRI). There is two-way exchange of information (wave spectra boundary forcing) between grids. The grid system achieves efficiency by minimizing the variability of grid spacing, which allows a larger time step size. The northern grid extends to very high latitude (89˚N), and propagation is computed at third order accuracy. IRI systems have been created for four different resolutions: approximately 1/2˚, 1/4˚, 1/6˚, and 1/8˚. The report includes a review of advantages and disadvantages of different grid approaches; a description of important design considerations; a description of the four designs; timings; validation samples; and a discussion of the operational context (existing and future systems).

15. SUBJECT TERMS

WAVEWATCH III | Wave Model | Polar Stereographic
Grid System | Irregular Grids

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			W. Erick Rogers
Unclassified Unlimited	Unclassified Unlimited	Unclassified Unlimited	Unclassified Unlimited	54	19b. TELEPHONE NUMBER *(include area code)* (228) 688-4727

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

This page intentionally left blank.

CONTENTS

Listing of tables and figures

Executive Summary

This report describes a grid system created for use with the wave model WAVEWATCH III®. The system uses three grids: irregular grids at high latitudes and a regular grid at low latitudes, and so it is referred to as "Irregular-Regular-Irregular" (IRI). There is two-way exchange of information (wave spectra boundary forcing) between grids. The grid system achieves efficiency by minimizing the variability of grid spacing, which allows a larger time step size. The northern grid extends to very high latitude (89°N). IRI systems have been created for four different resolutions: approximately 1/2°, 1/4° 1/6°, and 1/8°. The report includes a review of advantages and disadvantages of different grid approaches; a description of important design considerations; a description of the four designs; timings; validation samples; and a discussion of the operational context (existing and future systems).

This page intentionally left blank.

1. Introduction

The U.S. Navy operates numerical models which provide forecasts of wind-generated surface gravity waves. Implementations of these models exist at a number of scales. The present report is concerned with global- and basin-scale wave models. These forecasts assist in planning of Navy operations, for example, to determine if and when particular ship operations can be conducted safely based on predicted sea state, and whether to adjust a route to enable these operations. These large-scale wave models also are used to provide boundary forcing to coastal wave models, which in turn are important for planning of nearshore activities, such as the operation of small boats or movement through the surf zone. Prediction of swell using a numerical model is particularly useful: by definition, swell is decoupled from the local winds, so a meteorological forecast alone is insufficient to anticipate swell conditions. In the context of windsea, the numerical model uses state-of-the-art physics parameterizations to quantify the impact of the available fetch (e.g., where it is defined by the coastline, or sea ice), and the impact of spatial and temporal variability of the wind field on the wave conditions. The U.S. Navy uses the WAVEWATCH III® (WW3, Tolman et al. 2002) for global wave modeling (see review in Rogers et al. 2014). This phase-averaged model solves for the wave field in terms of directional spectra. Forecast products tend to be in the form of bulk parameters which are calculated from the spectra. Examples are significant waveheight, peak period, mean period, mean direction, mean square slope, and surface Stokes drift.

This report is concerned with two priorities for operational wave modeling. The first priority is efficiency. This requires little explanation: for a particular resolution, a more efficient global model finishes more quickly and is available sooner. Or, if the time available is fixed (e.g. 30 minutes runtime), higher efficiency permits higher resolution. Of course, the forecast range and computing resources also figure into this equation.

The second priority is to include high latitudes. When a single regular grid is used for a global wave model, the meridians converge toward the north pole, which means that if the grid is extended north, the minimum geographic spacing is reduced, and the geographic propagation time step must be reduced for numerical stability. Thus, to preserve efficiency (the first priority), the maximum latitude is often restricted in practice (e.g. 78°N). Unfortunately, there is—depending on season—significant open water north of 78°N, which is therefore omitted from computations. So, for these regions, there is no forecast, and no information available for providing boundary forcing to regional wave models. Also, insofar as wave energy can propagate southward, it would also affect accuracy south of the northern limit. The second priority is to address this problem.

The system presented here uses three grids: irregular grids at high latitudes and a regular grid at low latitudes, and so it is referred to as “Irregular-Regular-Irregular” (IRI). There is two-way exchange of information (wave spectra boundary forcing) between grids. The grid achieves efficiency by minimizing the variability of grid spacing, which allows a larger time step size. The northern grid extends to very high latitude (89°N). IRI systems have been created for four different resolutions: approximately 1/2°, 1/4° 1/6°, and 1/8°.

Manuscript approved December 21, 2018.

In Section 2, the grid designs are described. Section 3 presents timings from the IRI grids and other global grid designs. Section 4 describes verification in the form of propagation testing. Section 5 presents examples of validation. Section 6 summarizes the development timeline. Section 7 describes the role of the grid system in context of existing and planned operational systems. Section 8 is discussion and Section 9 lists the main findings. A glossary of terms is given in Section 10.

2. Description of "IRI" global grid designs

2.1. Background: description of the multi-grid concept

WAVEWATCH is described in Tolman (1991). WAVEWATCH III® (WW3) is described in a number of publications, e.g., Tolman et al. (2002), but we recommend WW3DG (2016) as a primary source for the model, since it is more recent, and the model has undergone many changes since 2002. The multi-grid approach of WW3 was developed by Tolman (2008). This approach allows the computation of multiple grids using a single model run. There is two-way communication between grids: wave spectra are exchanged in both directions at the model boundaries.

The advantages of the multi-grid approach are described in the following paragraph. However, it should be clearly understood that these advantages apply to the case where the multi-grid model is used for grid nesting and *do not apply to the IRI grid design*. In fact, the only major advantage of multi-grid that applies to IRI design is the most obvious one: that it combines grids of two types (regular and polar stereographic) to take advantage of the relative strengths of both grid types, using one or the other depending on latitude.

In the multi-grid approach to nested grids, it is possible to omit computations within the low rank (usually taken to mean lower resolution) grids at regions where the ocean is computed within higher rank (usually higher resolution) grids. This reduces geographically redundant computations associated with the traditional one-way nesting, whereby a low resolution grid simulation is run (e.g., a global model) and then higher resolution grids are computed in subsequent simulations. The multi-grid approach to nested grids also makes the system design easier (though not necessarily easy) by automating selection of boundary data output points. Also, the two-way approach allows the low resolution model (and therefore all child grids) to benefit from the high resolution computations. For example, tropical cyclones tend to be small and poorly resolved in a global grid, but in a multi-grid system, generation of waves by a hurricane can be modeled at high resolution in a regional grid, and when this wave energy later becomes non-local (swell), prediction of this swell will benefit. By contrast, in the traditional one-way nesting, the swell from a remote hurricane always originates from the low-resolution global model. Similarly, the multi-grid approach allows the entire global wave modeling system to benefit from the improved accuracy of high resolution atmospheric models. (This topic is discussed further in Section 8.2.) In the traditional one-way nesting approach, only the high resolution wave models benefit from the high resolution atmospheric models, and the benefit is strictly local. Lastly, with the multi-grid WW3, it is possible to define coastal grids via intelligent masking, such as the Alaska grid demonstrated in Tolman (2008).

The multi-grid approach is not without disadvantages though. These include:

1. (General) The code is more complex, which makes model development and debugging more challenging, especially with respect to parallelism (MPI broadcasts of data between grids) and coupling to other models (e.g., exporting fields from multiple wave model grids to a coupler for use in an ocean model which is invariably on a different grid).
2. (General) Start-up time is increased, since the multigrid model performs extra steps during initialization, such as computing data arrays which instruct the later computations how to handle boundary conditions. This penalty is more relevant to short simulations, e.g., a 1-day forecast, and is negligible for longer simulations, e.g., a 30-day hindcast.
3. (Relative to one-way nesting.) In an operational context, the delivery time of the coarse parent grid is delayed, since computations of all grids complete at the same time. In the traditional one-way nesting, output from the different grids are available in sequence: global grids, then regional grids, then sub-regional grids, etc.

Gridded output from a multi-grid run (e.g., spatial distribution of significant waveheight) is often displayed and/or analyzed on the individual grids. However, the multi-grid approach is assisted by an optional post-processing program (ww3_gint), which remaps output variables from multiple grids onto a single grid. For example, for a system composed of a combination of a 0.5° global grid and five 0.1° resolution regional grids, the gridded output can be remapped to a single 0.1° resolution global grid which is *not* used for computations.

2.2. "IRI" approach (brief description)

Our approach is to use an irregular-regular-irregular (IRI) multi-grid system: three grids: a grid with regular spacing (in terms of latitude and longitude) at low latitudes and polar stereographic (PS) grids at high latitudes. The advantage of the IRI system is that each grid has more uniform grid spacing (in terms of real distances) than alternative global systems, which means that the limiting time step for CFL stability is larger, which in turn makes the system more efficient.

2.3. Comparison to alternative strategies

Here, we compare the IRI method with three possible alternatives: 1) a single high resolution global grid with regular latitude/longitude spacing, 2) a single high resolution global grid with irregular latitude/longitude spacing (tripole grid), and 3) a multi-grid system that uses a mix of global and regional grids. A fourth approach is possible—and noteworthy—but not included in discussion here: the "SMC" (Spherical Multiple-Cell Grid) unstructured grid, global wave modeling approach of Li (2009, 2011, 2012) which was introduced to WW3 in Li and Saulter (2014) and Tolman et al. (2014).

2.3.1. Versus single regular global grid

The IRI system can be compared to a single high resolution global grid with regular latitude/longitude spacing. An example of such a system is the operational global model(s) used by FNMOC prior to 2018:

- (Con) IRI design has some redundant computations in limited regions where grids overlap (Section 2.6.3)
- (Con) IRI has increased initialization time (important for very short simulations).
- (Con) In context of coupling: import/export of fields between multiple WW3 grids and the coupler requires additional code.

- (Pro) IRI can include high latitudes efficiently. With regular grid, the higher the maximum latitude, the less efficient the model becomes.
- (Pro) In context of coupling, export of fields from a single global WW3 would not include high latitudes.

2.3.2. Versus single irregular global grid (tripole)

The IRI system can also be compared with a global system which has irregular latitude/longitude spacing. The most relevant example is the grid which we denote as the "tripole grid" herein. It is relevant because grids of this sort are used operationally by the Navy for ocean modeling (e.g., Barron et al. 2006, 2007) and ocean/ice modeling (e.g., Metzger et al. 2014a). It has three poles, which are all over land to reduce distortion of sea points: two poles in the northern hemisphere and one at the South Pole. The grid is created using the method of Murray (1996) by Dr. Alan Wallcraft (employee of NRL 7320 at the time). It is designated as the "GLBa0.24" grid and exactly overlays every third point of the HYCOM "GLBb0.08" grid. The name implies a grid resolution near 1/4°. However, that is the resolution near the equator and the resolution is much higher elsewhere (Figure 1).

The tripole grid is a logically rectangular irregular grid. Irregular grid capability was implemented in WW3 in 2009 (Rogers and Campbell 2009). However, the tripole grid requires additional coding, because it includes an "Arctic seam" where two sides of the grid are stitched together. We coded WW3 to permit calculation of spatial gradients and propagation across this seam in May 2014 (WW3DG 2016). This was implemented only for the first order propagation scheme "PR1". This type of grid has never been used operationally for wave modeling, but we have used it for hindcasts (Section 3). The tripole grid has highly non-homogeneous spacing (Figure 1 and Figure 4), which is disadvantageous in terms of propagation time step.

Here we compare the IRI design to the tripole grid approach:

- (Con) IRI design has some redundant computations in limited regions where grids overlap (Section 2.6.3)
- (Con) IRI has increased initialization time (important for very short simulations)
- (Con) In context of coupling: import/export of fields between multiple WW3 grids and the coupler requires additional code
- (Pro) IRI permits larger times steps for propagation, making it more efficient than the tripole grid.
- (Pro) At time of writing, in WW3 with tripole grids, only the first order propagation scheme (PR1) is permitted. With the IRI system, any propagation scheme may be selected (usually the third order scheme, PR3).

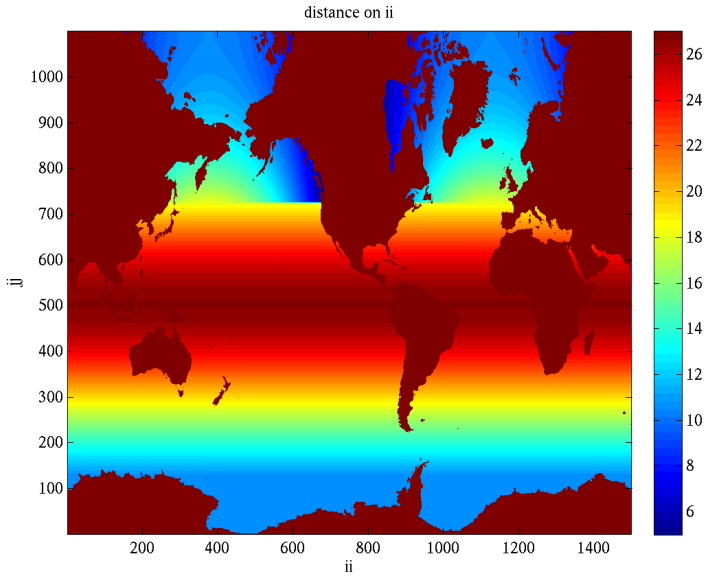

Figure 1. Spacing, in km, for the tripole grid. It is plotted in index space, and the spacing is computed on the "i" dimension (the horizontal coordinate here).

2.3.3. Versus multi-grid with mix of global and regional grids

Here we compare the IRI systems to the strategy using a global multi-grid (two-way nesting) that includes regional grids. Two examples of such systems would be 1) the one used by NCEP for many years (following the work of Tolman 2008) and 2) the multi-grid system transitioned by NRL to NAVOCEANO (Rogers et al. 2012, 2014).

- The IRI design uses one-way nesting to feed regional grids, which leads to the following advantages and disadvantages:
    - (Con) There is an increase in redundant computations with IRI one-way nesting to regional grids (multiple grids solving for the same geography).
    - (Con) There is an increase in labor to specify boundary output points for regional grids.
    - (Con) The wave energy generated using high resolution grids and/or high resolution atmospheric forcing does not improve the global grid solution.
    - (Con) This would likely prevent use of coastal (masked) grids such as the Alaska grid described in Tolman (2008). These are much less convenient with one-way nesting. To our knowledge, this is never done in practice. Specification of output points would be labor intensive.
    - (Pro) High resolution atmospheric fields are available later than regional atmospheric fields. Thus, compared to a two-way nesting system that exploits high resolution atmospheric forcing, the IRI system will be able to start sooner.
    - (Pro) Global output will be available sooner with IRI system. A multi-grid system that includes regional grids will have delayed completion time for the global wave model component.
- (Pro) Most of the IRI systems have global resolution that is finer than the coarsest grid (0.5°) used by Rogers et al. (2012). This implies an improvement in resolution for most of the ocean.

In some cases, the comparisons become conditional. The 10-grid system of Rogers et al. (2012)

prioritizes some parts of the ocean over others, e.g., to exploit COAMPS (Hodur 1997) forcing, or to better resolve coastlines. The IRI systems do not internally prioritize any part of the ocean: coastal regions are modeled subsequently using one-way nesting. Some of the IRI systems have global resolution that is similar to or finer than the regional grids used by Rogers et al. (2012). This implies that the regional grid does not need to be included via one-way nesting. However, there is obviously a computational cost associated with these higher-resolution IRI systems.

2.4. Design features and considerations

There are two basic questions to consider during the design process:

1. What grids to match up, e.g., do we match a 1/6° with a 14 km PS grid or 24 km PS grid?
2. At what latitude should they intersect?

This is governed by one primary concern: Grid spacing (in real distances) should be as uniform as possible across all latitudes. This results in a less strict CFL criterion. As an example, based on these considerations, we paired the 1/6° regular grid with 21 km PS grids, and joined them at 30N and 30S (Figure 3).

There was also a secondary concern, related to memory usage. During the design of IRI-1/6 and IRI-1/8, we were concerned about the WW3 job failing by running out of memory, because of the large numbers of sea points. This was a common mode of failure with WW3 on the IBM systems on the DSRC. With this in mind we designed the grids so that there is better balance between grids in terms of memory usage (via number of sea points) than was the case with IRI-1/2 and IRI-1/4, speculating that it may reduce the peak memory usage. However, the transition to the Cray systems on the DSRC occurred in a similar timeframe (Section 6), and with those systems, the problems with running out of memory became rare. Thus, this secondary concern evaporated. No tests were performed to determine the impact (if any) that this “grid balancing” has on peak memory usage.

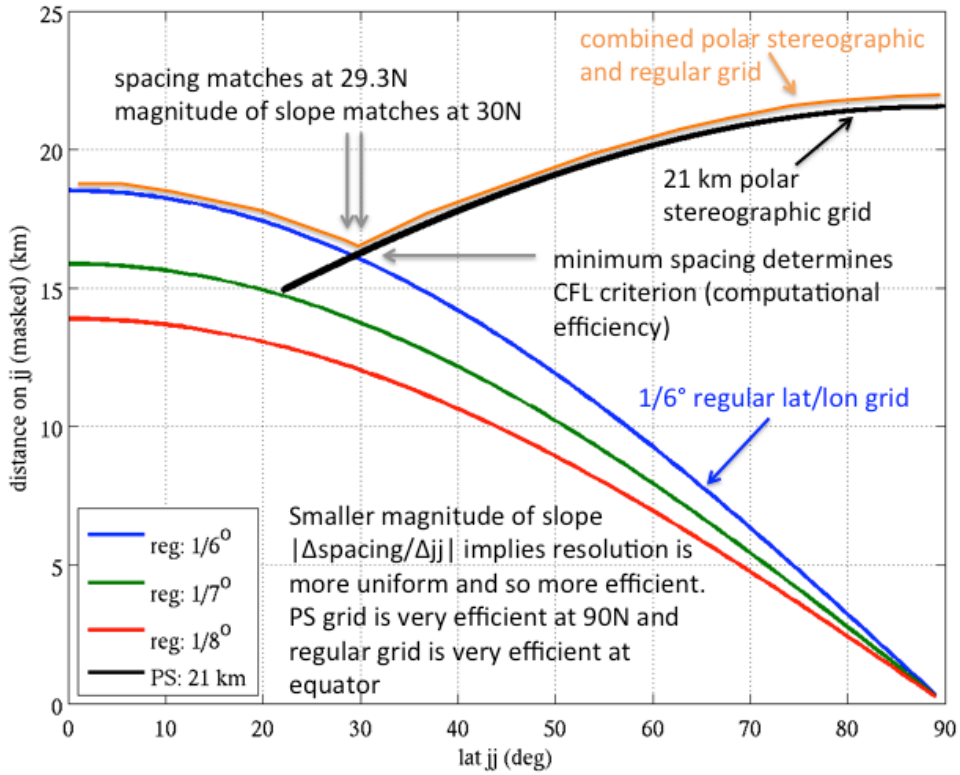

Figure 2. Spacing (in km) as a function of latitude (in degrees). This shows two of the grids in the 1/6° 3-grid IRI system: the low-latitude grid (blue) and the high latitude PS grid (black). The orange line indicates the IRI system as a whole. The grids exchange information at 30N. The high-latitude southern grid is not shown, but is used for 90S to 30S.

