AMERICAN
METEOROLOGICAL
SOCIETY

Weather and Forecasting

EARLY ONLINE RELEASE

This is a preliminary PDF of the author-produced manuscript that has been peer-reviewed and accepted for publication. Since it is being posted so soon after acceptance, it has not yet been copyedited, formatted, or processed by AMS Publications. This preliminary version of the manuscript may be downloaded, distributed, and cited, but please be aware that there will be visual differences and possibly some content differences between this version and the final published version.

The DOI for this manuscript is doi: 10.1175/WAF-D-15-0093.1

The final published version of this manuscript will replace the preliminary version at the above DOI once it is available.

If you would like to cite this EOR in a separate work, please use the following full citation:

Sampson, C., J. Hansen, P. Wittmann, J. Knaff, and A. Schumacher, 2016: Wave probabilities consistent with official tropical cyclone forecasts. Wea. Forecasting. doi:10.1175/WAF-D-15-0093.1, in press.

© 2016 American Meteorological Society

Manuscript (non-LaTeX)
Click here to download Manuscript (non-LaTeX)
WW3_ensemble_manuscript_with_evaluation_rev2.4.docx

Wave probabilities consistent with official tropical cyclone forecasts

Charles R. Sampson and James A. Hansen, NRL Monterey
Paul A. Wittmann, FNMOC
John A. Knaff, NOAA/NESDIS
Andrea Schumacher, CIRA

July 2016

Weather and Forecasting

(NCEP Notes)

Corresponding author:
Charles R. Sampson
7 Grace Hopper Ave
NRL Monterey, CA, USA 93943
831 656-4714
Buck.sampson@nrlmry.navy.mil

PRELIMINARY ACCEPTED VERSION

ABSTRACT

Development of a 12-ft seas significant wave height ensemble consistent with the official tropical cyclone intensity, track and wind structure forecasts and their errors from the operational U.S. tropical cyclone forecast centers is described. To generate the significant wave height ensemble, a Monte Carlo wind speed probability algorithm that produces forecast ensemble members is used. These forecast ensemble members, each created from the official forecast and randomly sampled errors from historical official forecast errors, are then created immediately after the official forecast is completed. Of 1000 forecast ensemble members produced by the wind speed algorithm, 128 of them are selected and processed to produce wind input for an ocean surface wave model. The wave model is then run once per realization to produce 128 possible forecasts of significant wave height. Probabilities of significant wave height at critical thresholds can then be computed from the ocean surface wave model-generated significant wave heights.

Evaluations of the ensemble are provided in terms of maximum significant wave height and radius of 12-ft significant wave height, two parameters of interest to both U.S. Navy meteorologists and U. S. Navy operators. Ensemble mean errors and biases of maximum significant wave height and radius of 12-ft significant wave height are found to be similar to those of a deterministic version of the same algorithm. Ensemble spreads capture most verifying maximum and radii of 12-ft significant wave heights.

1. Introduction

Tropical cyclones (TCs) have received great attention in the U.S. Navy since Halsey's Fleet encountered an intense TC on December 18, 1944 while some of the ships were refueling. The event resulted in the loss of 780 lives, three destroyers sunk, and 146 aircraft damaged beyond repair or destroyed. Eighty sailors suffered injuries and many of the other ships in Task Force 38 were damaged by the high seas and hurricane force winds (U.S. Navy, cited 2016). One of the legacies of this disaster was the creation of weather facilities in the western North Pacific and eventually the creation of the Joint Typhoon Warning Center (JTWC). Since ship safety and performance is highly dependent on sea state, wave models that produce forecasts of significant wave height and swell are of great interest to the Navy. Forecasting the state of the sea is also of great importance to commercial shipping, offshore oil/gas operations, and recreational boating, to name a few additional interests.

Traditionally, third generation spectral ocean wave models such as WAVEWATCH III ® (Tolman 1991, Tolman et al., 2002) are run with global Numerical Weather Prediction (NWP) model surface winds to produce significant wave height forecasts. However, the resolution of the global NWP models is often insufficient to capture the steep wind gradients associated with TCs, so there are applications to infuse more detailed TC structure from higher resolution NWP models into the global NWP model wind fields used as input to WAVEWATCH III (e.g., Tolman et al. 2005), which can result in more realistic wave fields in the vicinity of the TC (e.g., Chao and Tolman 2010).

One disadvantage to using exclusively NWP model winds is that they are inconsistent with the official forecasts from the operational centers (e.g., JTWC). To

alleviate this shortcoming, Sampson et al. (2010) implemented an algorithm to use the official forecast (track, intensity and wind structure) placed within a global NWP model's output winds as input to WAVEWATCH III. The winds used in WAVEWATCH III would have the track and structure of the official forecasts and use the global NWP model winds (with the NWP model TC vortex removed) only as background winds. An objective evaluation against the National Hurricane Center (NHC) Tropical Analysis and Forecast Branch (TAFB) real-time estimates of significant wave heights indicates that using the NHC forecast information produced forecasts of maximum significant wave heights and 12-ft[1] (1 ft = .3048 m) seas radii that were generally less biased and more accurate relative to forecaster estimates than using global model winds alone (Sampson et al. 2013). More important to Navy operations is that the forecast area of 12-ft seas is geographically consistent with the official TC forecast. The algorithm, currently named WW3_TC_OFCL, was subsequently implemented operationally at Fleet Numerical Meteorology and Oceanography Center (FNMOC).

One possible extension to producing 12-ft seas consistent with official forecasts is to produce 12-ft seas probabilities that are consistent with official TC wind speed probabilities (WSPs hereafter, DeMaria et al. 2007; DeMaria et al. 2013) disseminated through the operational TC forecast centers. The WSPs are designed to create a forecast ensemble that reflects the operational location and intensity errors and climatological wind structure errors. These would then provide 12-ft seas probabilities that could be used in conjunction with the wind speed probabilities in ship sortie and

[1] Throughout this manuscript we use imperial units instead of SI units because the application is designed for use in U.S. maritime operations, which is still in the habit of using imperial units. We also use the term "12-ft seas" interchangeably with "significant wave heights of 12 ft."

ship routing decisions for the entire Northern Hemisphere. Such forecasts would also be consistent with guidance for the Department of Defense's TC Conditions of Readiness that is also based on the WSPs (Sampson et al. 2012).