2.5. Design summaries (tabulated)

Table 1 shows the grid specifications for the polar stereographic (PS) grids in each of the four IRI designs. This table includes: 1) the resolution in km at 70° latitude (north or south), 2) the number of nodes (n_i and n_j) for each grid (north or south), and 3) number of sea points n_{sea} in each grid. The larger n_{sea} in parentheses "()" is the value prior to application of the "auto-mask" (Section 2.6.2). The smaller number, not in parentheses, is a better representation of the grid size.

Table 1. Grid specifications for irregular (polar stereographic) grids in IRI designs

design	resolution@70°	n_i	n_j	n_{sea}(north)	n_{sea}(south)
IRI-1/2	34 km	260	262	24k	42k
IRI-1/4	18 km	520	522	90k	159k
IRI-1/6	21 km	770	772	208k (235k)	364k (391k)
IRI-1/8	15 km	1060	1062	372k (450k)	638k (741k)

Table 2 shows the grid specifications for the low-latitude regular grids. The name of the design corresponds exactly to the grid spacing in degrees, e.g., IRI-1/4 has spacing of 0.25°.This table includes: 1) the latitude spacing in km, 2) the number of longitudes and latitudes (n_i and n_j), 3) number of sea points n_{sea}, and 4) the latitude of the overlap between irregular and regular grids (Section 2.6.3). Again, the larger n_{sea} in parentheses "()" is the value prior to application of the

"auto-mask" (Section 2.6.2). Note that n_i and n_j correspond to the number of longitudes and latitudes (n_i=n_x and n_j=n_y) only in the case of regular grids.

Table 2. Grid specifications for regular (low-latitude) grids in IRI designs

design	resolution (N/S)	n_i	n_j	nsea	IRI overlap (°)
IRI-1/2	55 km	720	340	116k (121k)	50 to 55 (both)
IRI-1/4	28 km	1440	441	455k (461k)	50 to 54 (both)
IRI-1/6	18.5 km	2160	362	577k	24.6 to 30.5 (north)
IRI-1/8	14 km	2880	497	1057k	25.1 to 30.9 (both)

Table 3 shows information on

1. whether the irregular or regular grids are treated as being higher rank (Section 2.6.2),
2. time step sizes and ratios between global and propagation time step sizes (Section 2.6.8), and
3. indication of whether a) this report includes timings that can be directly compared to timings on other grids, or if b) the timings in this report are not directly comparable due to different number of processes or different computer architecture (Section 3).

Table 3. Information on rank, time step, and timings for each of the IRI designs

design	higher rank	global Δt_g	prop Δt_p (all grids)	ratio $\Delta t_g/\Delta t_p$	Timings comparison
IRI-1/2	irregular grids	3600 s	1200 s	3	indirect
IRI-1/4	irregular grids	3600 s	720 s	5	direct
IRI-1/6	regular grid	3600 s	600 s	6	indirect
IRI-1/8	regular grid	3600 s	450 s	8	indirect

Table 4 shows information about the global models which are run using a single grid rather than using IRI design. Information includes:

1. Maximum latitude in the grid,
2. Number of nodes, n_i and n_j,
3. Number of sea points, n_{sea},
4. Time step sizes and ratios between global and propagation time step sizes (Section 2.6.8), and
5. Indication of whether a) this report includes timings that can be directly compared to timings on other grids, or if b) the timings in this report are not directly comparable due to different number of processes or different computer architecture (Section 3).

Table 4. Information about the single-grid global models

design	max lat. (°)	n_i	n_j	n_{sea}	global Δt_g (s)	prop. Δt_p (s)	ratio $\Delta t_g/\Delta t_p$	Timings comparison
1/2° global	77	720	340	153k	1800	450	4	indirect
1/4° global	78.5	1440	680	621k	3600	240	15	direct
"	"	"	"	"	1200	240	5	direct
"	84	"	"	648k	3600	120	30	direct
"	"	"	"	"	960	"	8	direct
"	"	"	"	"	720	"	6	direct
"	"	"	"	"	600	"	5	direct
tripole (PR1)	n/a	1500	1099	991k	3600	150	24	direct
"	"	"	"	"	750	"	5	direct

See Section 3.1 for tables with grid information and *timings* for cases with direct timing comparison.

2.6. Design discussion

2.6.1. Variability of resolution

The resolution (in km) of the IRI grids and the tripole grid is shown in Figure 3. The smallest resolution for each design determines the propagation time step. This figure is useful for visualizing the difference between maximum and minimum resolution for each system, $\Delta_{max} - \Delta_{min}$. However, the quantity $\Delta_{max} - \Delta_{min}$ is not the best indicator of efficiency. It is more appropriate to visualize $\Delta_{max}/\Delta_{min}$. This is presented in Figure 4. Here, we see that for all the IRI systems, the time step required by the smallest grid cell is never less than half the time step required by the largest grid cell, so the model is running near maximum efficiency at all grid cells. In other words, the Courant number of the fastest waves for all grid cells is between 0.5 and 1.0. In contrast, the time step required by the largest grid cell in the tripole grid is more than five times the time step required by the smallest grid cell, so the Courant number of the fastest waves is less than 0.2 for some grid cells, which implies extreme inefficiency. The figure also indicates that the IRI-1/6 and IRI-1/8 systems are somewhat more efficient than the IRI-1/2 and IRI-1/4 systems, since the $\Delta_{max}/\Delta_{min}$ is higher for the low latitude regular grids of IRI-1/2 and IRI-1/4 systems.

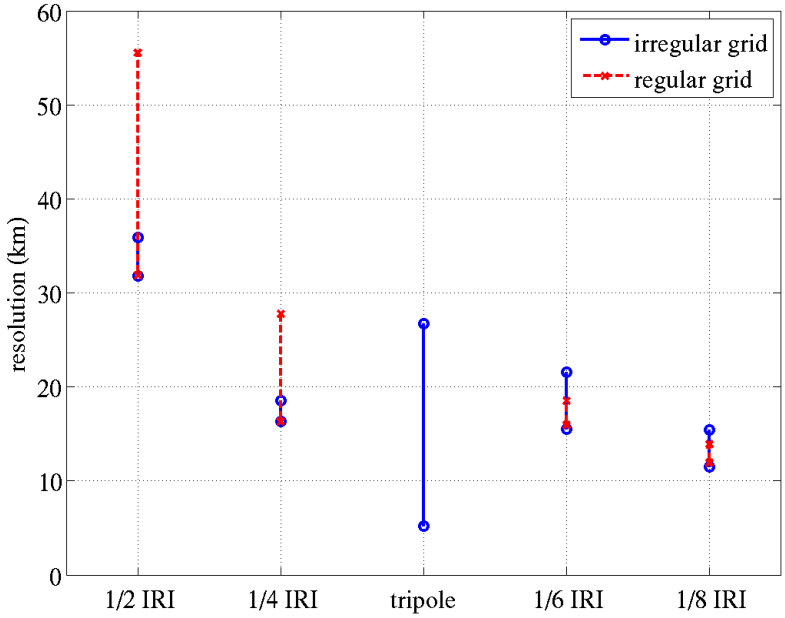

Figure 3. The resolution (in km) of the IRI grids and the tripole grid

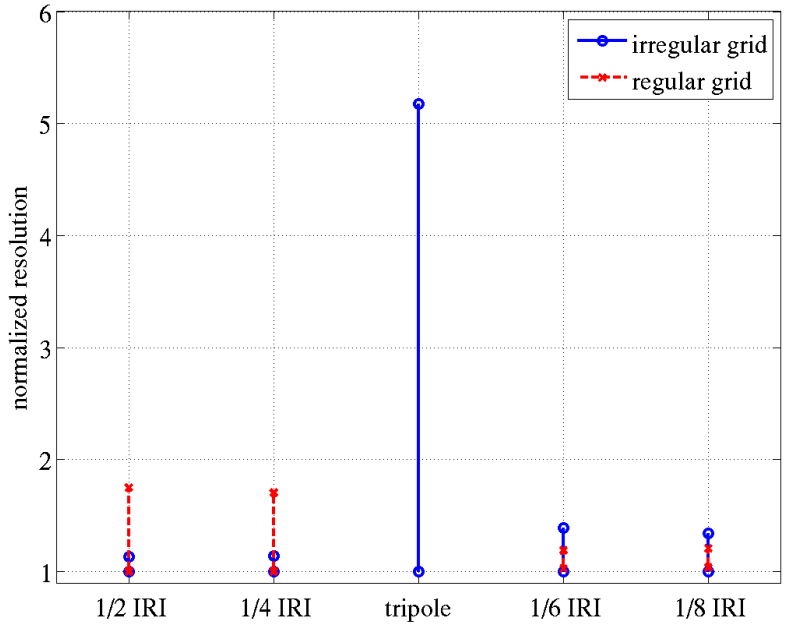

Figure 4. Like Figure 3, but resolution is normalized by the smallest resolution in each system.

2.6.2. Internal boundaries in IRI systems

Grid rank and conservative remapping

In multi-grid WW3, the user must assign a rank to each grid. This rank determines how information (wave spectra) is passed between grids at the boundaries. In principle, higher rank should be assigned to grids of higher resolution. For wave energy that is propagating from a low-rank grid to a high-rank grid, the situation parallels that of traditional one-way nesting, and as with traditional one-way nesting, simple bilinear interpolation is used to remap the spectra from

the source grid points (low resolution) to the destination grid points (high resolution). For wave energy that is propagating from a high-rank grid to a low-rank grid, the situation is unique to the multi-grid approach. Bilinear interpolation is not appropriate. To understand why, consider the extreme case where the high rank grid is many times finer than the low rank grid. With bilinear interpolation, the remapping would be nearly equivalent to simple "nearest neighbor" selection, and if there is strong variability in the high rank grid, much information can be lost, and artifacts such as aliasing can be expected. So, instead, a conservative remapping approach is used, where the high resolution source grid cells are averaged to compute a spectrum for the low resolution destination grids cells according to the region of overlap. This was implemented for regular grids by Tolman (2008). NRL implemented a system for conservative remapping with irregular grids (WW3DG 2016) using the "SCRIP" software of Jones (1999), which determines the region of overlap and the averaging "weights". This software is compiled into WW3. The computation of these weights is expensive for large grids because it is not parallelized and does not take masks into account. This can make model initialization very slow. Fortunately, it is possible in WW3 to store weights for subsequent model runs (WW3DG 2016), so the problem is effectively addressed. Since the use of SCRIP does not take masks into account, the same stored weights can be used even if masks are changed. It is only necessary to recompute weights if the latitude-longitude grids are changed.

For a multi-grid design using all regular grids, the user may assign two overlapping grids equal rank, and the two grids are reconciled as described in Tolman (2008). However, we have not extended the code of WW3 to permit the case where one or both of the overlapping grids are irregular grids. Thus, in the IRI systems, we must decide whether the PS grids or low-latitude grid are of higher rank. This cannot easily be decided based on resolution, since the strategy of the IRI system is to make resolution as uniform as possible, globally. However, small differences in resolution do exist (Figure 3 and Figure 4), and this is used to decide grid rank (Table 3):

- IRI-1/2 and IRI-1/4: PS grids are higher rank.
- IRI-1/6 and IRI-1/8: low-latitude grid is higher rank.

In the IRI design, since it uses overlapping grids of non-equal rank, there is (less precise) bilinear interpolation for wave energy moving across boundaries in one direction, and (more precise) conservative remapping for wave energy moving across boundaries in the other direction. It is reasonable to ask whether the system would be more accurate if WW3 were coded to allow overlapping irregular grids of equal rank. The short answer is: probably not, since the equal rank grid procedure uses interpolations in both directions. However, for large problems such as IRI-1/8, there may be benefits in terms of runtime, due to positive impact of equal rank grid procedure on scaling. This is discussed in Section 8.3.2.

Auto-masking

With WW3, the user runs a grid pre-processor program named "ww3_grid", which reads user-provided information about each grid. This includes latitude, longitude, bathymetry, sub-grid obstructions, grid masks, and model options such as calibration coefficients for physics parameterizations. The program produces a model definition ("mod-def") binary file which is used by subsequent WW3 programs. In an operational context, the ww3_grid program is run

infrequently, only when the information in the mod-def file needs to be changed. The ww3_grid program reports, via standard output, the number of sea points in the grid. In the context of high-rank grids, this is the actual number of sea points. In the case of low-rank grids, there is some ambiguity.

In Section 2.5, Table 1 and Table 2, we gave two values for the number of sea points for the low-rank grids. The larger values, given in parentheses, are the number of sea points reported by ww3_grid. The smaller values, however, are the more meaningful ones. These are the numbers of sea points after applying the "auto-mask", which is an optional feature of the multi-grid WW3. When ww3_grid is run, it only has information about that particular grid. It does not have information about higher rank grids, and so it cannot anticipate the number of sea points that will be removed by auto-masking.

The auto-mask concept is simple: any grid point of the low-rank grid which satisfies two criteria is disabled (masked). These two criteria are: 1) it is covered by a higher rank grid and 2) it is not close to the edge of the higher rank grid.

Note that (2) implies that for some geographic regions, both grids have active sea points. This "overlap" region, discussed in Section 2.6.3, is needed to ensure that the void from omitted calculations in the low-rank grid does not affect the overall solution. The definition of "close" in (2) defines the size of the overlap region (Section 2.6.3).

2.6.3. Latitude of grid overlap (handoff)

As just described, the two-way communication (multi-grid) approach requires overlap between grids. This overlap defines the region of the low-rank grid for which computations can affect the outcome of the high rank grid. By contrast, in the portion of the low-rank grid that is far from the boundary, computations can be omitted, i.e. those grid points are disabled using the "auto-masking" feature described in Section 2.6.2. The width of this overlap region depends on the numerical stencil (Tolman 2007, 2008).

With the lower resolution IRI systems (IRI-1/2 and IRI-1/4), the hand-off between grids occurs at a relatively high latitude (Table 2). Thus, a large majority of the work is done in the low-latitude regular grids.

With the higher resolution IRI systems (IRI-1/6 and IRI-1/8), the hand-off occurs at a relatively low latitude, i.e. we put increasing burden on the high latitude grids to reduce the number of sea points in the low-latitude regular grid. Having the hand-off at a lower latitude also makes the two systems more efficient by reducing the ratio between the minimum and maximum spacing (Section 2.6.1).

A goal of the IRI design is to have resolutions match at the hand-off region (as in Figure 2). Since the low-latitude grid resolution increases with latitude, this means that the cases with higher latitude hand-off (IRI-1/2 and IRI-1/4) have higher resolution polar stereographic (PS) grids. Thus, we have an unexpected situation: the PS grids for IRI-1/4 have higher resolution (18 km) than the PS grids for the IRI-1/6 case (21 km).

With the IRI-1/2 system, PS grid active points extend to position south of the Aleutians, around 50°N. The low-latitude grid reaches as far north as Ireland, around 55°N. With the IRI-1/4 system, the overlap is similar, though the low-latitude grids do not reach quite as far north: 54°N.

The IRI-1/6 case is the only one for which the overlap differs between north and south. The hand-off happens between 24.6 and 30.50°N, and 24.1 and 29.3°S.

The IRI-1/8 case also has overlap at a relatively low latitude, 25.1° to 30.9°N and S. An advantage of the 30.9°N cut-off is that the Gulf of Mexico, Red Sea, and Persian Gulf are fully within the higher-rank, low-latitude grid.

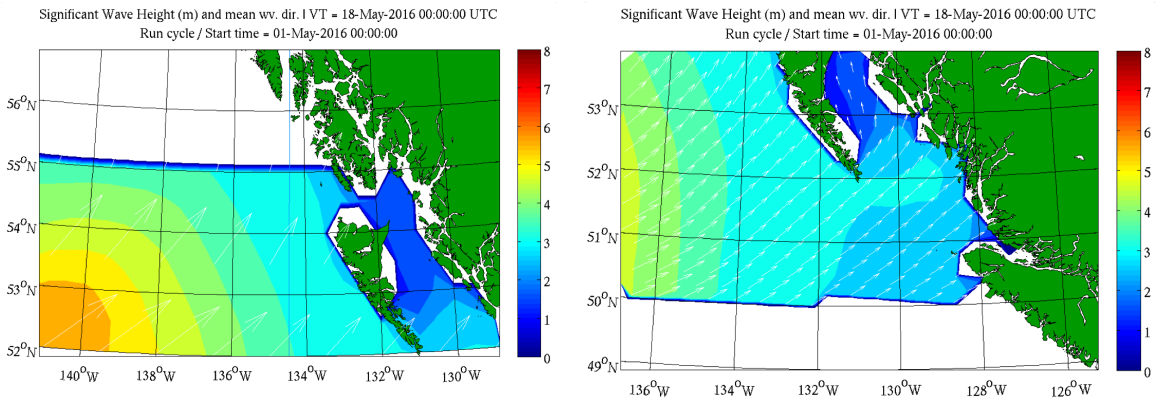

Figure 5. Example of overlap between the low-latitude grid (left) and PS north grid (right). This example is from the IRI-1/2 system, near western Canada and southern Alaska. Note that the regions shown in the two plots are not identical: they are offset by approximately 3° latitude and 4° longitude.

2.6.4. Number of points

The number of columns, rows, and sea points (n_i, n_j, n_{sea} respectively) for various grids are given in Table 1, Table 2, and Table 4. PS grids are often designed to be square in index space ($n_i = n_j$), with the center of the square at the north/south pole. However, with our PS grids, we use $n_j = n_i + 2$, since this makes it easier to catch some user errors, such as accidentally transposing a field.

The strategy for balancing (or not balancing) the number of sea points n_{sea} between PS and low-latitude grids was discussed in Section 2.6.3.

2.6.5. Maximum latitude

This section describes the maximum latitude of the global systems as a whole. The maximum latitude of the low-latitude components of the IRI systems was already discussed in Section 2.6.3.

Systems which include irregular grids. In the case of PS grids and tripole grids, the grids themselves are not limited in latitude. However, to avoid a singularity associated with the model

definition of wave direction (Rogers and Campbell 2009), latitudes north of 89°N are masked. This is done using either the model mask file, or by specifying the points as land in the depth file (the outcome is the same either way).

Systems with regular grids only. The maximum latitude of single grid systems is given in Table 4. In the case of the regular grid single-grid systems, there are no sea points above a particular latitude ϕ_{max}: points further north are either not part of the grid, or they are in the grid but masked out. The same is true for the southern limit, though this is typically not a concern since there are few sea points south of 78°S. Operationally, ϕ_{max} is typically between 77° and 82°N. There are no good options for selection of ϕ_{max}. If a large ϕ_{max} is used (e.g., 84°N), the convergence of meridians forces WW3, with its explicit propagation scheme, to adopt a small propagation time step, which slows the model. If a smaller ϕ_{max} is used, e.g., 78°N used operationally by Tolman et al. (2002), large sections of the Arctic Ocean are omitted. This is especially problematic in the present decade, since reduced ice cover (e.g., Stammerjohn et al. 2012) implies greater wave activity at high latitudes (e.g., Thomson and Rogers 2014; Thomson et al. 2016; Stopa et al. 2016; Ardhuin et al. 2016).

2.6.6. Corner trimming

Structured grids are logically rectangular. With a PS grid centered at the north/south pole, the result is that the corners reach to lower latitude than the sides of the "rectangle". In other words, for the north PS grid, the minimum latitude of the PS grid varies significantly with longitude (e.g., Figure 4-12 of Rogers and Campbell 2009 and Figure 2 of Rogers and Orzech 2013). This is not only unattractive: it is also undesirable insofar as the minimum latitude of the PS grid determines the limiting grid spacing (e.g., Figure 2) and thus the propagation time step of the PS grid. We address this by masking the "corners" of the PS grids in the IRI system. The unmasked points are therefore circular when viewed in Cartesian coordinates; i.e. the minimum latitudes of sea points in the PS grid are uniform with longitude (where there is not land, of course).

2.6.7. Process fraction

With multi-grid WW3, the user specifies how many processes are allocated to each grid. This is in terms of a fraction of the total available. However, when grids are of different rank, they are solved sequentially rather than concurrently. Thus, when there are no two grids of the same rank, the only logical approach is to instruct the model to assign all processes to work on all grids. For example, if 30% of processes are assigned to the grid of rank 3, then 70% of processes will be idle while this grid is being solved. The only possible advantage to this is the reduction of memory usage, which may make the difference between success and failure in the case of large problems such as IRI-1/8. However, the same memory reduction can be achieved more simply via instructions in the batch submission script, e.g., on the DSRC Crays at time of writing, to only use 16 cores per node instead of the full 32 cores per node. This is discussed further in Section 8.3.1.

2.6.8. Time step sizes

Time step sizes for the IRI systems are given in Table 3. All three grids in each IRI system use the same time steps. However, it should be understood that this is not a general requirement of

the multi-grid approach, but rather an outcome of a design feature: the IRI grids have similar resolutions. This, in turn, results in similar time step sizes.

The two time step sizes which are listed are the "global time step size" Δt_g and the "propagation time step size" Δt_p. The second is typically the first one that a modeler decides upon, since it is unambiguously defined: it is set such that the Courant number $C_{gx}\Delta t_p/\Delta x \leq 1$ for the fastest wave components (i.e. the lowest frequency bin). Maximum speed is achieved using $C_{gx}\Delta t_p/\Delta x =$1.0, but in practice, the Δt_p is selected so that the Courant number is between 0.8 and 1.0, since it is useful to have a round number, e.g., so that the ratio $\Delta t_g/\Delta t_p$ is an integer.

In cases where the modeler desires a Courant number very close to unity, he/she *should not use* the deep-water calculation for group velocity. Group velocity C_g is higher in intermediate water depth than in deep water. Thus, if the deep-water calculation is used, and Courant number is very close to unity, the model will be unstable in intermediate water depth. A WW3 user might also naturally assume that the impact of currents on stability should also be taken into account, since currents, along with group velocity, determine the advection velocity. However, in this case WW3, automatically adjusts the time step size, and so surface currents do not need to be included in the modeler's Courant number calculation.

The global time step determines the interval at which the propagation and source term operations are synchronized in the model. When the ratio $\Delta t_g/\Delta t_p$ is very large, say $\Delta t_g/\Delta t_p = 12$, energy will be propagated with 12 time steps before applying source term calculations, and if the system is designed such that the Courant number $C_{gx}\Delta t_p/\Delta x$ is near unity for the fastest wave components, then it further implies that the these wave components will propagate a distance of 12 geographic grid cells Δx before applying source term calculations. Thus, advection becomes desynchronized (or less tightly coupled) with the source terms. This is expected to be important if there is variability in geographic space of something that may affect the source terms, e.g., a feature of the wind, coastline, or bathymetry.

To help us understand the implication of the ratio $\Delta t_g/\Delta t_p$, we can compare two cases at either extreme. Say for Case A, $\Delta t_g/\Delta t_p$=3600/1200=3. And Case B: $\Delta t_g/\Delta t_p$=3600/100=36. In what sense is Case B worse? Both have the same synchronization interval, Δt_g=3600 s. The group velocity is a physical quantity, so the distance travel during the synchronization interval is identical. Neither is more or less likely to skip past an important feature in the wind or coastline without seeing it. Thus, we can see that Case A is *not more accurate* than Case B. The difference is that Case B takes many steps in propagation before synchronizing. Thus, the real defect of Case B is that it is doing more work (taking many small steps) but is discarding much of the benefit of this extra work via infrequent synchronization. Case B is wasteful relative to A.

Now take Case C: $\Delta t_g/\Delta t_p$=300/100=3. Both B and C have the same Δt_p, but C is more accurate (and expensive) than B because of the more frequent synchronization. This time, Case B is inaccurate. From these two examples we can conclude that a large $\Delta t_g/\Delta t_p$ ratio is worse, but the ratio on its own does not tell us *why* it is worse: it depends on the two cases that we are comparing, since it can be worse due to waste, inaccuracy, or both.

This leads to the obvious question: What should the ratio $\Delta t_g/\Delta t_p$ be? There are multiple approaches.

Approach 1. Follow advice given in the WW3 User Manual. The guidance in WW3 v3.14 (Tolman 2009) (repeated in manuals for v4 and v5) is to use a ratio between 2 and 4: "accuracy does not suffer significantly if the overall time step is [taken] as 2 to 4 times the critical CFL time steps."

Approach 2. Follow the example given by NOAA/NCEP for their WW3 implementations. Fortunately NCEP have traditionally liberally shared their grid designs. In the time of WW3 version 2 (Tolman 2002), the NOAA/NCEP WW3 (NWW3) global grid had 1.25° longitude by 1° latitude (288 by 157 cells) resolution, and $\frac{\Delta t_g}{\Delta t_p} = \frac{3600}{1300} = 2.8$. In the time of WW3 version 3 (Tolman 2009), the NWW3 grid had resolution of 0.5°, and $\frac{\Delta t_g}{\Delta t_p} = \frac{3600}{480} = 7.5$. Thus, we can see that the settings used by NCEP were consistent with the manual when that part of the manual was written, but by version 3, there was a discrepancy between the recommendations in the manual and the settings used by the author of that manual. In any case, Approach 2 is to use a ratio near 7.5.

Approach 3. We can devise our own rule. In the case of the global model, our primary objective is to resolve the features of the wind and coastlines. Resolution of bathymetry, ice, and currents are important but not primary, as they could arguably be for a regional model. Let us address the case of the winds, and take the case where winds are prescribed at a resolution of 0.25°, and conservatively assume that we should resolve 0.25° spatial variability in the winds, and that any parcel of wave energy should have opportunity to be acted on by source terms within a (roughly) 28 km × 28 km square before advecting to a place where the winds will be different. Then let us assume that a representative long (but not necessarily longest) wave has group velocity C_g around 13 m/s. This tells us that the global time step should be $\Delta t_g \sim 2200$ s. On the other hand, if C_g=7.8 m/s, the global time step should be $\Delta t_g \sim 3600$ s. In deep water, this is a 10 second wave. Note that in this example of Approach 3, Δt_p and the ratio $\Delta t_g/\Delta t_p$ are not considered. Another example is to assume that resolution of the coastline is paramount. This is dictated by the WW3 computational grid resolution Δx. If we use a medium group velocity again, C_g=7.8m/s, this implies that if $\Delta x < 28$ km, then we should use $\Delta t_g < 3600$ s.

For the IRI designs, we use $\Delta t_g = 3600$ s for all cases (Table 3). Thus the ratio varies between $\Delta t_g/\Delta t_p$=3 for IRI-1/2 and $\Delta t_g/\Delta t_p$=8 for IRI-1/8. This effectively follows Approach 2. It also follows the Approach 3 example for using the wind resolution and a medium group velocity. Only one of the four designs (IRI-1/2) is consistent with Approach 1.

Table 4 shows the time step sizes used for the single-grid global model examples which are used as points of comparison for the computational efficiency of the IRI systems. Direct comparison of computation time is made in Table 5 to the IRI-1/4 system, which has a ratio $\Delta t_g/\Delta t_p$=5. This presents a dilemma. What ratio should be used in these single-grid global model examples, keeping in mind that the Δt_g value has a major impact on model economy? Since the answer is

not obvious, we use examples with ratios from 5 to 30. The larger ratios in Table 5 (values of 24 and 30) enable these single-grid models to run relatively quickly, but these may be ill-gotten gains, since they are in major violation of the explicit suggestion of Approach 1 and the implicit suggestion of Approach 2.

As noted in Table 3, the IRI-1/8 system uses Δt_p=450 s. It is interesting to point out that this is also the time step that would be required if we use a 0.5° grid that extends to 78°N, because the resolution in that case is 11.55 km at 78°N! This highlights the inefficiency of a regular grid at high latitude.

2.7. Design illustrations

2.7.1. IRI-1/2

The IRI systems were originally referred to as "PSNAS" (Polar Stereographic North and South). IRI-1/2 was implemented as the "mww3_PS34km" regression test December 2014. Figure 6 illustrates the spacing used in the PS grids and low-latitude grid of the IRI-1/2 system. Figure 7 shows the mask of the north PS grid of the IRI-1/2 system.

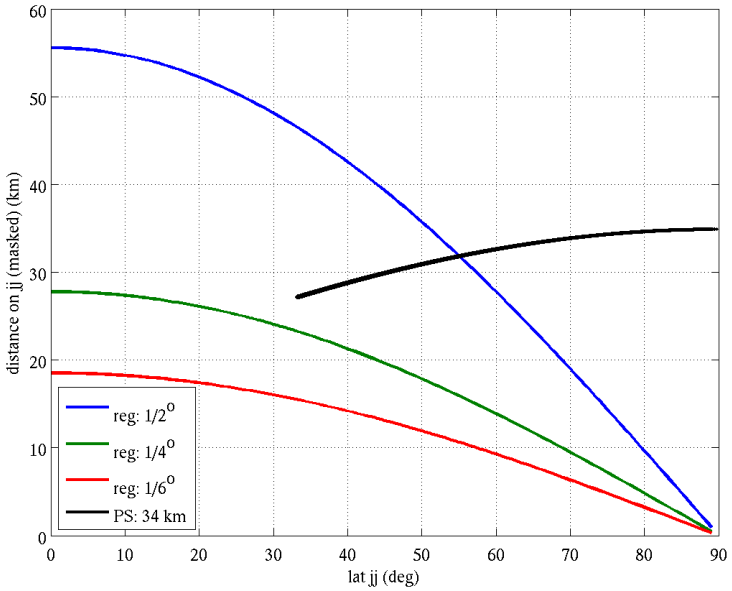

Figure 6. Spacing as a function of latitude. Black and blue lines are the PS and low-latitude grids, respectively, used in the IRI-1/2 system. Green and red lines are low-latitude grids used in other IRI systems.

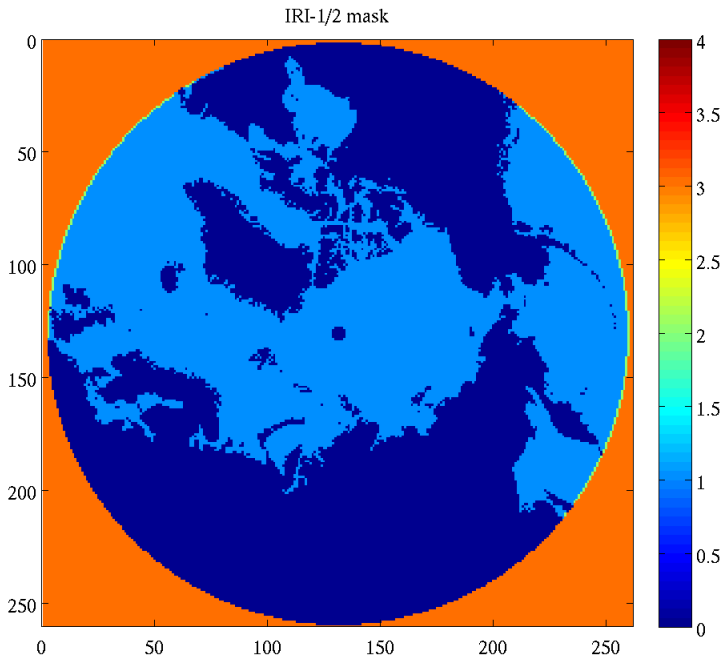

Figure 7. Mask used for north PS grid in IRI-1/2 system, plotted in index space. Colorbar indicates mask type, where: 0=land, 1=sea, 2=boundary, 3=excluded.

2.7.2. IRI-1/4

IRI-1/4 was implemented as the "mww3_PS18km" regression test December 2014. Figure 8 illustrates the spacing used in the PS grids and low-latitude grid of the IRI-1/4 system. Figure 9 shows the mask of the south PS grid of the IRI-1/4 system.

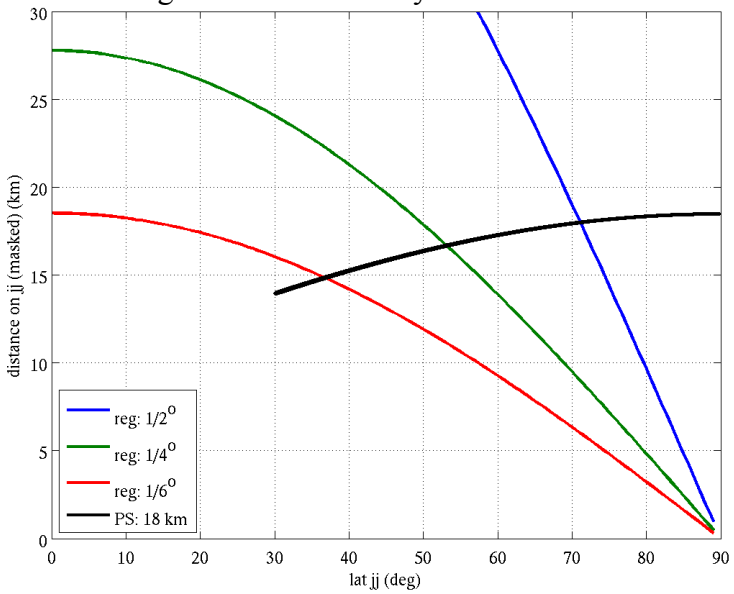

Figure 8. Spacing as a function of latitude. Black and green lines are the PS and low-latitude grids, respectively, used in the IRI-1/4 system. Blue and red lines are low-latitude grids used in other IRI systems.

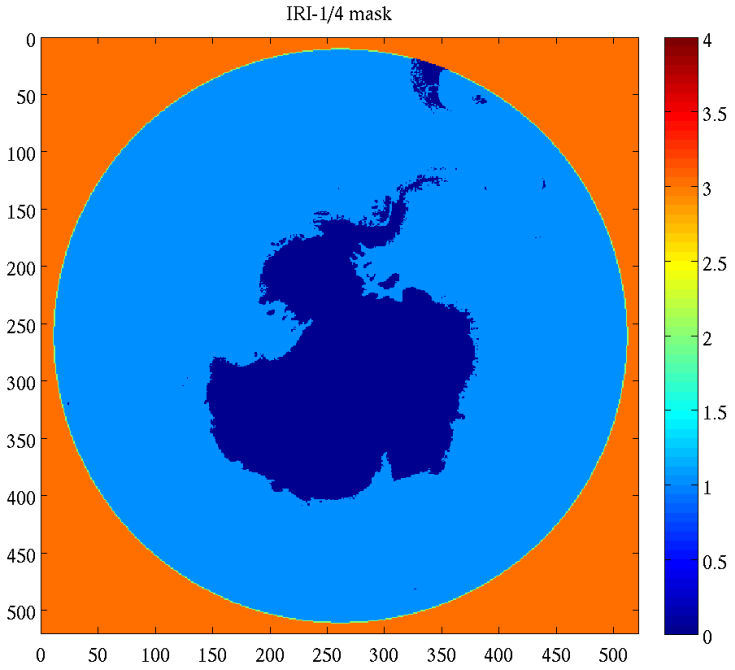

Figure 9. Mask used for south PS grid in IRI-1/4 system, plotted in index space. Colorbar indicates mask type, where: 0=land, 1=sea, 2=boundary, 3=excluded.

2.7.3. IRI-1/6

IRI-1/6 was implemented as the "mww3_PSNAS_1_6th" regression test July 2016. Figure 2 illustrates the spacing used in the PS grids and low-latitude grid of the IRI-1/6 system. Figure 10 shows the active grid points of the north PS grid of the IRI-1/6 system. Here, we do not show the mask provided to WW3, because in the IRI-1/6 system, the PS grids are the lower rank grids, which means that WW3 performs internal masking (or "auto-masking", see Section 2.6.2) of this grid. Thus, the mask provided as input to WW3 does not provide a good indication of the active grid points. Figure 10 reflects the auto-masking.