Impetus for using probabilities in forecasts can be shown using an example. Figure 1 shows a deterministic 48-h TAFB forecast for Joaquin from September 29, 2015 at 00 UTC. At this time the S. S. El Faro, a steam turbine container ship 790 ft long capable of a speed of 24 kt departed Jacksonville for San Juan Puerto Rico. El Faro never made it to Puerto Rico as planned, but was sunk near Crooked Island Passage on October 1, 2015. Navigating solely with the TAFB forecast in Figure 1 (and assuming the forecast is 100% accurate), a ship could hypothetically take a route with "fair winds and following seas" just north of the Bahamas to Puerto Rico. The TAFB forecasted maximum significant wave heights associated with Joaquin are far removed from the El Faro route with a maximum significant wave height of 18 ft. In retrospect, this forecast had large errors in both track and intensity relative to the averages (NHC, cited 2015). The TAFB analyzed TC position at 00 UTC on October 1 is directly in the path of El Faro with an intensity of 100 kt and estimated maximum significant wave heights of 28 ft along El Faro's route. The forecast errors were higher than average, so the case is a good one to use as a case study for ensembles such as the WSPs that are intended to capture large errors in the official forecast.

The purpose of this work is to describe a wave forecast ensemble that is consistent with the forecast and errors of official TC forecast centers. The algorithm to generate significant wave height probabilities is described in section 2, the dataset is discussed in

section 3, verification of the input data and algorithm output is shown in section 4, conclusions and operational considerations are discussed in section 5.

2. The algorithm

Forecast ensemble members (128 randomly selected from the original 1000, which is a number that will be explained later in this work) are generated using the WSP algorithm described in DeMaria et al. (2013). Each of the 128 ensemble members is made available to the WW3_TC_OFCL algorithm independently. The ensemble member is essentially the same as an official forecast defined at 0, 12, 24, 36, 48, 72, 96, and 120 h. A series of hourly forecasts is generated and those hourly vortex forecasts are then converted to high resolution hourly storm-scale gridded fields using the tessalation routine from O'Reilly and Guza (1993), which will later be inserted into the NWP model surface wind fields. The NWP model surface wind fields are preprocessed by removing the NWP model vortex, which is likely geographical displaced from and structurally different than the ensemble member vortex. The NWP model vortex removal process eliminates the vortex using forecast information produced by the National Centers for Environmental Prediction (NCEP) vortex tracker (Marchok 2002) and replaces the removed area with bilinearly interpolated data from the sides of the removed area. For our NWP model surface wind fields we used the Navy Global Environmental Model (NAVGEM, Hogan et al. 2015) operational 1 degree resolution 10 m winds. The final step of the gridded surface wind processing is inserting the hourly storm-scale gridded fields into the NWP model surface winds at the prescribed resolution of the NWP model. The resultant set of gridded surface wind field forecasts at 1-h forecast intervals serve

as input in WAVEWATCH III (version 2.2.3). The WAVEWATCH III is then run on each of the 128 members, producing significant wave height grids. The version of WAVEWATCH III used in the ensemble is the same as used in Sampson et al. (2013) except the grid resolution is 0.4° instead of the 0.2° used in the deterministic version of OFCL/WW3 in order to reduce the computation time. Two WAVEWATCH III domains are defined for the ensemble, one for the Atlantic basin (0° to 50°N and 100° to 30°W) and one for the western North Pacific (5°N to 46°N and 100°E to 166°E). There is no input at the boundaries and the model is cold started since it only runs once a day for each domain.

To compute probabilities, the number of ensemble members with significant wave height over a threshold (e.g., 12 ft) are counted at each grid point in the domain, and then divided by the total number of runs (e.g., 128) to produce a field of probabilities above the threshold.

One of the topics of discussion for implementation of the ensemble in operations concerns the number of realizations required to yield reasonable results. The WSP algorithm described in DeMaria et al. (2013) sets the number of realizations at 1000, which yields smooth wind probability fields. Although 1000 realizations (the same number as in the WSP algorithm) would be ideal for our application, we found this number to be untenable for running the WAVEWATCH III since it is resource intensive. Based on computational restraints we surmised that we could run on the order of 100 realizations, but would prefer to run less if we could. DeMaria et al. 2009 estimates that the error introduced by limiting the number of realizations to 100 instead of 1000 is on the order of 1-2%, which is acceptable for our purposes. We also ran a sensitivity study

whereby we compute and plot the cumulative probabilities of both winds and waves for 10 to 120 realizations by increments of 10. We investigated this with a 96-h forecast for Yagi (wp162006) on September 19, 2006 at 12 UTC because that particular TC has some of the attributes we care most about (i.e., it is a TC that accelerates into the westerlies near Japan). As seen in Figure 2, the cumulative probabilities for 34-kt winds are noisy for 120 realizations while the cumulative probabilities of 12-ft seas are much smoother. Based on these results we selected 128 realizations for the wind probabilities, which is a number that fits well with many multi-processor systems, being a multiple of 16 and 64. We chose the WAVEWATCH III because that is the wave model used at FNMOC. Another option for operational centers with less computational resources would be to run a simple model such as done in Lazarus et al. (2013).

3. Data Used In this Study

The TC track and structure information used in this study come from the JTWC and NHC as stored on their Automated Tropical Cyclone Forecast System (ATCF®, Sampson and Schrader 2000) work stations. ATCF storm identifiers are used throughout the manuscript and are in the form bbnnyyyy (e.g., al112015), where bb (e.g., al) is the basin, nn (e.g., 11) is the TC sequential number for the season, and yyyy (e.g., 2015) is the season in which the TC developed.