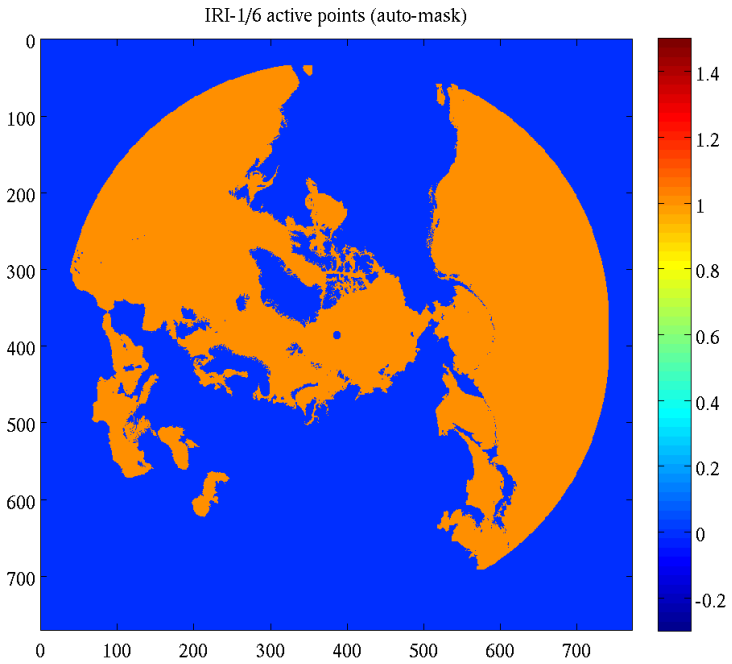

Figure 10. "Auto-mask" for north PS grid in IRI-1/6 system, plotted in index space. Colorbar indicates mask type, where: 0=inactive, 1=active.

2.7.4. IRI-1/8

IRI-1/8 was implemented as the "mww3_PSNAS_1_8th" regression test August 2016. Figure 11 illustrates the spacing used in the PS grids and low-latitude grid of the IRI-1/8 system. Figure 12 shows the active grid points of the south PS grid of the IRI-1/8 system. Figure 13 shows example output.

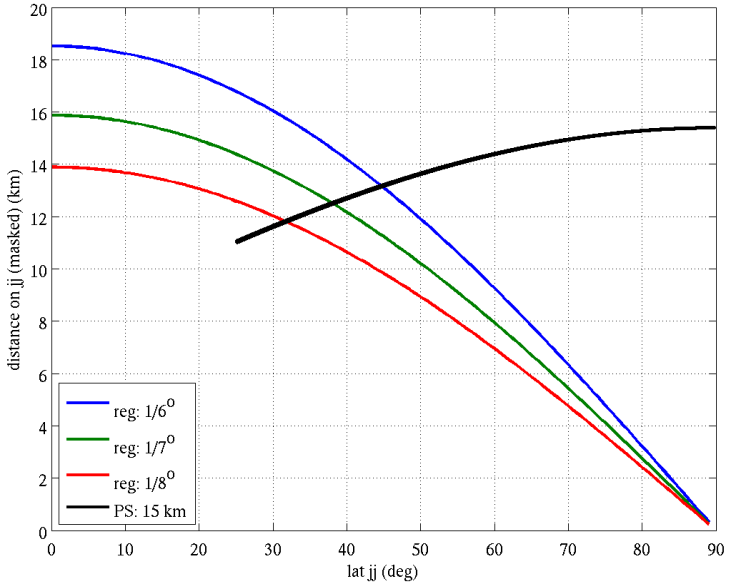

Figure 11. Spacing as a function of latitude. Black and red lines are the PS and low-latitude grids, respectively, used in the IRI-1/8 system. Blue and green lines are low-latitude grids used in other IRI systems.

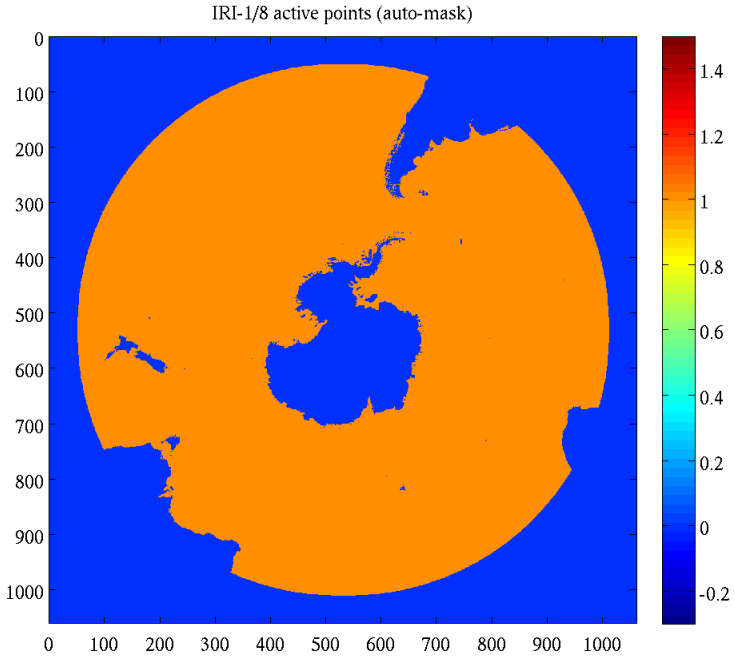

Figure 12. "Auto-mask" for south PS grid in IRI-1/8 system, plotted in index space. Colorbar indicates mask type, where: 0=inactive, 1=active.

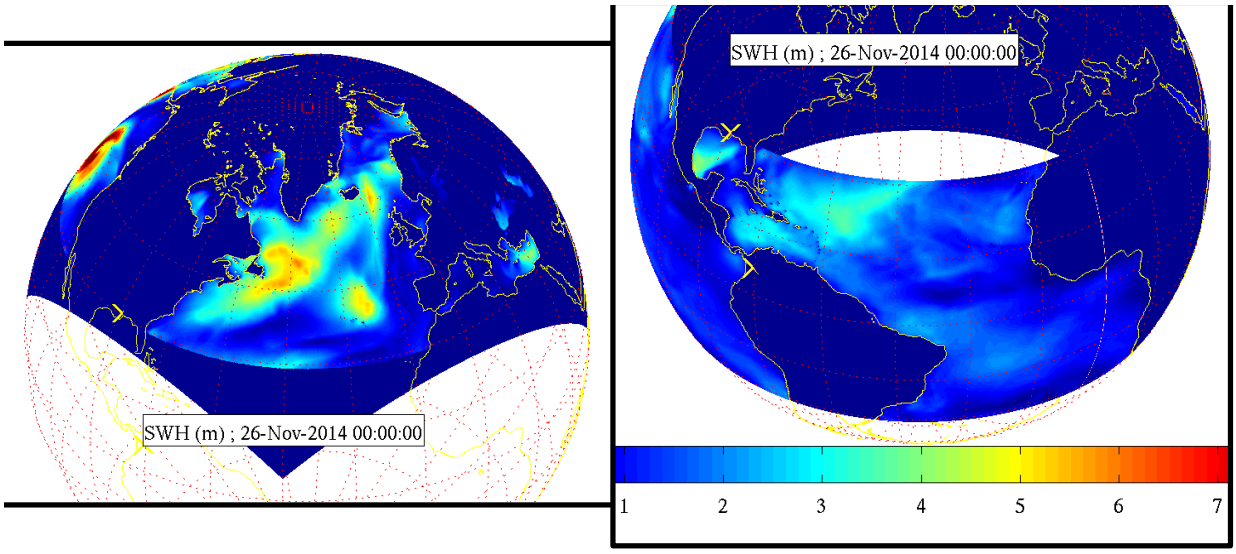

Figure 13. Example of output from IRI-1/8. Significant waveheight is shown in meters. The north grid (15 km resolution at 70˚N) and low-latitude grid (1/8˚ regular spacing) are shown. The south grid is included in this simulation but is not shown here.

3. Timings

3.1. IRI-1/4 vs. single-grid systems (direct comparison)

Direct comparison of the IRI-1/4 system against global systems based on single grids is tabulated in Table 5. Here we will refer to the case of a single global grid with regular 1/4° spacing as

"single-1/4". That model's efficiency is determined by a) the maximum latitude, which determines the required Δt_p, and b) the ratio $\Delta t_g/\Delta t_p$ (Section 2.6.8). In terms of (a), we include two possible values: 78.5° and 84° latitude. The smaller value, 78.5°, is convenient because 78.5°S corresponds to the southernmost sea point, with bathymetry at 1/4° resolution. The higher value, 84°, is disadvantageous to the single-1/4 design, since it requires a small Δt_p. In terms of (b), we include several variations, and some of them are strongly advantageous to both the single-1/4 and tripole designs.

When comparing these timings, it should also be kept in mind that the tripole grid simulations are performed with first order propagation (PR1), while all other simulations use higher order propagation scheme (PR3). In terms of timings, this is of course advantageous to the tripole simulations.

Table 5 quantifies the relative benefit of the IRI-1/4 system over the single-grid systems with comparable number of sea points. The benefits depend on which model it is being compared with; these include: 1) Smaller ratio $\Delta t_g/\Delta t_p$ (Section 2.6.8), 2) inclusion of a larger fraction of the Arctic, 3) faster computations, and/or 4) more accurate (higher order) propagation scheme. Only one of the comparison models has speed comparable to IRI-1/4: the single-1/4 which uses a large $\Delta t_g/\Delta t_p$ and excludes the Arctic north of 78.5°N. If we eliminate (1) by comparing only cases with $\Delta t_g/\Delta t_p$=5 (Table 6), we see that IRI-1/4 is superior to single-1/4 in terms of (2) and (3), and that it is superior to tripole in terms of (3) and (4). The speed advantage relative to tripole is a factor 4.9, which is a large advantage in context of operational turnaround.

Table 5 shows grid information and timings for cases with direct timing comparison. Timing metrics are given in two forms, where WMPWS=420/DSPWH:

1. DSPWH = "days simulated per wallclock hour"
2. WMPWS = "wallclock minutes per week simulated"

Timings are created using Dell PowerEdge M630 blade servers with Intel Xeon CPUs at 2.60GHz: five (5) nodes with 28 cores each (140 MPI processes).

The subset of cases in Table 5 with time step size ratio of 5 are given in Table 6.

Table 5. Grid information and timings for cases with direct timing comparison.

design	max latitude	nsea	global Δt_g (s)	prop. Δt_p (s)	ratio $\Delta t_g/\Delta t_p$	DSPWH (days)	WMPWS (minutes)
IRI-1/4	n/a	704k	3600	720	5	13.08	32.1
1/4° global	78.5°	621k	3600	240	15	11.39	36.9
"	"	"	1200	"	5	5.75	73.0
"	84°	648k	3600	120	30	6.33	66.4
"	"	"	960	"	8	3.50	120.0
"	"	"	720	"	6	3.29	127.7
"	"	"	600	"	5	3.21	130.8
tripole (PR1)	n/a	991k	3600	150	24	4.81	87.3
"	"	"	750	"	5	2.68	156.7

Table 6. Like Table 5, but only showing cases with time step size ratio of 5

design	max latitude	DSPWH (days)	WMPWS (minutes)
IRI-1/4	n/a	13.08	32.1
1/4° global	78.5	5.75	73.0
1/4° global	84	3.21	130.8
tripole (PR1)	n/a	2.68	156.7

3.2. Example timings

We have performed a number of simulations with the various global grid designs. However, most do not use the same computer architecture and/or same number of processes as are used in the simulations compared in Section 3.1. These timings are given here in Table 7.

Notation in the table are as follows:

- "Local" indicates that the simulation was run on a local cluster, composed of Dell blades with Intel® Xeon® CPUs, PowerEdge M610 or M630.
- "IBM" indicates IBM Dataplex system, which is the prior generation of primary supercomputer at the DoD Supercomputing Resource Center (DSRC).
- "Cray" indicates a Cray XC40, which is the current generation of primary supercomputer at the DSRC.
- “Full CPU” indicates that all processes are working on all grids. This is the optimal setting for IRI systems, except for some applications of IRI-1/8. See Section 2.6.7.
- “Partial CPU” indicates that less than 100% of processes are distributed to each grid. This will tend to cause the IRI system to run at slower-than-normal speed (e.g., half speed).
- "Cores" is the number of cores used. This is equal to the number of processes used unless otherwise noted. It is given in format “`X*Y = # of cores`”, where `X` is number of nodes and `Y` is number of cores on each node.
- In the tripole cases, the number of processes is indicated under “additional information” and is different from the number of cores. On the IBM machines, these simulations would run out of memory if every core was used.
- D, H, DSPWH: see table caption.
- "Cold start" indicates that the model was started with an initial condition based on local wind speed and grid spacing. A cold-started model may run faster than another model during the first several days, since more spectral bins are empty.
- "Spun up" means that the timing is taken from a model well after the cold-start. This also implies that initialization operations are not included in timings.
- "IC0" indicates that the sea points are disabled when ice concentration is greater than 75%. This will tend to make the model run faster. However, the enabling/disabling of sea points is not optimized for parallelization, so the speed-up is not necessarily proportional to the number of points disabled.
- "IC4" indicates that no sea points are disabled due to ice.
- Where IC0/IC4 is not indicated, assume IC0 was used.
- PR1 or PR3 indicates propagation scheme: first or third order, with the latter being more accurate and expensive.

- "With currents" indicates that surface currents are included in the forcing. Where it is not specified, assume the model is run without currents. A single test is included here to quantify the slow-down associate with including the effects of surface currents.
- In some cases, the "additional information" includes a breakdown of the model time into two parts: 1) initialization, and 2) time-stepping. The latter is the main part of the computations.

Though the timings here are not directly comparable to the timings in Section 3.1, in some cases, there are pairs of simulations that can be compared to each other. We list these by their simulation numbers which are given in Table 7:

- 1 vs. 2: compares single-1/2 to IRI-1/2
- 8 vs. 9: compares PR1 vs. PR3
- 9 vs. 11: compares IRI-1/4 vs. tripole
- 10 vs. 11: compares "with currents" vs. "no currents"

One noticeable feature of these timings is the very fast computation times for the IRI-1/4 system on the IBM machines, performed in March 2015. It is much faster than the tripole system on the same architecture, and is much faster than the IRI-1/4 system on the newer Cray machines. Unfortunately, these results are not reproducible, as the IBM machines have been retired. Since they are not reproducible, they should not be considered in any decision-making.

Table 7. Example timings for other grid designs. The first column indicates the simulation number, used for referencing in the text. "D" indicates days simulated. "H" indicates wallclock time in hours. "DSPWH" is D/H, i.e. days simulated per wallclock hour. Further explanation is given in the text.

	machine	cores	D (days)	H (hours)	DSPWH (days)	additional information
	single-1/2 with maximum latitude of 77.5°N					
1	local	8*8=64	1	0.0583	17.15	cold start, included for comparison to first "IRI-1/2" example below
	IRI-1/2					
2	local	8*8=64	1	0.0761	13.14	full CPUs, cold start
3	local	12*12=144	1	0.0255	39.22	partial CPUs, fully spun-up model
4	local	12*12=144	141	3.82	36.91	partial CPUs, from cold start
	IRI-1/4					
5	local	10*12=120	1	0.2	5.00	no other info
6	local	12*12=144	1	0.2	5.00	fully spun up
7	IBM	12*16=192	7.75	0.291	26.63	no other info
8	IBM	12*16=192	60	1.966	30.52	cold start, IC0, PR3, no currents, Δt_g=3600 s
9	IBM	12*16=192	60	1.75	34.29	cold start, IC0, PR1, no currents, Δt_g=3600 s
	tripole					
10	IBM	12*16=192	60	18.00	3.33	cold start, IC0, PR1, with currents, Δt_g=600 s, 12*12=144 processes
11	IBM	12*16=192	60	15.817	3.79	cold start, IC0, PR1, no currents, Δt_g=600 s, 12*12=144 processes
	IRI-1/6					
12	local	96	1	0.6917	1.45	cold start, IC0, w/out init.: H=0.625

13	local	264	1	0.2917	3.43	cold start, IC0, w/out init.: H=0.233
	IRI-1/8					
14	local	11*12=132	1	1.3333	0.75	cold start, IC4, Δt_g=1800 s
15	local	11*18=198	1	1.1389	0.88	cold start, IC4, Δt_g=1800 s
16	local	15*18=270	1	1.1111	0.90	cold start, IC4, Δt_g=1800 s
17	local	11*18=198	1	0.9167	1.09	cold start, IC4, Δt_g=3600 s, (init.: H=0.264), (time stepping: H=0.653)
18	Cray	16*16=256	1	0.6611	1.51	cold start, IC0
19	Cray	16*32=512	1	0.4389	2.28	cold start, IC0

Six timings results were made for the same test case, so these are shown separately (Table 8). A remarkable finding is that the test on the local machine with 144 processes ran much faster than the test with 160 processes on a Cray machine. This is currently being investigated. Possible explanations are: different compilers used (e.g., Portland on local machines vs. Intel compiler on the DSRC), different impact of i/o on runtime, different model versions (though this should not have significant impact on runtime), and code changes to address the smaller limit on MPI tag numbers on the Crays. The outcome implies that additional optimization could be performed on the Crays to reduce runtime.

Table 8. Timings for IRI-1/4. All have 1.5 day duration, with cold start, and use IC0.

machine	cores	H (hours)	DSPWH (days)	additional inf.
Cray	12*32 = 384	0.25	6.11	partial CPU
Cray	12*32 = 384	0.25	6.09	partial CPU
Cray	12*32 = 384	0.22	6.76	partial CPU
Cray	12*32 = 384	0.15	10.02	full CPU
Cray	5*32 = 160	0.33	4.49	full CPU
local	12*12 = 144	0.18	8.40	full CPU

4. Propagation tests

As discussed in Section 2.6.2, two different methods are used for passing information between grids. For lower-rank to higher-rank, bilinear interpolation is used. For higher-rank to lower-rank, conservative remapping is used. It is important to verify that both methods work properly for all of the IRI grid systems. For such an exercise, it is important to keep in mind that problems can be masked by wave generation: if a storm is moving from one grid to another, it may appear that energy is propagating between grids even when it is not. The functionality in question is essentially one of swell propagation. Therefore, it is most appropriate to run a propagation test without winds and without source terms. We do this by specifying a signal that exists only within one grid (during model initialization), propagating it until it exists only in the other grid, and tracking its progress during that time. For each of the four IRI systems, there are four possible tests (moving from south grid to low-latitude grid, moving from low-latitude grid to north grid, etc.) which implies 16 possible combinations to test. We tested 14 of the 16 possible combinations. Two are shown in Figure 14 and Figure 15. All tests completed without any indication of problems.