The significant wave height forecasts were generated from real-time runs of the WAVEWATCH III ensemble at the Naval Research Laboratory (NRL) during the 2013-2015 seasons. These real-time runs were available approximately 12 h late; however, at FNMOC the real-time runs potentially could be available within 3 h of the release of

the official forecast. Due to computational limitations and communications issues between NRL and the operational centers, we were only able to run the ensemble a maximum of once a day and for only one TC in the western North Pacific and one in the Atlantic. The western North Pacific ensemble was usually run at 00 UTC and the North Atlantic ensemble was run at 12 UTC. Although the number of forecasts is limited, there is a great deal of serial independence between the forecasts compared to a data set generated from sequential forecasts (i.e., every six hours).

A common way of evaluating significant wave height is to compare the results of wave model forecasts with altimeter data (see Alves et al., 2013) or buoy observations. This was attempted in Sampson et al. (2013), but the altimeter pass footprint was small and rarely passed over the area where the significant wave heights were greater than 12 ft. We found only 20 passes for our entire dataset, many for the same TC, so we did not attempt this type of evaluation in this work. Instead, we found that the 6-hourly real-time analyses of maximum significant wave height and 12-ft seas radii (i.e., the radii of 12-ft significant wave height) generated by TAFB to be both a convenient and frequently available source of data for ground truth. Sources of data that go into these analyses are buoy reports, ship reports, altimeter passes, and WAVEWATCH III output. The 12-ft-sea radii estimates (in the compass quadrants NE, SE, SW, NW from the center of the TC as defined by NHC) are part of the NHC advisory messages and are stored in the ATCF database. Maximum significant wave heights are not saved in the ATCF database, but are part of the TAFB high-seas forecasts issued every 6 h. If we treat these TAFB analyses of maximum significant wave height and 12-ft seas as ground truth, we can evaluate both the forecasts of maximum seas within and the 12-ft-sea radii

surrounding the TC circulation using a modified version of the NCEP tracker (Marchok 2002) with a maximum radius of 300 nm. These parameters (maximum significant wave height and the radii of 12-ft seas) should provide us with a reasonable evaluation of large waves in the vicinity of TCs. It should be noted that the maximum significant wave height estimates are not necessarily at the center of the TC wind circulation.

Coincidentally, JTWC also provides estimates of maximum significant wave height in their warnings (JTWC warnings are the equivalent of NHC advisories). These estimates are based on altimetry and a wave analysis and forecasting nomogram based on wind speed, duration and fetch (Bretschneider 1970). They provide another independent dataset for evaluation.

4. Results

The results section is divided into four parts: (1) evaluation of the track, wind intensity and wind radii data to provide justification for this algorithms utility, (2) an evaluation of our wave algorithm output of maximum significant wave height near the center of TCs and radii of 12-ft seas against estimates from the NHC and JTWC, and (3) application of our algorithm to the El Faro case.

4.1 Track, wind intensity and wind radii evaluation

Questions frequently arise about the need for an algorithm such as the one we describe above since we already have global NWP model ensembles (e.g., NAVGEM) that generate probabilities of significant wave height, and that those ensembles may be consistent with official forecasts from the TC forecast centers. To answer those

questions, the authors evaluated forecast tracks, intensities and wind radii from the WSP algorithm against those from the NAVGEM and Global Forecast System (GFS; NOAA cited 2016) ensembles since those are critical metrics to discern whether the winds that generate the waves are consistent[2] with the official forecasts. Visual inspection of the tracks, intensities and wind radii reveals that the ensemble forecasts are generally inconsistent with official forecasts. Figure 3 shows an example of track and intensity forecasts from the GFS and NAVGEM ensembles and the 128 members used in our algorithm for Joaquin in the Atlantic on 29 September, 2015 at 00 UTC. By the 72-h forecast time, the 20 member GFS ensemble (Fig. 3 top) clearly has a right of track and negative intensity bias relative to the NHC forecast. The 20 member NAVGEM ensemble also have a negative intensity bias. The WSP ensemble members form an envelope around the NHC forecast for both track and intensity. Even though the NHC forecast is far right of the verifying track with lower intensity than verified, the highest intensity forecast from the WSP ensemble intensifies the TC to approximately 95 kt, much closer to the verifying intensity of 115 kt than the highest intensity forecasts (50 kt) from each of the global NWP model ensembles. The 34-kt wind radii were also inspected for this case, and relatively few 34-kt wind radii (about 25%) are generated by the GFS and NAVGEM ensembles since their intensities are low biased, and the radii range between 50 and 300 nm. The WSP algorithm has more 34-kt wind radii forecasts (about 60%) that range between approximately 50 nm and 350 nm. The official forecast 34-kt wind radii were 50-100 nm so all three ensembles had at least some 34-kt wind radii enveloping the official forecast.

[2] Consistency is preferred for forecasters, but more consistent forecasts don't necessarily mean more accurate forecasts.

To test whether the NAVGEM and GFS ensemble intensity biases were isolated to the El Faro case, we evaluated these ensembles for our dataset. The intensity evaluation indicates that the NAVGEM and GFS ensembles have a 15-20 kt negative intensity biases relative to official forecasts on average, and that the biases become more negative for TCs of 65 kt or greater intensity. For example, the NAVGEM ensemble mean intensity forecast bias for the 2014-2015 western North Pacific TCs is -14 kt at t=0 h (307 cases) and -23.5 kt at 72 h (155 cases) while the JTWC intensity forecast bias is 0.3 kt and 10.7 kt for 0 and 72 h, respectively. The NAVGEM ensemble mean intensity forecast bias for the 2015 western North Pacific TCs verifying with intensities 65 kt or greater is –28.1 kt at t=0 h (156 cases) and -33.1 kt at t=72 h (113 cases) while the JTWC bias is 0 kt and 7.3 kt at 0 and 72 h, respectively. A similar trend is seen in the Atlantic with the GFS ensemble, but with far fewer cases.

We also investigated NAVGEM ensemble wind radii tendencies relative to the JTWC forecasts by verifying two years of NAVGEM ensemble mean wind radii forecasts for the western North Pacific 2014-2015 seasons. These were verified against the JTWC real-time analyzed wind radii (over 500 cases at 72 h), and indicate that the NAVGEM 34-kt forecast wind radii are 50-60 nm larger than the analyzed radii and about 70-90 nm larger than the JTWC forecast wind radii. The NAVGEM ensemble mean 50-kt wind radii are larger by approximately 50-70 nm than both the JTWC analyzed and forecast wind radii. The 128-member WSP ensemble mean radii biases are within 7 nm of the JTWC forecasts and so are consistent with the JTWC forecasts in size estimates.

Finally, we investigated the GFS ensemble wind radii tendencies relative to the NHC forecasts, we verified GFS ensemble mean wind radii forecasts for the Atlantic 2013-

2015 seasons (approximately 80 cases at 72 h) against the NHC wind radii estimates. This evaluation shows that both the WSP and GFS ensemble mean forecast 34-kt wind radii biases are reasonable, between -20 and 20 nm. The GFS ensemble 50-kt forecast wind radii are biased 10 to 20 nm larger than the WSP ensemble mean forecasts, which are in turn larger than the NHC forecasts by 5-15 nm. The probability of detection for WSP ensemble mean forecasts of 34- and 50 –kt radii is 100% at all forecast times (0 – 120 h). For the same forecast times, the GFS ensemble mean forecasts have probability of detection for 34-kt radii ranging from 58% to 97% at with an average of 65%, and have probability of detection for 50-kt wind radii ranging from 80-100% with an average of 84%. This is disconcerting since only one ensemble radius needs to be present to compute the ensemble mean radius, so rates lower than 100% indicate cases where none of the ensemble members intensified the TC above the verifying threshold. This is an undesirable quality for computing wind probabilities because the probabilities will be unrealistically low above these thresholds and was noticed in many cases we investigated.

4.2 Maximum significant wave height and radii of 12-ft seas

As discussed in the data section, the ground truth for evaluation of the maximum significant wave height come from the forecast center products. The maximum significant wave height is the parameter for which we have most data since real-time estimates exist in both NHC and JTWC products. Figure 4 shows an evaluation of the ensemble mean maximum significant wave height and radius of 12-ft seas for the Atlantic 2013-2015 seasons verified against the NHC analyses. Maximum significant

wave height mean errors start at approximately 4 ft in the analysis and grow to approximately 11 ft by 96 h for both the ensemble and deterministic versions of WW3_TC_OFCL, though the number of cases is very small. The errors are approximately 20-40% of the verifying maximum significant wave height, and the biases in the Atlantic are generally small and negative. These means and biases are consistent with those found for the deterministic version in 2010-2011 (Sampson et al. 2013). The 12-ft seas radii errors for both deterministic and ensemble mean range between 58 and 73 nm, which is about 20-35% of the average 12-ft seas radius in NHC analyses. The biases for the 12-ft seas radii have range of 0 to 50 nm for the ensemble and -10 to 10 nm for the deterministic WW3_TC_OFCL.

To test consistency between the ensemble mean and the deterministic versions of WW3_TC_OFCL is to use the deterministic WW3_TC_OFCL analyses in lieu of the TAFB analyses. An evaluation of the maximum significant wave height and 12-ft seas radii against the WW3_TC_OFCL estimates at analysis time is shown in Figure 5. The mean errors and biases of the deterministic WW3_TC_OFCL start at 0 (as they should) and the ensemble mean is low biased (recall that our algorithm is only run every 24 h so it takes time to generate a reasonable sea state for intensifying TCs), but within 24 h of forecast time the deterministic and ensemble means in the top left of Figure 5 become highly correlated. The 12-ft seas radii errors for the ensemble mean tend to be about 10 nm larger than those of the deterministic model, but biases are similar. The only notable difference is between the maximum significant wave height biases at 120 h where the number of data points evaluated is small (n=60). We also evaluated the maximum significant wave height from both the deterministic and ensemble

WW3_TC_OFCL against the estimates in the JTWC warning messages. The errors are about 5 ft higher and the biases 5 ft more negative at all forecast times. It is not obvious whether the JTWC estimates are high biased, or the WW3_TC_OFCL estimates are low biased, or both.

Ensemble spread metrics were estimated for the WW3_TC_OFCL ensemble in both basins (Fig. 6). The maximum distance, selected from all members, of a member from the ensemble mean (dashed line in Fig. 6) increases with forecast time out to 72 h, and then level off. For the Atlantic, this metric, on average, increases from approximately 10 ft at 24 h to over 20 ft at 72 h and beyond. For the western North Pacific, this metric increases from approximately 10 ft at 24 h to over 15 ft at 72 h and beyond. The medians (not shown) are within a few ft of the means and usually slightly higher (1-3 ft). The largest maximum distance of a member from the ensemble mean was found to be 41 ft in the Atlantic and 31 ft in the western North Pacific. Using the maximum significant wave height estimates as ground truth, the hit rate (where the estimated maximum significant wave height lies within the spread of the ensemble) in the Atlantic ranged from 86% to 100% at 24-120 h. In the western North Pacific the hit rate was somewhat lower (70% to 87%). As expected, the mean distance (purple line in Fig. 6) of the ensemble members from the ensemble mean gradually increases to 5-6 ft through the forecast.

4.3 El Faro

Although objective analysis of the ensemble is useful, it is still important to scrutinize individual cases, especially the difficult ones such as the El Faro case described in the

Introduction where the NHC and most NWP model forecast errors were above seasonal averages. Figure 7 shows the WW3_TC_OFCL ensemble forecast for the case discussed in Figure 1. It is encouraging that the maximum significant wave height forecasts by the 128 ensemble members encompass the maximum significant wave height estimated by forecasters during the event. It also highlights the importance of including many members in the ensemble because the first 20 members forecast maximum significant wave height less than observed (the closest is within 5 ft). The 12-ft seas radii forecasts in the ensemble also encompass most of the forecaster estimated radii and provide an indication of probabilities of 12-ft seas greater than 40% along El Faro's approximate route north of the Bahamas. This is in contrast to the deterministic forecast shown in Figure 1, which if assumed to be 100% accurate would allow for passage south of Joaquin with following seas less than 12 ft.