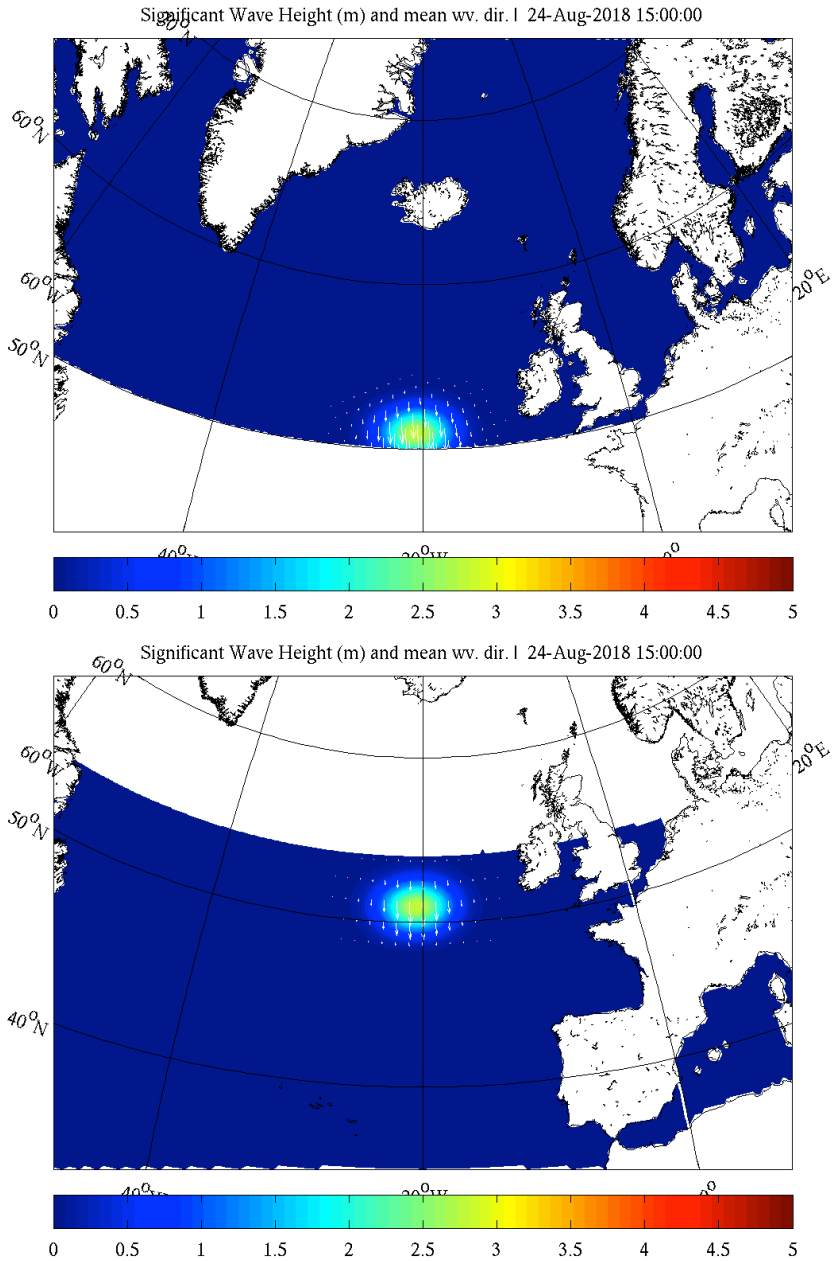

Figure 14. Propagation test for IRI-1/4. Feature starts at a position that exists only in the north PS grid (phase 1, not shown), propagates south, is in the overlap region for a period of time (phase 2), propagates further, and then exists only in the low-latitude grid (phase 3, not shown). This graphic corresponds to phase 2.

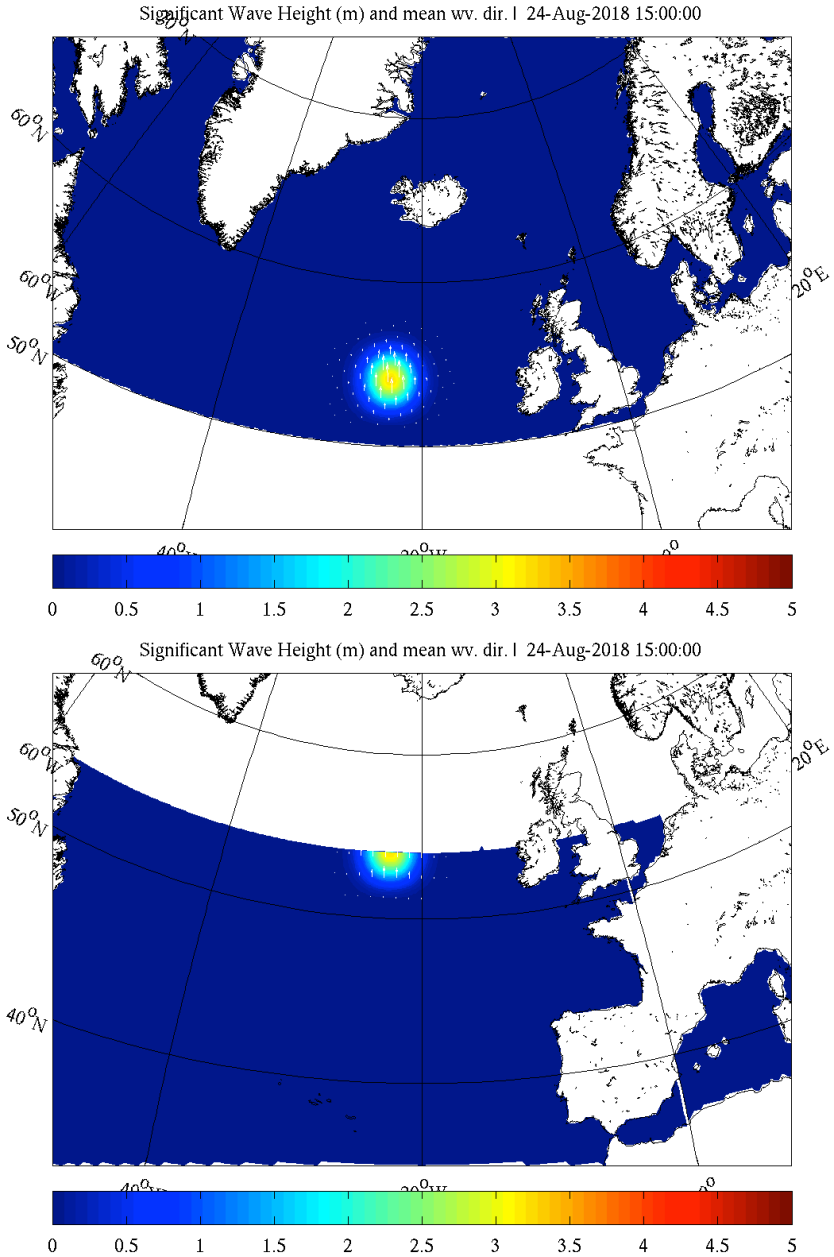

Figure 15. Propagation test for IRI-1/4. Feature starts at a position that existing only in the low-latitude grid (phase 1, not shown), propagates north, is in overlap region for a period of time (phase 2), propagates further, and then exists only in the north PS grid (phase 3, not shown). This graphic corresponds to phase 2.

5. Validation samples

In this section, we provide some limited examples of validation with the IRI systems, and comparable scores from non-IRI systems are included for comparison. Model skill in global hindcasts is primarily determined by the accuracy of the wind forcing and the skill of the deep-water source term package. (Source term packages are denoted as ST4, ST6, etc., and are defined in the glossary.) These factors are independent of the grid design. Thus, validation has secondary importance in the present report, and is why we only present limited results here.

We have applied all four IRI systems in hindcasts with realistic inputs (e.g., winds) and full spectral range[1]. The IRI-1/4 system has been exercised the most, followed by IRI-1/2. For example, IRI-1/4 has been applied for at least 11 unique multi-month time periods, for various purposes, and is used currently in a short-range, real-time system operated by James Dykes (NRL 7322) that includes data assimilation. The IRI-1/6 and IRI-1/8 systems have only been applied in short-term hindcasts, e.g., Table 7, due to computational cost.

5.1. IRI-1/2 hindcast, validated with a single buoy

In this section, we present sample results from the IRI-1/2 system. The selection of IRI-1/2 over IRI-1/4 in this case is not significant: IRI-1/2 was selected simply because comparisons had already been made which are particularly appropriate to our purposes here.

As mentioned above, validation has secondary importance in the present report. The primary question is whether the grid design and multi-grid features are functioning as intended. The most complex issue here—and thus most likely to suffer from human error or coding bugs—is the passing of information between the grids (two-way boundary forcing). We have already tested this with idealized propagation tests (Section 0). Validation of a realistic hindcast with the full complement of source terms provides additional confidence. With this motivation in mind, we select a geographic location that lies within the grid overlap, where the boundary forcing occurs (50°N to 55°N in this case, see Table 2). The University of Washington Applied Physics Laboratory (UW/APL) maintains a Datawell Waverider buoy at Ocean Station Papa. It has the CDIP (Coastal Data Information Program) designation of buoy number 166 and the NDBC (National Data Buoy Center) buoy number is 46246. The buoy was positioned at 214°E 50.03°N during the time period considered. Ocean Station Papa is far from any coastline, and the wave climate at the position is dominated by large extra-tropical storms. These two factors imply that geographic resolution is not important: i.e., the coarser IRI-1/2 system is sufficient. Because the Datawell Waverider is a dedicated wave buoy with a long history of success (e.g., O'Reilly et al. 1996), we believe that it has higher accuracy than larger buoys which mount additional equipment, such as meteorological instruments. The smaller size also means that it can measure wave spectra that extend into somewhat higher frequency range (to 0.58 Hz), relative to larger NDBC buoys. There is an important caveat however: small buoys suffer more from biofouling on long deployments, which has a large impact on the high frequency portion of the spectrum (Thomson et al. 2015). For this reason, we use a time period immediately following the deployment of a clean buoy.

The first three weeks are treated as spin-up, and output is compared to observations 5 January to 1 July 2015. The model is forced by winds from the Navy global atmospheric model NAVGEM (Hogan et al. 2014), and the associated ice concentration analysis, derived from satellite passive microwave radiometer.

Settings used for all grids are as follows:

[1] This refers to the number of frequency bins used. For idealized propagation tests, it is possible to run the model with just a few frequency bins. In a realistic hindcast, the number of frequency bins follows that which would be used operationally.

- Run time is 15 December 2014 to 1 July 2015.
- 36 directional bins
- 31 frequency bins, from 0.0418 to 0.7294 Hz
- Minimum source time step is 300 s.
- Ardhuin et al. (2010) physics (denoted "ST4").
- ST4 physics has a parameter, β_{max} ("BETAMAX" in the WW3 code), intended to be adjusted to accommodate gross wind bias. (Altering this parameter is, in practice, similar to a "blunt tool" adjustment of the wind speeds.) In these simulations, we used $\beta_{max} = 1.2$. This is a relatively low value, appropriate to the positive bias of the NAVGEM wind forcing[2].

This hindcast does not include data assimilation.

We use four different wave parameters in the validation examples here:

1. Waveheight: We use the spectral significant waveheight, or zero-moment waveheight is calculated as $H_{m0} = 4\sqrt{m_0}$, where $m_0 = \int_{f_{min}}^{f_{max}} E(f)df$. The integration limits f_{min} and f_{max} are chosen so that they are consistent (identical) between model and buoy. Here, as is typically true, f_{min} is limited by the model spectrum range and f_{max} is limited by the buoy spectral range: $f_{min} = 0.0418$ Hz and $f_{max} = 0.58$ Hz.
2. Partial waveheight: $H_{m0,p} = 4\sqrt{E_p}$, where $E_p = \int_{f_1}^{f_2} E(f)df$. Here, f is frequency (units of Hz), $E(f)$ is the non-directional (i.e. one-dimensional) wave energy spectrum with units of m^2/Hz, and f_1 and f_2 are specific frequencies bounding a portion of the spectrum $E(f)$.
3. m_4: This is simply the fourth moment of $E(f)$. Similarly, m_0, introduced above, is the zeroth moment. The general definition is: $m_n = \int_{f_{min}}^{f_{max}} E(f)f^n df$. The higher moments are of course more sensitive to the energy level in the higher frequencies. The m_4 is proportional to the mean square slope of the sea surface.
4. Mean wave period: There are several ways to define the mean wave period. Here we use $T_{m,-1,0} = m_{-1}/m_0$. This definition is favored over other mean period definitions because it is less influenced by energy in higher frequencies, and so gives a value roughly corresponding to the peak period T_p, e.g., $T_{m,-1,0}$ might be $0.90T_p$ or $0.93T_p$. At the same time, it is preferred over direct use of T_p since T_p does not vary smoothly. Rather, T_p has a limited number of possible discrete values, depending on the resolution of the frequency bins. For example, in our WW3 hindcast, the WW3 spectral bins correspond to T=…9.2, 10.1, 11.2, 12.3, 13.5, 14.9, 16.3…(in seconds).

[2] For example, we can see this in the most recent (at time of writing) JCOMM report (October 2017). Documents describing these can be found at www.jcomm.info. These comparisons are made by Dr. Jean Bidlot of ECMWF. A second type of validation is at https://portal.fnmoc.navy.mil/verify_cgi/, but this is available only to users within the Navy internet domain. We checked multiple results at this site, most recently the one labeled "DTG 2018100700". In all cases, NAVGEM 10-meter wind speed bias is positive, though it is usually not large, e.g. +0.03 m/s in the JCOMM report. Also relevant: the β_{max} parameter is customarily set to a larger value—e.g., $\beta_{max} = 1.45$ in the v5 WW3 public release documentation—when ECWMF winds are used, and this same JCOMM document (October 2017) reports a negative bias for ECMWF winds (-0.09 m/s).

We include five different skill metrics here:

1. Bias, i.e., mean error.
2. Root-mean-square error (RMSE)
3. Pearson's Correlation coefficient (CC).
4. Scatter index (SI), the standard deviation of the differences (model vs. observation), normalized by the observed mean.
5. Normalized bias (nbias): bias normalized by the observed mean.

Figure 16 compares partial waveheight. The energy is grouped into four frequency bands here:

1) a band containing energy that is typically to the left of the spectral peak, f_1=0.0418 Hz and f_2=0.06 Hz,
2) a first energy band that is typically near the spectral peak, 0.06 to 0.09 Hz,
3) a second energy band that is typically near the spectral peak, 0.09 to 0.2 Hz, and
4) a band containing energy that is typically to the right of the spectral peak, 0.2 to 0.3 Hz.

Accuracy is high in bands 2, 3, and 4. Accuracy is less in band 1, which we ascribe to two factors. First, we use the Discrete Interaction Approximation (DIA) for the four-wave interaction source term S_{nl4}, which has a tendency to push too much energy to left of the peak (see Rogers and van Vledder 2013). Second, this band has the highest fraction of older swell energy, which typically contains more error than windsea and young swell, since it is more afflicted by numerical propagation error, as well as error in the spectral distribution of the deep-water source terms. Errors in directional distribution for band 1 become more evident as energy disperses as swell.

Figure 17 evaluates significant waveheight. This is, of course, the most common and traditional type of wave model evaluation. The accuracy is good: correlation is 0.973 and RMSE is 28.5 cm. Accuracy for waveheight is better than the accuracy for constituent parts (Figure 16), because fortuitous cancellation of errors is not uncommon.

Figure 18 evaluates m_4. This parameter is not often included in model evaluation, so it is unsurprising that it is less accurate than waveheight. Accuracy is still good, however: correlation is 0.94. The m_4 parameter is more responsive to local winds than H_{m0}, so the m_4 correlation is probably a good indicator of how well the wind forcing tracks short-term changes in the true winds. Notably, where H_{m0} has a normalized bias of +3%, m_4 has normalized bias of -8%. This implies that the highest frequencies in the wave spectrum (up to 0.58 Hz) are somewhat underpredicted. Also, since the biases are of opposite sign, it is unlikely that both are caused by systematic bias in the wind forcing.

Figure 19 evaluates $T_{m,-1,0}$. Correlation is lower than for waveheight and m_4 (0.89) and normalized bias is +6%. The latter is consistent with the +10 cm bias in the first and second frequency bands presented in Figure 16.

Taken as a whole, and in context with other validations of the same set of wave parameters, the results indicate no issues of concern within the grid overlap region.