5. Conclusions

We have described an algorithm to produce an ensemble of significant wave heights from forecasts and forecast errors consistent with track, wind and structure forecasts from official forecasts at the U.S. tropical cyclone forecast centers (JTWC, NHC and CPHC). The algorithm was evaluated in terms of maximum significant wave height and radius of 12-ft significant wave height, two parameters of interest to both U.S. Navy meteorologists and U. S. Navy operators. The ensemble mean errors and biases of maximum significant wave height and ensemble mean errors and biases of the radius of 12-ft significant wave height are found to be similar to a deterministic version of the same algorithm. The ensemble spreads also appears to capture a very poorly forecast event, which is essentially what wave ensembles should do.

If implemented in operations, the WW3_TC_OFCL ensemble can be employed to generate probabilities of significant wave heights at critical levels (e.g., greater than 12-ft seas) used by Navy forecasters in applications such as sortie timing and ship routing. The algorithm implemented at FNMOC can process all active TCs at once, and uses NAVGEM ensemble background 10-m wind fields in the WAVEWATCH run on a .4° global band. The entire process to generate significant wave height probabilities on this global band takes about an hour, employing 128 processors (one for each ensemble member). The FNMOC implementation would also solve issues with initial and boundary conditions since it would be run every 6 or 12 hours. There is a significant impact on operational resources, but the wave probability product is consistent with the official forecasts from the U.S. tropical cyclone forecast centers, the deterministic WW3_TC_OFCL product, and the WSP products, so it has value as part of a consistent suite of operational center products.

The WW3_TC_OFCL ensemble would be further enhanced using ensemble background wind fields as discussed in Alves et al. (2013) and would also benefit from improvements in the extremely active topic of NWP model ensemble development. The WW3_TC_OFCL algorithm will also be improved with incremental enhancements to the WSP product. Efforts are currently underway to address a number of known shortcomings in the WSP product as part of the Joint Hurricane Testbed. These include improving the hourly interpolation of track information by replacing linear interpolation with cubic spline interpolation, and bias correcting the Wind Radii CLIPER model. Bias correcting the Wind Radii CLIPER will be accomplished by developing a method to use

all available wind radii (34, 50 and 64 kt) from the NHC forecast to consistently bias correct the Wind Radii CLIPER model, and to use the error serial correlation to extend the influence of the bias correction beyond the time when the NHC radii are available (72 h for 34 and 50 k, and 36 h for 64 kt). This task may be easier now since the official NHC 34-kt wind radii forecasts through 72 hours have been skillful (better than the Wind Radii CLIPER) the last several years (Knaff and Sampson 2015). Work remains to correct Wind Radii CLIPER biases in the western North Pacific, where concerted efforts are currently underway to best track wind radii and improve wind radii forecasts.

With regard to making this product available everywhere, the WSP product is also being extended to the Southern Hemisphere and North Indian Ocean similar to what Brownlee et al. (2013) did for the Australian Bureau of Meteorology.

Acknowledgements

We would like to acknowledge the support of the Office of Naval Research and Commander Naval Meteorology and Oceanography Command. Also, we would like to acknowledge the COLA/IGES for use of the GrADS software and the three anonymous reviewers who made this a better manuscript. We also use software written by Efren Serra, Tim Marchok, Mark DeMaria, Mike Frost and Ann Schrader. Rachel Knaff is acknowledged for processing years of NHC and JTWC messages for us. The views, opinions, and findings contained in this report are those of the authors and should not be construed as an official National Oceanic and Atmospheric Administration or U.S. Government position, policy, or decision.

References

Alves, J.-H., and Coauthors, 2013: The NCEP–FNMOC combined wave ensemble product: Expanding benefits of interagency probabilistic forecasts to the oceanic environment. Bull. Amer. Meteror. Soc., 94, 1893–1905, doi:10.1175/BAMS-D-12-00032.1.

Bretschneider, 1970: Forecasting relations for wave generation. Look Lab., Hawaii, I (3) (1970), pp. 31–34.

Brownlee, J., M. Foley, A. Donaldson, M. DeMaria, and J. Knaff, 2013: Estimating TC wind probabilities in the Australian region. 19th National Conference of the Australian Meteorological and Oceanographic Society (AMOS), Poster Sesion 8, Poster # 8.

Chao, Y.Y., Burroughs, L.D. and Tolman, H.L., 2003. The North Atlantic hurricane wind wave forecasting system (NAH). Technical Procedures Bulletin No, 478.

_____ and H. L. Tolman, 2010: Performance of NCEP regional wave models in predicting peak sea states during the 2005 North Atlantic hurricane season. Weather and Forecasting, 25, 1543-1567.

DeMaria, M., J .A. Knaff, R. Knabb, C. Lauer, C. R. Sampson, and R. T. DeMaria, 2009: A new method for estimating tropical cyclone wind speed probabilities, Wea. Forecasting, 24, 1573-1591.