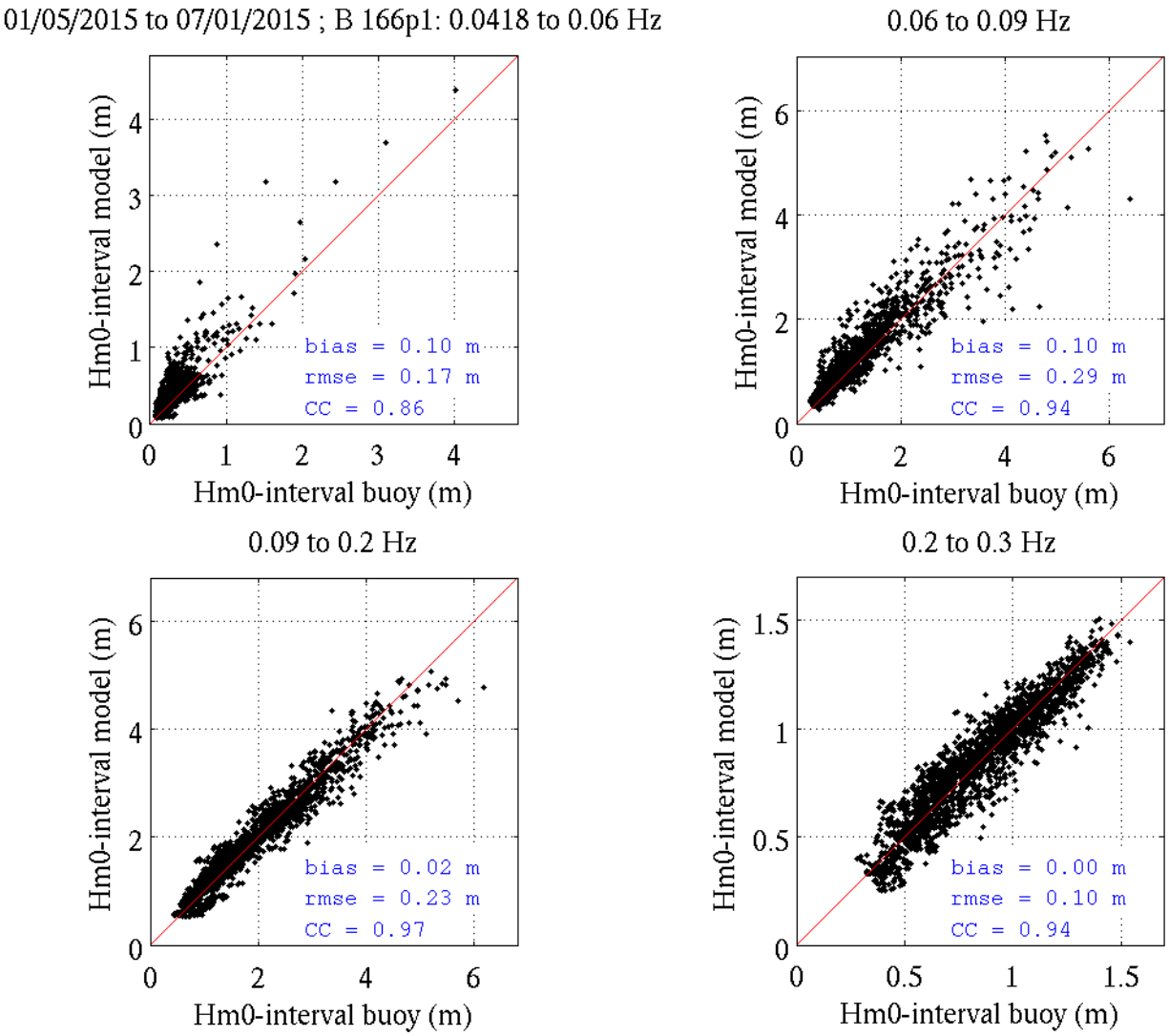

Figure 16. Scatter plots for partial waveheight, IRI-1/2 vs. UW/APL buoy at OS Papa (CDIP 166). The time considered is 5 January to 1 July 2015. The frequency range (f_1 and f_2) are given above each plot.

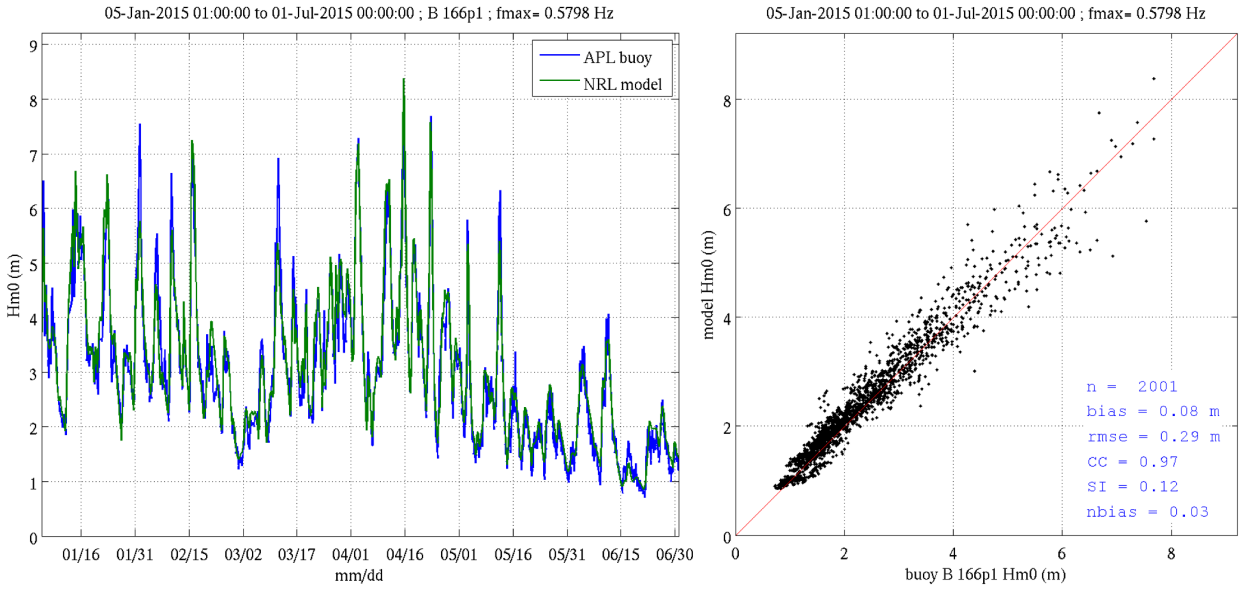

Figure 17. Like Figure 16, but showing time series and scatter plot for zero-moment waveheight, H_{m0}. "n" indicates the number of co-locations, and the five skill metrics are explained in the text.

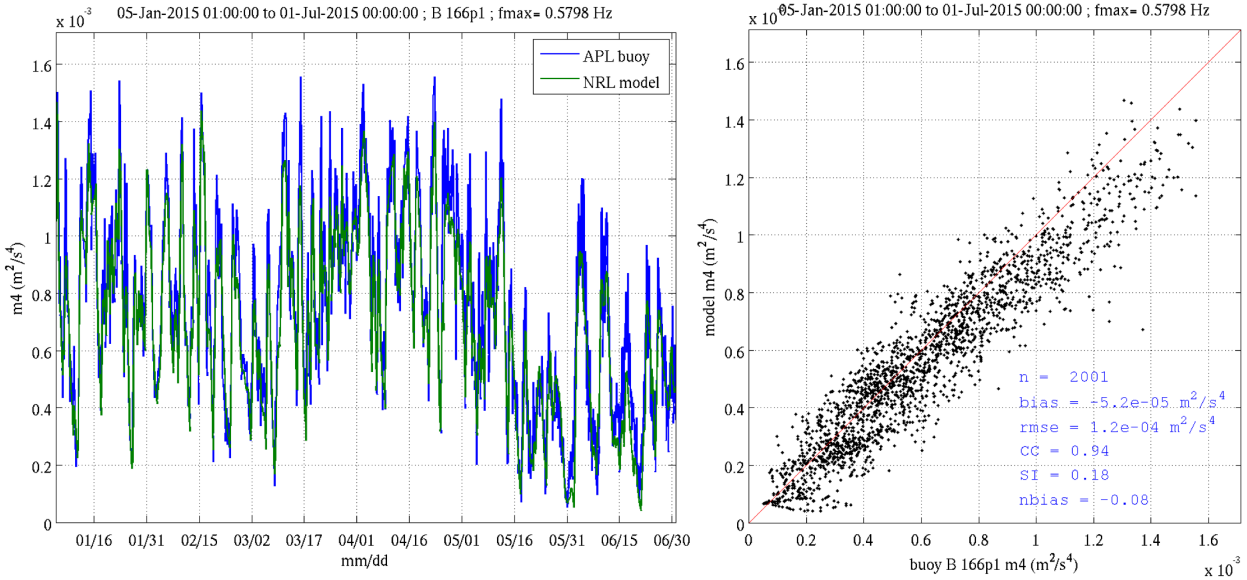

Figure 18. Like Figure 16, but showing time series and scatter plot for the fourth moment of the wave spectrum, m_4

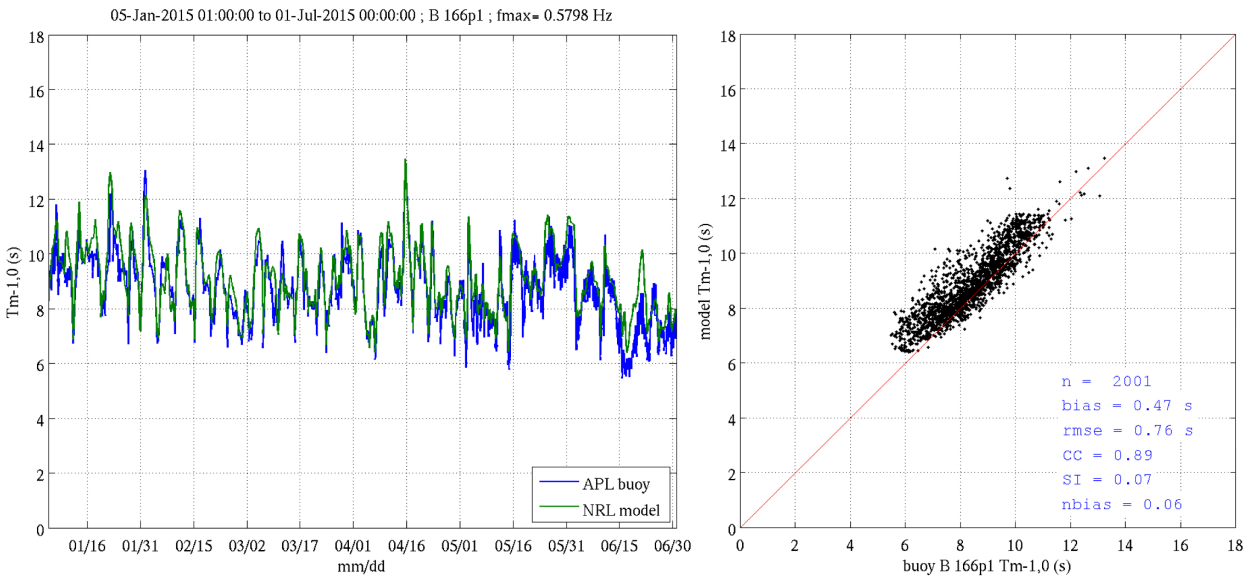

Figure 19. Like Figure 16, but showing time series and scatter plot for the mean period, $T_{m,-1,0}$

5.2. Other IRI results (Fan and Dykes)

Here, we report H_{m0} statistics from others at NRL who are using the IRI-1/4 grid system (Table 9). These include:

1. Work by Dr. Yalin Fan (NRL 7322, denoted "YF" in Table 9): A hindcast for all of 2015, compared to buoy data provided by Dr. Jean Bidlot (ECMWF, European Centre for Medium-Range Weather Forecasts), who collected it for operational wave model evaluation as part of a JCOMM (Joint Technical Commission for Oceanography and Marine Meteorology) collaboration (e.g., Bidlot and Holt 2006). The quality control and averaging methods used by YF are likely different from those used in the model-data comparisons published by Dr. Bidlot. This hindcast does not include data assimilation.
2. Like (1), but for a hindcast of 2017.
3. Work by Dr. Fan: Same hindcast as (1), but comparing to NDBC data.
4. Like (3) but for the 2017 hindcast.
5. Work by Mr. James Dykes (NRL 7322, denoted "JD" in Table 9). Mr. Dykes compared output from the 2015 hindcast of (1) and (3) to altimeter waveheights. These are separated by month, for all 12 months. We report only the January 2015 values here.
6. Like (5) but for the 2017 hindcast used in (2) and (4).
7. Work by Mr. Dykes. Output from a cycling real-time, short-term forecast system is compared to altimeter waveheights (data from 0 to 12 hours, i.e. nowcast to short-term forecast). This hindcast does not include data assimilation.
8. Like (7), but includes data assimilation from NCODA (Navy Coupled Ocean Data Assimilation, Cummings and Smedstad 2013). Altimeter waveheights are used in the DA to create the analysis, and the short-term forecast is compared to subsequent altimeter waveheights (0 to 12 hours).

The hindcasts used in 1-6 represent a baseline model, to be compared against a global coupled modeling system, the Navy's Earth System Prediction Capability (ESPC) system. Results from both the baseline model and coupled model will be included in a future report, led by Dr. Fan[3].

The system used in item 8 is a demonstration system and will form the basis for the future Navy operational global wave model. This is planned for transition during FY19-FY20. It is an uncoupled, or standalone, system. This single system will replace two existing operational models which are semi-redundant. This situation could be confusing: the emergence of an entirely new system, ESPC, and the merging of two existing systems into one, the standalone WW3. Therefore, we summarize the old and new systems in Section 7.

Table 9. Significant waveheight statistics for IRI-1/4 systems run by others at NRL.

Case number	Description	n	RMSE (cm)	bias (cm)	R
1	YF JCOMM 2015	2.8M	49	-3	0.92
2	YF JCOMM 2017	3M	40	-4	0.95
3	YF NDBC 2015	512k	34	+12	0.95
4	YF NDBC 2017	448k	34	+11	0.95
5	JD Jan 2015	6k	52	+10	0.92
6	JD Jan 2017	60k	39	+17	0.96
7	JD Jan 2016	3.6k	36	14	0.97
8	JD Jan 2016+	3.6k	31	6	0.97

5.3. Non-IRI results

We can compare correlation values in Section 5.2 to values published by JCOMM[4] (e.g., Bidlot and Holt 2006), showing statistics from a number of operational centers. For example, in the October 2017 JCOMM report, H_s correlation values range from 0.959 to 0.978 with a median near 0.973, with 10 centers reporting. It should be assumed that most of these operational models include some form of data assimilation, and some may even include the verification data in the assimilation. This implies values given in Table 9 are in the expected range. Table 10 presents four more non-IRI examples. All are based on hindcasts with single grids. Three are global and one is regional. Definitions of ST1, ST4, ST6, NOGAPS, CFSR are provided in the glossary.

1. Rogers et al. (2012). This is the WW3 Validation Test Report. We report the numbers from the global model, compared to altimetry, using ST4 physics with winds form NOGAPS and CFSR.
2. Rogers and van Vledder (2013). This is a SWAN (see glossary) hindcast for Lake Michigan, using ST1 and ST6 physics (see glossary), compared to buoy measurements.
3. Zieger et al. (2015). These are again numbers from the global model, compared to altimetry. ST4 and ST6 results are reported here.

[3] That report will probably be published in late 2019 or early 2020. Prior to publication of the report, readers within the Navy internet domain can find these results from Dr. Fan and Mr. Dykes at https://alvin.nrlssc.navy.mil/index.php/WW3_ESPC_VTR_progress.

[4] Comparisons are made by Dr. Jean Bidlot of ECMWF. Documents describing these can be found at www.jcomm.info.

4. Liu et al. (2018/19). This paper compares global WW3 (ST4 and ST6 physics) against altimetry.

Table 10. Error metrics from other studies. All metrics are for significant waveheight, H_s.

Case number	Description	n	RMSE (cm)	bias (cm)	R
1	Rogers et al. (2012)	16M	41 to 60	-10 to 32	0.96
2	Rogers and van Vledder (2013)	2k	17	-2 to 2	0.96
3	Zieger et al. (2015)	1.4M	33 to 36	3 to 6	0.97
4	Liu et al. (2018/19)	2M	34 to 39	0 to 4	0.96 to 0.97

These results indicate that model skill with the IRI system are comparable to skill shown in other wave model evaluations.

6. Development timeline

Here, we provide a timeline for milestones relevant to the IRI systems:

- May 2008: First test of WW3 using spherical (latitude/longitude) coordinates and an irregular grid, a Lambert Conformal grid.
- Aug. 2009: First test of WW3 on polar stereographic (PS) grid.
- Sep. 2009: Rogers and Campbell (2009) published. This includes use of PS grid.
- Aug. 2011: WW3 multi-grid modified to operate with irregular grids (unpublished work by NRL), but is inefficient for large grids because of slow initialization.
- May 2013: WW3 multi-grid with irregular grids is made efficient by allowing pre-computed weights (from Kevin Lind, a High Performance Computing Modernization Program employee, working with NRL).
- Mar. 2014: Public release of WW3 version 4 includes irregular grids.
- Jun. 2014: James Dykes (NRL) implemented pre-operational real-time multi-grid system that includes PS grid for the Arctic.
- Dec. 2014: IRI-1/2 and IRI-1/4 created.
- Mar. 2015: Tripole grid and IRI-1/4: completed 60 day hindcasts on DSRC (IBM system).
- Apr. 2015: Obstruction grid created by Arun Chawla (NOAA/NCEP) for IRI-1/4. Associated new mask files and regression tests were also created.
- Aug. 2016: IRI-1/6 and IRI-1/8 created.
- Sep.. 2016: Systematic testing of IRI-1/8 system on DSRC (Cray system)
- Apr. 2017: Systematic testing of IRI-1/4 system on DSRC (Cray system)
- July 2018: Obstruction grids added for IRI-1/2, IRI-1/6, and IRI-1/8, with new mask files and new regression tests.

7. Operational systems

As noted in Section 5.2, the Navy currently has two operational global wave models, and its plan is to have two new operational global wave models. This is not a simple one-for-one swap. Instead one system is entirely new, and another is replacing the two existing systems. The two existing systems are described here for the two operational centers at which they reside.

7.1. Existing systems

1: FNMOC-Monterey
Historically, FNMOC has been responsible for global wave modeling, e.g., Wittmann (2002), as well as larger regional grids, e.g., for the western Atlantic or eastern Pacific. Until recently, the global model was a 0.5° resolution single-grid implementation of WW3. In April 2018, FNMOC-Monterey implemented a three-grid system with regular grids: 0.5°×0.25° resolution at high latitudes, and 0.25°×0.25° at low latitudes. The latter is identical to the low-latitude grid of the IRI-1/4 system. Coverage with this system stops at 80° latitude. This uses NCODA version 3, which is not the most recent system, and requires that WW3 grids must be regular. This three-grid system could be called an RRR (Regular-Regular-Regular) approach. It is an interim solution, to be used until the IRI-1/4 system is transitioned to operations. Official global graphics produced by Navy customers are taken from this model.

2: FNMOC-Stennis
This uses a multi-grid implementation of WW3 (Rogers et al. 2012, 2014). It includes a global grid and nine regional grids, such as western Atlantic, eastern Pacific, and Arctic. Most regional grids are 0.2° resolution, and areas not covered by the regional grids are represented by the 0.5° global grid. This modeling system, which replaced a prior system based on WAM (Jensen et al. 2002), exists to provide boundary forcing for coastal wave models (SWAN implementations) and is historically separate from the FNMOC-Monterey global model because of the difficulty associated with communication of boundary forcing from one center to another. FNMOC-Stennis has responsibility for coastal wave modeling, and produces corresponding official products to the customer. However, graphics from the FNMOC-Stennis global model are not provided to the customer, since this would conflict with the FNMOC-Monterey mission; to our knowledge, such graphics are not routinely created. The primary design motivation of the FNMOC-Stennis system is alluded to in Section 2.3.3: most of the regional grids correspond to implementation of the higher resolution atmospheric model COAMPS. Therefore, the system was intended to exploit the improved accuracy of the higher resolution winds. However, this has two problems. First, COAMPS wind fields are available later, and for some grids, much later. This implies a long delay (e.g., 1 or 2 hours) in the completion of the global wave model. Second, the underlying assumption that COAMPS wind forcing is more accurate that NAVGEM wind forcing has become increasingly difficult to support (see Discussion Section 7). Because of the first problem, FNMOC-Stennis has actually reverted to using NAVGEM forcing for the entire multi-grid system. In short, the original design of this system is inconsistent with the current situation, making the system obsolescent. This system does not include data assimilation.

Rogers et al. (2014) give a more detailed overview, contrasting the different approaches at these two centers. Two problems with model (2) above were already mentioned. A third can be added: the two systems are strongly bifurcated. The only key similarity is the underlying code (WW3), though the version numbers are generally different. The differences may be justified on grounds that the two centers have different priorities. However, the disconnected development and validation represent a redundancy cost. Improvements to one system—e.g., fixing a bug in a script, adopting a more accurate physics parameterization, or adding data assimilation—are not readily incorporated in the other system. Such improvements to one system therefore require

modifications to be implemented in the second system. Discussion of redundancies are not fully addressed in the present report, but are continued later in this section.

Lastly, because the FNMOC-Stennis WW3 system is used primarily to provide boundary forcing for regional models, without routine inspection of graphical output from the model, there is enhanced risk that problems associated with human error, scripting bugs, etc., could go unnoticed for an extended period of time, while having a negative impact on those regional models. By merging the FNMOC-Stennis and FNMOC-Monterey models, this danger is essentially eliminated: any problems in the global WW3 which affect the boundary forcing to regional grids will likely be quickly noticed in the routine graphical products for the global WW3.

7.2. Planned systems

The two planned systems earlier mentioned are described in more detail here.

1: Standalone WW3
This is an IRI-1/4 global system that will take the role of the two existing systems. This will run on the DSRC. Data assimilation will use NCODA, as with the existing systems, but will use NCODA 4.1, which accommodates the irregular grids of IRI. This system is already being run in real time by James Dykes (NRL-Stennis) but is not yet an operational system. At time of writing, Mr. Dykes is updating the system to use ice concentration fields from GOFS instead of SSM/I, which should improve accuracy of the wave model at high latitudes (e.g., Rogers et al. 2018). This system, like the two existing systems, will prioritize fast delivery.

2: Coupled system (ESPC)
A WW3 implementation will be part of a coupled system operated for the ESPC. Here, atmosphere, ocean, ice, and wave models are compiled into a single executable (Metzger et al. 2014b). The atmosphere, ocean, and ice models are fully coupled, while the wave model is one-way coupled, meaning that it takes information from the other models (10-meter wind vectors, surface currents, ice concentration) but does not send information back to the other models. The latter is omitted because the software for exporting from three wave model grids to the other models does not yet exist within the coupling framework. Data export is only possible within the coupling framework from a single wave model grid, such as the tripole grid.

At time of writing, a "reforecast" for 2017 is being run with the coupling system. It does not include WW3 because of a pair of bugs in WW3 which we have only recently identified and corrected. New tests of the system with WW3 are stable, and so a reforecast which includes WW3 is imminent. Since the coupling to WW3 is one-way at present, the inclusion of WW3 is non-critical: running WW3 post facto should yield identical results. However, inclusion is meaningful as an exercise of the next-generation forecasting system. Resolutions of WW3 within the future system are as follows: short-term deterministic with IRI-1/8, long-term deterministic with IRI-1/4, and short-term ensemble (probabilistic) with IRI-1/4 (Metzger et al. 2014b). The IRI-1/8 (short term deterministic) model is obviously not expected to run quickly. The general guidelines are that it should run in under 60 minutes per model day, using 200 to 500 cores[5]. This

[5] Alan Wallcraft, personal communication

appears achievable, since our 1-day simulation takes 26 minutes using 512 cores (Table 7). The reader is referred to Metzger et al. (2014b) for further information: the resolutions quoted here are unchanged from the early plans stated in Table 1 of that report.

As already stated, this system (in the context of WW3) is entirely new and is not intended to replace any existing system. This system does not prioritize fast delivery. In the context of the short-range deterministic, it emphasizes resolution. In the context of the one-way coupling, it prioritizes the addition of surface currents to the forcing, and exploitation of the expected higher accuracy of the traditional inputs (10-meter winds in particular). In the context of two-way coupling of the system in the longer term future, it prioritizes the potential benefits to other models by including additional physical feedbacks, e.g., the dependence of atmospheric drag on wave age, and the enhanced mixing associated with vertical shear in Stokes drift.

One criticism of the existing systems is the redundancy. Since the number of systems remains unchanged at two, there will still be some redundancy. However, redundancy will be reduced if the new systems have:

1. Common grid design (IRI-1/4 and IRI-1/8). In the case of IRI-1/4, the mask files, obstruction grids, and bathymetry are identical between the two systems. IRI-1/8 will be used only with ESPC, but at least the two grid designs share common features and creation methods.
2. Common compile-time settings. These are known as "switches" and determine the pre-processing of the code based on selection of things such as source term package (STx), ice dissipation routine (ICx), and propagation scheme (PRx). We can think of no reason why this would not be consistent between the two systems.
3. Common run-time settings. These include the calibration coefficients used in ST4, the parameterization settings used in IC4, and a large number of other settings. Again, we can think of no reason why this would not be consistent between the two systems.
4. Common code versions. Of course, the codes will not be identical if the ESPC code has special modifications for coupling (presently the case), but it is desirable that they at least be built from the same model version.

In short, improvements made to one system can be applied directly to the other system, which improves efficiency and reduces the probability of coding and user errors.

8. Discussion

8.1. Discussion: grid integration

Grid integration in the standalone system refers to the application of the post-processing program ww3_gint, which takes fields (e.g., significant waveheight as a function of longitude and latitude) from the three grids and puts the information onto a common grid. In the coupled modeling system, the procedure must be built into the main executable and it does not exist yet. The present version of ESPC produces separate model- and grid-specific output files in native formats for each component model. Post-processing of WW3 output is performed using ww3_gint, ww3_ounf, and ww3_ounp to produce global netCDF files. In either case, the common grid will have regular spacing, e.g., the IRI-1/4 system may use a 1/4° regular global grid that includes latitudes from 79°S to 89°N. In the context of the export of fields for purposes of coupling, it is preferable that discontinuities at the hand-off between grids be small. Thus far,

we have not noticed any discontinuities, but this does not mean that they cannot exist. Our recommendation is, in the coupled modeling system, to implement a grid integration procedure which smoothly weights the output from the two grids according to the distance from each grid edge, similar to the procedure used to reconcile grids of equal rank in Tolman (2008).

8.2. Discussion: COAMPS forcing vs. NAVGEM forcing

The underlying assumption for the present FNMOC-Stennis multi-grid WW3 system (Rogers et al. 2012) was that a) winds from the regional atmospheric model (COAMPS) are more accurate than winds from the global atmospheric model (NOGAPS at the time, NAVGEM now), and b) that if COAMPS has been implemented for a particular region, then it is has been established as a high-priority region and therefore justifies a higher resolution wave model. Assumption (a) is derived from three factors. First, a higher resolution atmospheric model is less likely to suffer from negative bias caused by under-representation of these gradients in pressure fields. Second, we had our own limited comparisons of the accuracy of the two models, e.g., Rogers (2002), which tended to favor COAMPS. The third is admittedly non-scientific. It is based on the logic that if someone has gone through the trouble to implement COAMPS, then they must have had a good reason: i.e. it must be better than the default alternative (NOGAPS or NAVGEM). Now, in 2018, we must recognize that these arguments are insufficient. First, the NAVGEM resolution is now 0.28°, and though COAMPS resolution is always higher, it is not by a large margin, e.g., 15 km resolution COAMPS implies a difference near factor two. Second, spot checks of skill do not indicate an advantage using COAMPS. For example Rogers et al. (2018) look at a research cruise in the western Arctic Ocean. The report finds that the main wind/wave event of that 5-week study is much more accurately predicted by NAVGEM. We can also use the validation reports produced by FNMOC[6]. We inspected the 10-meter wind speed nowcast verifications for August/September 2018 for eight regional areas corresponding to eight of the nine regional grids used in the existing FNMOC-SSC WW3 multi-grid system. It was found that three showed mixed performance (Europe, northeast Pacific, and Arctic), three showed better performance with COAMPS (equatorial Americas, northwest Pacific, and Hawaii), and two indicated better performance with NAVGEM (northwest Atlantic, central Indian Ocean). There does not now appear to be a compelling argument for using regional COAMPS for WW3 forcing, especially since it implies a delay in the start time of the WW3 model.

Finally, we consider the separate issue of regional priority. It is logical to assume that a region that has been chosen to be covered by a COAMPS grid is likely to have greater operational significance than a region is not. In that case, it may logically follow that this priority justifies the implementation of a higher resolution WW3 grid. However, this argument is undercut by the fact that these regions are already well-covered by high resolution SWAN implementations at FNMOC-SSC. We find no compelling evidence for treating these regions at special resolution within the global multi-grid system. Until such evidence is produced, they can be treated using one-way nesting from the global system to regional WW3 or SWAN grids.

We recognize that this situation is not permanent, so the question can be revisited later. Further, we do not discourage use of regional model atmospheric forcing in context of one-way wave model nesting: this decision should be made on a case-by-case (grid-by-grid) basis.

[6] Users within the Navy domain can find these at https://portal.fnmoc.navy.mil/verify_cgi/ .

8.3. Discussion: Nesting sequence

8.3.1. Grids of non-equal rank

When the two-way nesting feature was created (Tolman (2007, 2008)), it was replacing—or at least being compared with—traditional one-way nesting which was already well-established in the context of wave models. In that traditional approach, the high resolution model is provided with boundary forcing corresponding to the duration of its calculation. For example, if the boundary forcing ends at time t_{end}, the computations do not extend past t_{end}. There was a specific intention by Tolman (2008) to retain this conservative treatment: "wave conditions in lower resolution grids need to be computed before higher resolution grids can be considered". The nesting sequence for two grids of non-equal rank is sequential. We therefore refer to it as the *Sequential Approach*:

1) Compute for outer (low rank) grid, from (t_0) to $(t_0 + \Delta t)$. Regions of the grid that are covered by the high rank grid and far from the edge are optionally omitted from this computation.
2) Use time/space interpolation to determine boundary forcing for inner (high rank) grid.
3) Compute for inner grid, from (t_0) to $(t_0 + \Delta t)$.
4) Use grid-cell averaging from inner grid at $(t_0 + \Delta t)$ and apply this to the outer grid at $(t_0 + \Delta t)$. The averaging is based on rigorous calculation of area overlap. Note that there is no time interpolation: it is just applied at $(t_0 + \Delta t)$, and the low rank grid is allowed to evolve freely at these locations until the next update.
5) Compute for outer (low rank) grid, from $(t_0 + \Delta t)$ to $(t_0 + 2\Delta t)$.
6) Etc.

As already mentioned, the high rank grid has boundary information for the entire period of its computation. By contrast, the outer grid does not have this level of information. It computes over the interval Δt without any update from the high rank grid. It "runs ahead" of the other grid in step-wise fashion. It is reasonable to ask, "What if both grids were allowed to do this?" The sequence would be:

1) Compute for both grids from (t_0) to $(t_0 + \Delta t)$.
2) Use space (not time) interpolation to determine boundary forcing for high rank grid at $(t_0 + \Delta t)$. At the same time, use grid-cell averaging from inner grid at $(t_0 + \Delta t)$ and apply this to the outer grid at $(t_0 + \Delta t)$.
3) Compute for both grids, from $(t_0 + \Delta t)$ to $(t_0 + 2\Delta t)$.
4) Etc.

The latter *Parallel Approach* has a computational advantage in context of large problems. Since the high rank and low rank grids are solved at the same time, they are using fewer processes on a per-grid basis, so scaling is improved. Good load balancing is required for this approach. Also relevant to large problems: WW3 has a practical limit on the number of processes per grid, corresponding to the number of spectral grid points, typically from 1000 to 1300. In the most extreme (or futuristic) cases, the *Parallel Approach* reduces the likelihood of hitting this limit.

There are disadvantages of the *Parallel Approach*. What is lost by allowing the high rank grid to "run ahead" of its boundary forcing? We compare the two approaches with two examples. In both examples, there are two grids of non-equal rank. In the first example, the outer grid has a

global time step of Δt_g=3600 s, the inner grid has a global time step of Δt_g=1800 s, and the model is being solved for the period from t=0 to 3600 s.

A) *Sequential Approach.* High rank grid has its boundary forcing updated at…
   a. …t=0 using information from low rank grid at t=0.
   b. …t=1800 s using temporal interpolation from low rank grid, t=0 and t=3600 s.
   c. …t=3600 s using information from low rank grid at t=3600 s.

B) *Parallel Approach.* High rank grid is updated at…
   a. …t=0 using information from low rank grid at t=0.
   b. …t=1800 s is using information from low rank grid last updated at t=0.
   c. …t=3600 s using information from low rank grid at t=3600 s.

In the second example, both grids have Δt_g=3600 s, similar to the IRI design.

A) *Sequential Approach.* High rank grid is updated at t=0 and t=3600 s using information from low rank grid at t=0 and t=3600 s respectively.
B) *Parallel Approach.* Identical to (A).

Thus, we see that for the typical scenario where the high rank grid has a smaller Δt_g than the low rank grid, there is some disadvantage to the *Parallel Approach*, due to non-update of boundary forcing at some time steps, and the fraction of non-updated time steps corresponds to the ratio between the time step sizes, $\Delta t_{g,1}/\Delta t_{g,2}$. In the case of $\Delta t_{g,1} = \Delta t_{g,2}$ (the IRI design), there is no disadvantage apparent with the *Parallel Approach.* In fact, it would be interesting to test whether results are bitwise identical.

Only the *Sequential Approach* is taken in WW3 at time of writing. We recommend that the *Parallel Approach* be implemented as an option.

8.3.2. Grids of equal rank

Grids of equal rank are solved simultaneously, so they already enjoy the advantages for large problems listed above (most importantly: improved scaling). Equal rank grids provide information to each other using spatial interpolation in both directions: the more rigorous method using fraction of area overlap is not used. The grids are reconciled at synchronization intervals using averaging that is (appropriately) based on distance from the edge of the overlapping grids. At time of writing, WW3 is not coded to permit equal rank grids if one or more of the grids are irregular. Therefore the method cannot be applied to the IRI problem.

9. Summary and conclusions

The key points about the IRI grid design are as follows:

- The design achieves efficiency by minimizing the variability of grid spacing, which allows a larger propagation time step size. Variability of grid spacing is minimized by combining two grid types (polar stereographic and regular). Each grid is used to represent the latitude range for which it has the most uniform spacing.
- The design allows computations at high latitudes, unlike the "regular grid" alternatives.
- The design allows computation of propagation with third order accuracy, while the single tripole grid alternative is limited to first order propagation, using the present version of WW3.

This report is primarily concerned with *description* of the IRI design, *discussion* of design decisions, and *review* of relevant concepts. However, some findings are made via experimentation and analysis:

- The IRI grid design is more efficient (faster, at given resolution) than either single-grid alternative (tripole or regular).
- Verification, in the form of propagation tests where energy travels from grid to grid do not indicate any problems.
- Validation, in the form of hindcasts and comparison to observations also do not indicate any problems. We show that error metrics are within the expected range.
- All WW3 global grid designs (IRI or otherwise) are slower than expected on the DSRC system (Crays). This is still under investigation. We expect that improvements can be made on the timings (on the Crays) made in this report.

Combining the benefits known a priori (high latitudes, third order propagation) and the findings demonstrated here (faster computations), the content of this report suggests that it would be appropriate to use the IRI grids for operational systems, either as standalone or fully-coupled systems (Section 7.2).

Finally, two recommendations for the WW3 code are as follows:

1. For non-equal rank grids, the code should be extended to perform simultaneous computations (Section 8.3.1). This can be a non-default user option. It would improve scalability for large problems.
2. Alternatively, for equal rank grids, the code could be extended to allow irregular grids (Section 8.3.2). This would also improve scalability for the larger IRI grids. In the context of this scalability, this code enhancement is redundant with (1).

10. Glossary

APL	Applied Physics Laboratory at the University of Washington
CC	correlation coefficient
CDIP	Coastal Data Information Program. This has multiple functions, but the function most relevant here: CDIP deploys and maintains wave buoys and wave gages. Data is provided on a free website. CDIP is associated with the state of California. Most (but not all) CDIP buoys are near California and most are buoys manufactured by Datawell Inc.
CFL	Courant–Friedrichs–Lewy, usually referring to a limit on numerical stability, e.g., in context of maximum stable time step.
CFSR	Climate Forecast System Reanalysis, Saha et al. (2010)
COAMPS®	Coupled Ocean/Atmosphere Mesoscale Prediction System, The Navy's regional atmospheric model (Hodur 1997), or the Navy's regional coupling system (Allard et al. 2014).
DIA	Discrete Interaction Approximation. Hasselmann et al. (1985). This is the fast but not particularly accurate method of computing the source term for four-wave nonlinear interactions in models such as WAM, SWAN, and WW3. More accurate solvers are available in these models, but are prohibitively expensive, so DIA is the only one being used operationally at present.
DSPWH	Days simulated per wallclock hour. A metric for computation time. It must be taken in context with information about the number of nodes and CPUs used.
DSRC	DoD Supercomputing Resource Center.
ECMWF	European Centre for Medium-Range Weather Forecasts.