_____, J. A. Knaff, M. Brennan, D. Brown, C. Lauer, R. T. DeMaria, A. Schumacher, R. D. Knabb, D. P. Roberts, C. R. Sampson, P. Santos, D. Sharp, K. A. Winters, 2013: Operational tropical cyclone wind speed probabilities part I: Recent model improvements and verification, Wea. Forecasting, 28, 586-602. doi: http://dx.doi.org/10.1175/WAF-D-12-00116.1

Falvey, R., cited 2012: Summary of the 2011 Western Pacific/Indian Ocean Tropical Cyclone Season. Presentation to the 66th Interdepartmental Hurricane Conference, Charleston, SC, 5 March 2012. [Available on-line at http://www.ofcm.gov/ihc12/Presentations/01b-Session/05-JTWC_2012_IHC_Final.pdf]

Lazarus, S. M., S. T. Wilson, M. E. Splitt, and G. A. Zarillo, 2013: Evaluation of a Wind-Wave System for Ensemble Tropical Cyclone Wave Forecasting. Part II: Waves. Wea. Forecasting, 28, 316–330. doi: http://dx.doi.org/10.1175/WAF-D-12-00053.1

Hogan, T. F., M. Liu, J.A. Ridout, M.S. Peng, T.R. Whitcomb, B.C. Ruston, C.A. Reynolds, S.D. Eckermann, J.R. Moskaitis, N.L. Baker, J.P. McCormack, K.C. Viner, J.G. McLay, M.K. Flatau, L. Xu, C. Chen, and S.W. Chang, 2014: The Navy Global Environmental Model. Oceanography, 27(3), 116–125, doi:10.5670/oceanog.2014.73.

Knaff, J.A. and C.R. Sampson 2015: After a decade are Atlantic tropical cyclone gale force wind radii forecasts now skillful?, Submitted to *Wea. Forecasting*.

_____, _____, M. DeMaria, T. P. Marchok, J. M. Gross, and C. J. McAdie, 2007: Statistical tropical cyclone wind radii prediction using climatology and persistence, Wea. Forecasting, 22, 781–791.

Marchok, T. P., 2002: How the NCEP Tropical Cyclone Tracker works. Preprints, 25th Conf. on Hurricanes and Tropical Meteorology, San Diego, CA, Amer. Meteor. Soc., P1.13. [Available online at

http://ams.confex.com/ams/pdfpapers/37628.pdf.]

NHC, cited 2016: "2015 National Hurricane Center Forecast Verification Report. NHC annual verification reports." NOAA/NWS/NHC. [Available on-line at 6http://www.nhc.noaa.gov/verification/pdfs/Verification_2015.pdf]

NOAA, cited 2016: "Global Forecast System." [Available on-line at http://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forcast-system-gfs]

O'Reilly, W. C., and R. T. Guza, 1993: A comparison of two spectral wave models in the Southern California Bight. Coastal Eng., 19, 263–282.

Sampson, C. R., and A. J. Schrader, 2000: The automated tropical cyclone forecasting system (Version 3.2). Bull. Amer. Meteor. Soc., 81, 1231-1240.

_____, P. A. Wittmann, and H. L. Tolman, 2010: Consistent tropical cyclone wind and wave forecasts for the U.S. Navy. Wea. Forecasting, 25, 1293-1306.

_____, _____, E. A. Serra, H. L. Tolman, J. Schauer, and T. Marchok, 2013: Evaluation of wave forecasts consistent with tropical cyclone wind forecasts, Wea. Forecasting, 28, 287-294. doi: http://dx.doi.org/10.1175/WAF-D-12-00060.1

_____, A.B. Schumacher, J.A. Knaff, M. DeMaria, E.M. Fukada, C.A. Sisko, D.P. Roberts, K.A. Winters, H.M. Wilson, 2012: Objective guidance for use in setting tropical cyclone conditions of readiness. *Wea. Forecasting*, **27**:4, 1052–1060.

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

_____, B. Balasubramaniyan, L. D. Burroughs, D. V. Chalikov, Y. Y. Chao, H. S. Chen, and V. M. Gerald, 2002: Development and implementation of wind generated ocean surface wave models at NCEP. Weather and Forecasting, 17, 311-333.

_____, J. H. G. M. Alves, and Y. Y. Chao, 2005: Operational forecasting of wind-generated waves by Hurricane Isabel at NCEP. Wea. Forecasting, 20, 544-557.

U.S. Navy, cited 2016: Naval History and Heritage. [Available on-line at http://www.history.navy.mil]

Figure Captions

Figure 1. (Top) TAFB 48-h forecast and (bottom) analyzed significant wave heights (contoured in ft) for Joaquin verifying 00 UTC October 1, 2015. Blue arrow is approximate path of El Faro and dotted lines indicate latitude and longitude in 10 degree increments. Boxed value (28) indicates maximum significant wave height in ft.

Figure 2. Sensitivity analysis for wind and wave probabilities. Cumulative probability (0-120 h) of wind exceeding 34 kt for 40 (top left), 80 (middle left) and 120 realizations (lower left). Probability of significant wave heights exceeding 12 ft for 40 (top right), 80 (middle right) and 120 realizations (lower right). The case is a 96-h forecast for Yagi (wp162006) September 19, 2006 at 12 UTC.

Figure 3. (Left) Ensemble forecast tracks and (right) forecast intensity (kt) out to 72 h for GFS (top), NAVGEM (middle), and the WSP (bottom). Case is Joaquin (al112015) on September 29th, 2015. Ensemble forecasts are gold, NHC forecast is orange and best track is black with the southernmost position being the 72-h verifying position.

Figure 4. Maximum significant wave height (top) and radius of 12-ft seas (bottom) mean errors and biases for the Atlantic 2013-2015 seasons. WW3_TC_OFCL ensemble mean of 128 realizations (solid line) and WW3_TC_OFCL deterministic (dashed) real-time runs. The forecast period (h) is shown on the x-axis. Number of cases for the top two graphs is 53, 38, 23, 14, 9, and 5 for 0, 24, 48, 72, 96, and 120 h, respectively. Number of cases for the bottom two is 109, 120, 83, 57, 43, and 19 for 0,

24, 48, 72, 96, and 120 h, respectively, and standard error bars shown on ensemble means.

Figure 5. Same as in Figure 4, but for the western North Pacific 2014-2015 seasons and ground truth being the WW3_TC_OFCL deterministic real-time run. The forecast period (h) is shown on the x-axis. Number of cases for maximum significant wave height is 189, 149, 122, 96, 76, and 60 for 0, 24, 48, 72, 96, and 120 h, respectively. Number of cases for 12-ft seas radii is 627, 528, 466, 363, 289, and 224 for 0, 24, 48, 72, 96, and 120 h, respectively, and standard error bars shown on ensemble means.

Figure 6. Mean (purple) and maximum (dashed blue) distance from the ensemble mean of maximum significant wave height for the WW3_TC_OFCL ensemble. The forecast period (h) is shown on the x-axis. Number of cases for the Atlantic is 42, 25, 15, 8, and 7 at 0, 24, 48, 72, 96, and 120 h, respectively. Number of cases for the western North Pacific is 142, 116, 99, 79, and 64 for 24, 48, 72, 96, and 120 h, respectively.

Figure 7. (Upper left) 120-h WW3_TC_OFCL ensemble forecasts of maximum significant wave height in ft, (lower left) 0-120 h cumulative probabilities of 12-ft seas with approximate El Faro route (blue arrow). (Upper right) 72-h forecast 12-ft seas radii for first member of ensemble. Blue (red) shades indicate 0-40% (80-100%) probabilities. (Lower right) NHC best track and estimated 12-ft seas radii. Initial time in all panels is 00 UTC September 29, 2015.

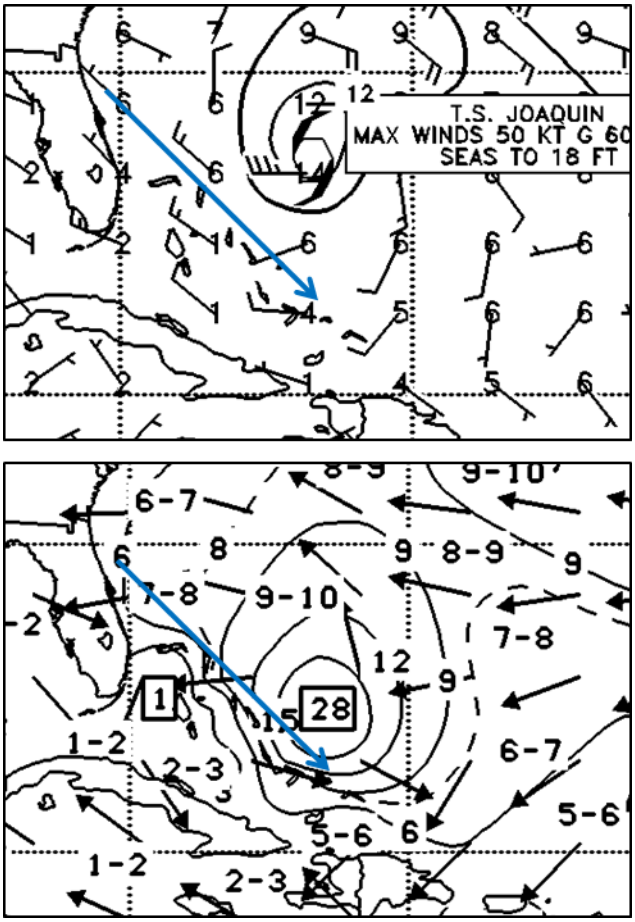

Figure 1. (Top) TAFB 48-h forecast and (bottom) analyzed significant wave heights (contoured in ft) for Joaquin verifying 00 UTC October 1, 2015. Blue arrow is approximate path of El Faro and dotted lines indicate latitude and longitude in 10 degree increments. Boxed value (28) indicates maximum significant wave height in ft.