ESPC	Earth System Prediction Capability.
FNMOC	Fleet Numerical Meteorology and Oceanography Center. Historically, a large majority of FNMOC is at Monterey, California. Recently however, the modeling group of NAVOCEANO (Stennis Space Center, Mississippi) has become part of FNMOC. Therefore, we now often specify "FNMOC-Monterey" or "FNMOC-Stennis". Alternatives are "FNMOC-MRY" and "FNMOC-SSC" (shorter to type, but not shorter to say aloud).
HYCOM	Hybrid Coordinate Ocean Model. Chassignet et al. (2003).
GOFS	Global Ocean Forecasting System, e.g., Metzger et al. (2009) and Posey et al. (2015).
IC0	Designation in WW3 for the primitive treatment of ice, Tolman (2003). Ice is treated as open water, or land, as porous land (i.e. via partial blocking), depending on whether ice concentration is low, high, or intermediate (respectively).
IC4	Designation in WW3 for several parametric and empirical representations of dissipation of wave energy by sea ice using a source term, Collins and Rogers (2017).
IRI	Irregular-Regular-Irregular. The WW3 multi-grid system with two polar stereographic grids and one regular grid, described by this report.
JCOMM	Joint Technical Commission for Oceanography and Marine Meteorology
MPI	Message Passing Interface, a method for parallelization of code, using in SWAN and WW3.
MRY	Monterey.
NAVGEM	Navy Global Environmental Model. Hogan et al. (2014). This is the Navy's global atmospheric model. This model replaced the older NOGAPS model.
NAVOCEANO	Naval Oceanographic Office, located at Stennis Space Center, MS.
NCODA	Navy Coupled Ocean Data Assimilation, Cummings and Smedstad (2013).
NDBC	National Data Buoy Center, located at Stennis Space Center, MS.
NOGAPS	Navy Operational Global Atmospheric Prediction System, Hogan and Rosmond (1991). This is the predecessor of NAVGEM.
PR1	Designation in WW3 for the first order geographic propagation method (upwind explicit)
PR3	Designation in WW3 for the higher order geographic scheme used in WW3. It is known as QUICKEST (e.g., Leonard 1979). See description in WW3DG (2016). For one-dimensional propagation, it has third order accuracy.
PS	Polar Stereographic.
PSNAS	Polar Stereographic North and South. This is our older name for the IRI system.
nbias	normalized bias
NOAA	National Oceanic and Atmospheric Administration of the National Weather Service (NWS)
NCEP	National Centers for Environmental Prediction (within NOAA). Responsibilities of this agency include the development and operation of NOAA's global atmospheric and wave models.
NRL	Naval Research Laboratory. A detachment of NRL at Stennis Space Center, Mississippi includes the Oceanography Division.
NWW3	NOAA WAVEWATCH III. This label is used by Tolman et al. (2002) to describe a forecasting system.
OS	Ocean Station
RMSE	Root-mean-square error
SCRIP	Spherical Coordinate Remapping and Interpolation Package, remapping software created by Jones (1999).
SI	Scatter Index
SMC	Spherical Multiple-Cell Grid. The unstructured grid, global wave modeling approach of Li (2009, 2011, 2012) which was introduced to WW3 in Li and Saulter (2014)
SSC	Stennis Space Center
SWAN	Simulating WAves Nearshore, Booij et al. (1999).
tripole	An irregular global grid with two poles in the northern hemisphere (both displaced from the north pole such that they are over land), plus the customary south pole. It is sometimes referred to as a "bipolar" grid, referring to the two poles in the northern hemisphere.
UW	University of Washington.

WAM	WAve Model, WAMDIG (1988). The original "third generation" wave model which is distinguished from earlier models by the fact that it does not make a priori assumptions about spectral shape, but instead relies on source terms as the primary control on spectral shape.
WAVEWATCH III	According to Alves and Banner (2003), the acronym was something like "WAVE height, WATer depth, and Current Hindcasting". If true, this usage is deprecated.
WMPWS	wallclock minutes per week simulated, inversely proportional to DSPWH
WW3	WAVEWATCH III®
ww3_gint	A post-processing program within WW3. Here, "gint" stands for "grid integration".
ww3_grid	A pre-processing program within WW3. This processes user-provided information, including grid parameters, and creates a model definition file which is used by subsequent WW3 programs.
ww3_ounf	A post-processing program within WW3. Here, "ounf" stands for "output netCDF fields."
ww3_ounp	A post-processing program within WW3. Here, "ounp" stands for "output netCDF points."

Acknowledgments

Section 5.2 of this report includes a sample of statistics provided by Yalin Fan and James Dykes, both of NRL Code 7322. Complete results from their validation will be presented in a forthcoming report led by Dr. Fan. We thank them for making their results available early.

We also thank Dr. Arun Chawla (NOAA/NCEP) for adapting his "gridgen" software (used for creating obstruction grids for WW3) so that it works with irregular grids. The first obstruction grid created for our polar stereographic grids was actually provided to us by Dr. Chawla, and we used his software to create subsequent grids.

This work was funded as part of the Earth System Prediction Capability project funded by the Office of Naval Research under program element 0603207N and managed by Daniel Eleuterio.

This is NRL contribution number NRL/MR/7320-18-9835 and is approved for public release.

References

Allard, R., E. Rogers, P. Martin, T. Jensen, P. Chu, T. Campbell, J. Dykes, T. Smith, J. Choi, and U. Gravois, 2014: The US Navy Coupled Ocean-Wave Prediction System. *Oceanography,* **27** (3), 92-103, doi:10.5670/oceanog.2014.71.

Alves, J. H. G. M., and M. L. Banner, 2003: Performance of a saturation-based dissipation-rate source term in modeling the fetch-limited evolution of wind waves. *J. Phys. Oceanogr.*, **33**, 1274-1298.

Ardhuin, F., W. E. Rogers, A. Babanin, J.-F. Filipot, R. Magne, A. Roland, A. van der Westhuysen, P. Queffeulou, J.-M. Lefevre, L. Aouf, and F. Collard, 2010: Semiempirical Dissipation Source Functions for Ocean Waves Part I: Definition, Calibration, and Validation. *J. Phys. Oceanogr.,* **40**, 1917-1941.

Ardhuin, F., P. Sutherland, M. Doble, and P. Wadhams, 2016: Ocean waves across the Arctic: Attenuation due to dissipation dominates over scattering for periods longer than 19 s. *Geophys. Res. Lett.,* **43**, 5775–5783, doi:10.1002/2016GL068204.

Barron, C. N., A. B. Kara, P. J. Martin, R. C. Rhodes and L. F. Smedstad, 2006: Formulation, implementation and examination of vertical coordinate choices in the Global Navy Coastal Ocean Model (NCOM). *Ocean Modelling*, **11**, 347-375

Barron, C. N., A. B. Kara, R. C. Rhodes, C. Rowley and L. F. Smedstad, 2007: Validation Test Report for the 1/8 Global Navy Coastal Ocean Model Nowcast/Forecast System. *NRL Tech Report NRL/MR/7320--07-9019*, 144 pp.

Bidlot, J.R. and M.W. Holt, 2006: Verification of Operational Global and Regional Wave Forecasting Systems against Measurements from Moored buoys. *JCOMM Technical Report No. 30*.

Booij, N., R. C. Ris, and L. H. Holthuijsen, 1999: A third-generation wave model for coastal regions, Part 1: Model description and validation. *J. Geophys. Res.,* **104** (C4), 7649-7666.

Chassignet, E. P., L. T. Smith, G. R. Halliwell, and R. Bleck, 2003: North Atlantic simulations with the HYbrid Coordinate Ocean Model (HYCOM): Impact of the vertical coordinate choice, reference pressure, and thermobaricity. *J. Phys. Oceanogr.*, **33**, 2504–2526, doi: 10.1175/1520-0485.

Collins, C.O. and W. E. Rogers, 2017: A Source Term for Wave Attenuation by Sea ice in WAVEWATCH III ®: IC4. *NRL Report NRL/MR/7320--17-9726*, 25 pp.

Cummings, J. A., and O. M. Smedstad, 2013: Variational data assimilation for the global ocean. pp. 303–343 in *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications* (Vol. II). S.K. Park and L. Xu, eds, Springer-Verlag, Berlin Heidelberg, doi: 10.1007/978-3-642- 35088-7_13.

Hasselmann, S., K. Hasselmann, J. H. Allender, and T. P. Barnett, 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.

Hodur, R. M., 1997: The Naval Research Laboratory's Coupled Ocean/Atmospheric Mesoscale Prediction System (COAMPS). *Mon. Wea. Rev.*, **125**, 1414-1430.

Hogan, T. F. and Rosmond, T. E., 1991: The description of the U.S. Navy Operational Global Atmospheric Prediction System's spectral forecast models. *Mon. Wea. Rev.,* **119**, 1786-1815.

Hogan, T., and 16 Coauthors, 2014: The Navy Global Environmental Model. *Oceanogr.,* **27**(3), 116-125.

Jensen, R. E., P. A. Wittmann, and J. D. Dykes, 2002: Global and regional wave modeling activities. *Oceanogr.,* **15**, 57-66.

Jones, P. W., 1999: First- and second-order conservative remapping schemes for grids in spherical coordinates. *Monthly Weather Rev.*, **127**, 2204-2210.

Leonard, B. P., 1979: A stable and accurate convective modeling procedure based on quadratic upstream interpolation. *Computer Methods in Applied Mechanics and Engineering,* **19**, 59-98.

Li, J.-G., 2009: Spherical Multiple-Cell Grid to Include the Arctic in Global Ocean Wave Model. *Proceedings of the 11th International Workshop on Wave Hindcasting and Foreecasting and 2nd Coastal Hazards Symposium, Halifax, Nova Scotia, Canada, 18-23 October 2009*, 15 pp.

Li, J.-G., 2011: Global transport on a spherical multiple-cell grid. *Monthly Weather Rev.*, **139**, 1536-1555.

Li, J.-G., 2012: Propagation of ocean surface waves on a spherical multiple-cell grid. *J. Comput. Physics*, **231**(24), 8,262–8,277.

Li, J.-G., and A. Saulter, 2014: Unified global and regional wave model on a multi-resolution grid. *Ocean Dynamics*, **64**, 1657-1670.

Liu, Q., W. E. Rogers, A. V. Babanin, I. R. Young, L. Romero, S. Zieger, F. Qiao, C. Guan, 2018/2019: Observation-based source terms in the third-generation wave model WAVEWATCH III: updates and verification. (Submitted to *J. Phys. Oceanogr.*)

Metzger, E. J., O. M. Smedstad and S. N. Carroll, 2009: User's Manual for the Global Ocean Forecast System (GOFS) Version 3.0. *NRL Report NRL/MR/7320--09-9175*.

Metzger, E. J., O. M. Smedstad, P. G. Thoppil, H. E. Hurlburt, J. A. Cummings, A. J. Wallcraft, L. Zamudio, D. S. Franklin, P. G. Posey, M. W. Phelps, P. J. Hogan, F. L. Bub and C. J. Dehaan, 2014a: U.S. Navy Operational Global Ocean and Arctic Ice Prediction Systems. *Oceanography*, **27** (3), doi: 10.5670/oceanog.2014.66.

Metzger, E. J., J. D. Dykes, A. J. Wallcraft, L. F. Smedstad, B. C. Ruston, T. R. Whitcomb, S. Chen and J. Chen, 2014b: Operational Implementation Design for the Earth System Prediction Capability (ESPC): A First Look. *NRL Report* NRL/MR/7320--14-9498, 27 pp.

Murray, R.J., 1996: Explicit generation of orthogonal grids for ocean models. *J. of Comp. Phys.*, **126**, 251–273, doi: 10.1006/jcph.1996.0136.

O'Reilly, W.C., T. H. C. Herbers, R. J. Seymour, and R. T. Guza, 1996: A comparison of directional buoy and fixed platform measurements of Pacific swell. *J. Atmos. Oceanic Technol.* **13**, 231-238.

Posey, P. G., E. J. Metzger, A. J. Wallcraft, D. A. Hebert, R. A. Allard, O. M. Smedstad, M. W. Phelps, F. Fetter, J. S. Stewart, W. N. Meier, and S. R. Helfrich, 2015: Improving Arctic sea ice edge forecasts by assimilating high horizontal resolution sea ice concentration data into the US Navy's ice forecast systems. *The Cryosphere,* **9**, 1735-1745.

Rogers, W. E., 2002: The U.S. Navy's Global Wind-Wave Models: An Investigation into Sources of Errors in Low-Frequency Energy Predictions. *NRL Formal Report 7320-02-10035*, 63pp.

Rogers, W. E., and T. J. Campbell, 2009: Implementation of Curvilinear Coordinate System in the WAVEWATCH III Model. *NRL Memorandum Report: NRL/MR/7320-09-9193*, 42 pp.

Rogers, W. E., and M. D. Orzech, 2013: Implementation and Testing of Ice and Mud Source Functions in WAVEWATCH III®. *NRL Memorandum Report, NRL/MR/7320-13-9462*, 31 pp.

Rogers, W. E., Van Vledder, G. Ph., 2013: Frequency width in predictions of windsea spectra and the role of the nonlinear solver. *Ocean Modelling*, **70**, 52-61, dx.doi.org/10.1016/j.ocemod.2012.11.010.

Rogers, W. E., A. V. Babanin, D. W. Wang, 2012: Observation-Consistent Input and Whitecapping Dissipation in a Model for Wind-Generated Surface Waves: Description and Simple Calculations. *J. Atmos. Oceanic Tech.,* **29***(9), 1329-1346.*

Rogers, W. E., J. D. Dykes, and P. A. Wittmann, 2014: US Navy Global and Regional Wave Modeling. *Oceanogr.*, **27**(3):56-67, doi: 10.5670/oceanog.2014.68.

Rogers, W.E., P. Posey, L. Li, R. A. Allard, 2018: Forecasting and hindcasting waves in and near the marginal ice zone: Wave modeling and the ONR "Sea State" field experiment. *NRL Report NRL/MR/7320--18-9786*, 179 pp. [available from www7320.nrlssc.navy.mil/pubs.php].

Saha, S., and 51 co-authors, 2010: The NCEP climate forecast system reanalysis. *Bull. Am. Meteorol. Soc.*, **91**, 1015–1057.

Stammerjohn, S., R. Massom, D. Rind, and D. Martinson, 2012: Regions of rapid sea ice change: An inter- hemispheric seasonal comparison. *Geophys. Res. Lett.,* **39**, L06501, doi:10.1029/2012GL050874.

Stopa, J. E., F. Ardhuin, A. Babanin, and S. Zieger, 2016: Comparison and validation of physical wave parameterizations in spectral wave models. *Ocean Modelling*, **103**, 2–17, doi: 10.1016/j.ocemod.2015.09.003

Thomson, J., and W. E. Rogers, 2014: Swell and sea in the emerging Arctic Ocean. *Geophys. Res. Lett.*, **41**, doi:10.1002/2014GL059983.

Thomson, J., J. Talbert, A. de Klerk, A. Brown, M. Schwendeman, J. Goldsmith, J. Thomas, C. Olfe, G. Cameron, C. Meinig, 2015: Biofouling effects on the response of a wave measurement buoy in deep water. *J. Atm. Ocean. Tech.*, **32**, 1281-1286, doi: 10.1175/JTECH-D-15-0029.1.

Thomson, T., Y. Fan, S. Stammerjohn, J. Stopa, W. E. Rogers, F. Girard-Ardhuin, F. Ardhuin, H. Shen, W. Perrie, H. Shen, S. Ackley, A. Babanin, Q. Liu, P. Guest, T. Maksym, P. Wadhams, C. Fairall, O. Persson, M. Doble, H. Graber, B. Lund, V. Squire, 2016: Emerging trends in the sea state of the Beaufort and Chukchi seas. *Ocean Modelling*, **105**, doi:10.1016/j.ocemod.2016.02.009.

Tolman, H. L., 1991: A Third-generation model for wind-waves on slowly varying, unsteady, and inhomogeneous depths and currents. *J. Phys. Oceanogr.,* **21**(6*),* 782-797.

Tolman, H. L., 2002: User manual and system documentation of WAVEWATCH-III version 2.22. *Tech. Note 222, NOAA/NWS/NCEP/MMAB*, 133 pp.

Tolman, H. L., 2003: Treatment of unresolved islands and ice in wind wave models. *Ocean Modelling,* **5**, 219-231.

Tolman, H. L., 2007: Development of a multi-grid version of WAVEWATCH III. *Tech. Note 256, NOAA/NWS/NCEP/MMAB*, 88 pp. + Appendices.

Tolman, H. L., 2008: A mosaic approach to wind wave modeling. *Ocean Modelling*, **25**, 35-47.

Tolman, H. L., 2009: User Manual and System Documentation of WAVEWATCH III™ Version 3.14, *Tech. Note, NOAA/NWS/NCEP/MMAB,* 220 pp.

Tolman, H. L. and the WAVEWATCH III® Development Group, 2014: *User Manual and System Documentation of WAVEWATCH III® version 4.18*, *Tech. Note 316, NOAA/NWS/NCEP/MMAB*, 282 pp. + Appendices.

WAMDIG, 1988: The WAM model – a third generation ocean wave prediction model. *J. Phys. Oceanogr.*, **18**, 1775-1810.

Wittmann, P. A., 2002: Implementation of WaveWatch III at Fleet Numerical Meteorology and Oceanography Center. *Proceedings: MTS/IEEE: Conference and Exposition: An Ocean Odyssey: November 5-8, 2001, Honolulu, Hawaii*, 1474-1479.

The WAVEWATCH III® Development Group (WW3DG), 2016: User manual and system documentation of WAVEWATCH III® version 5.16. *Tech. Note 329, NOAA/NWS/NCEP/MMAB, College Park, MD, USA*, 326 pp. + Appendices.

Zieger, S., A. V. Babanin, W. E. Rogers and I. R. Young, 2015: Observation-based source terms in the third-generation wave model WAVEWATCH. *Ocean Modelling*, **96**, doi:10.1016/j.ocemod.2015.07.014