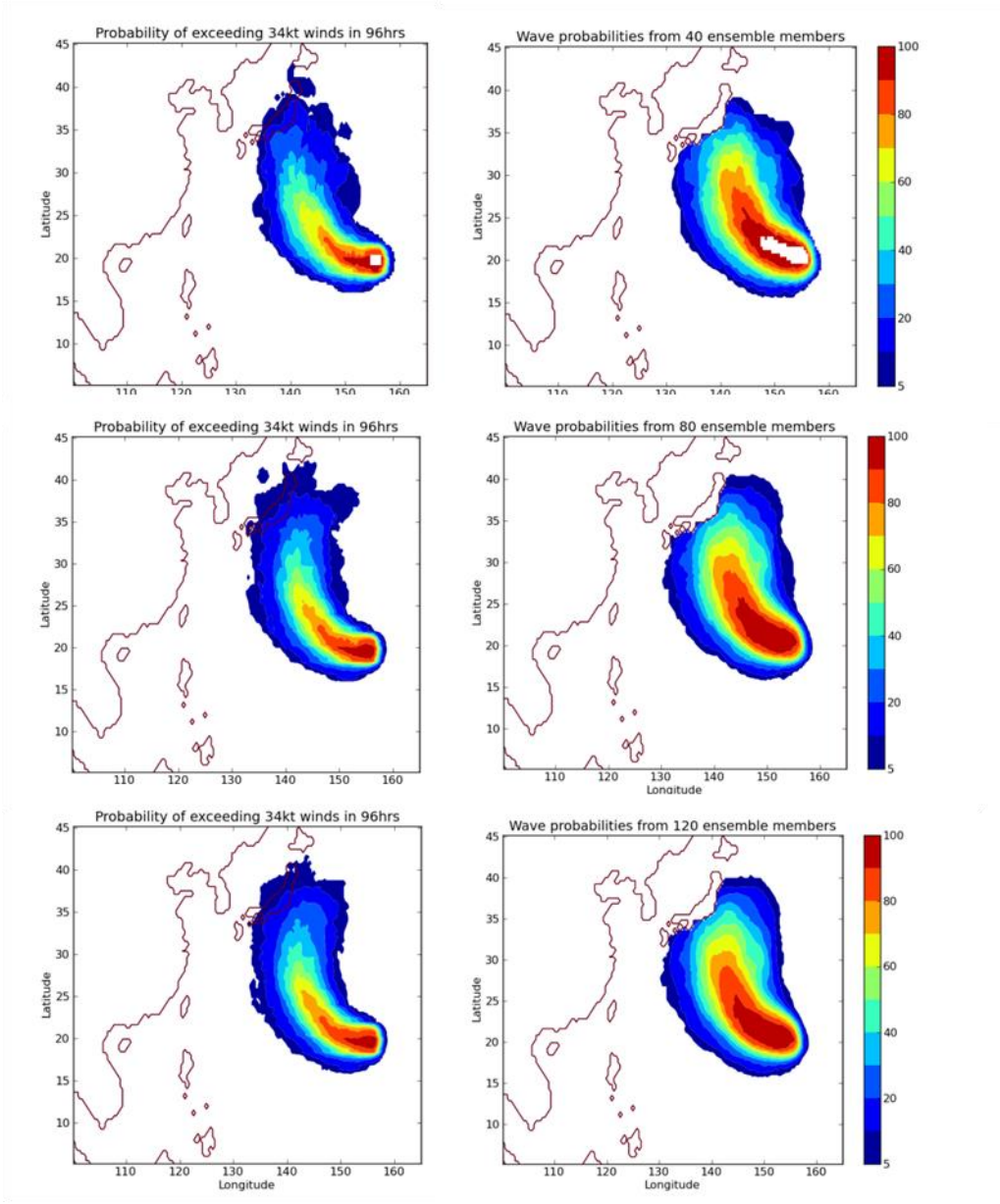

Figure 2. Sensitivity analysis for wind and wave probabilities. Cumulative probability (0-120 h) of wind exceeding 34 kt for 40 (top left), 80 (middle left) and 120 realizations (lower left). Probability of significant wave heights exceeding 12 ft for 40 (top right), 80 (middle right) and 120 realizations (lower right). The case is a 96-h forecast for Yagi (wp162006) September 19, 2006 at 12 UTC.

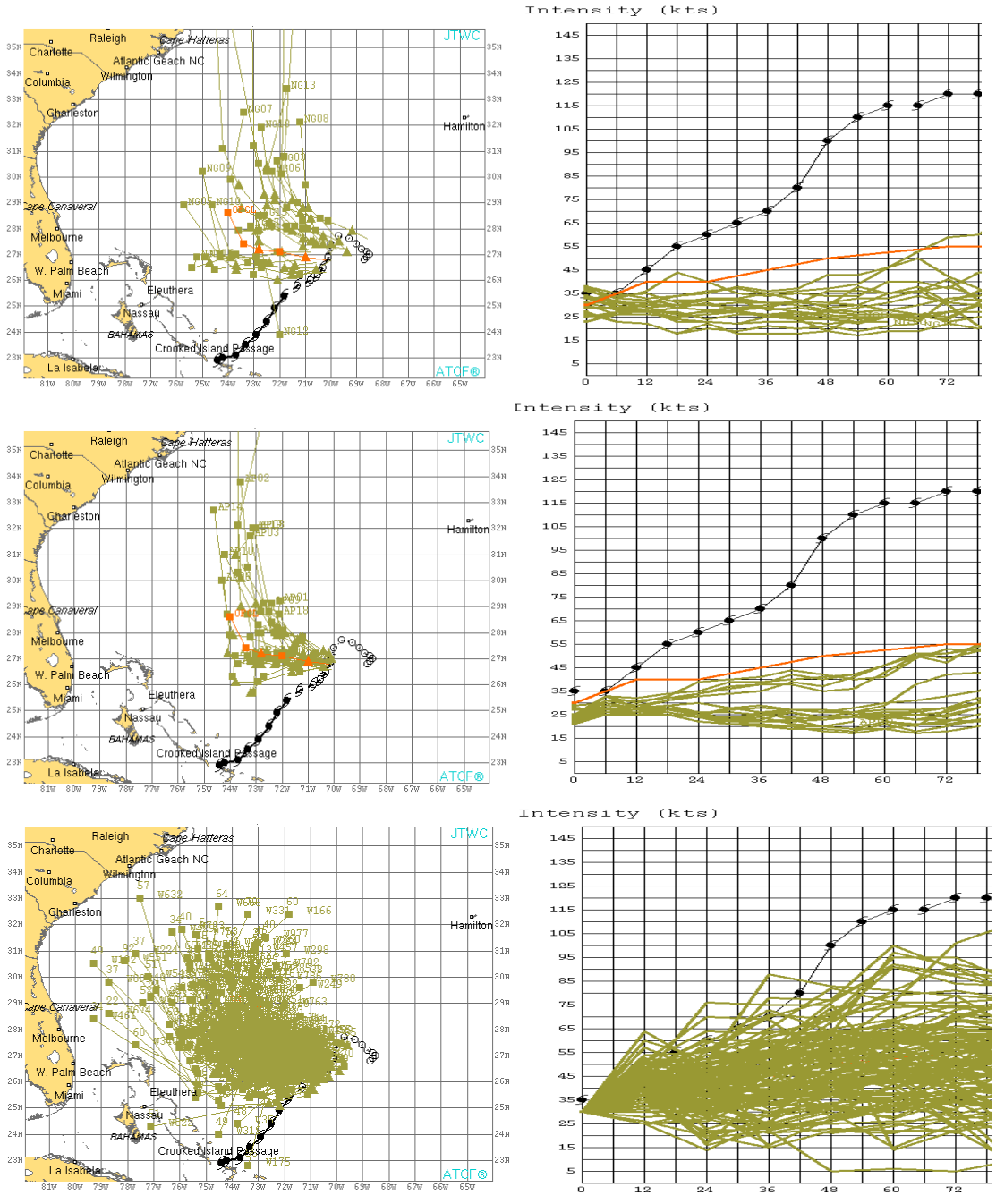

Figure 3. (Left) Ensemble forecast tracks and (right) forecast intensity (kt) out to 72 h for GFS (top), NAVGEM (middle), and the WSP (bottom). Case is Joaquin (al112015) on September 29th, 2015. Ensemble forecasts are gold, NHC forecast is orange and best track is black with the southernmost position being the 72-h verifying position.

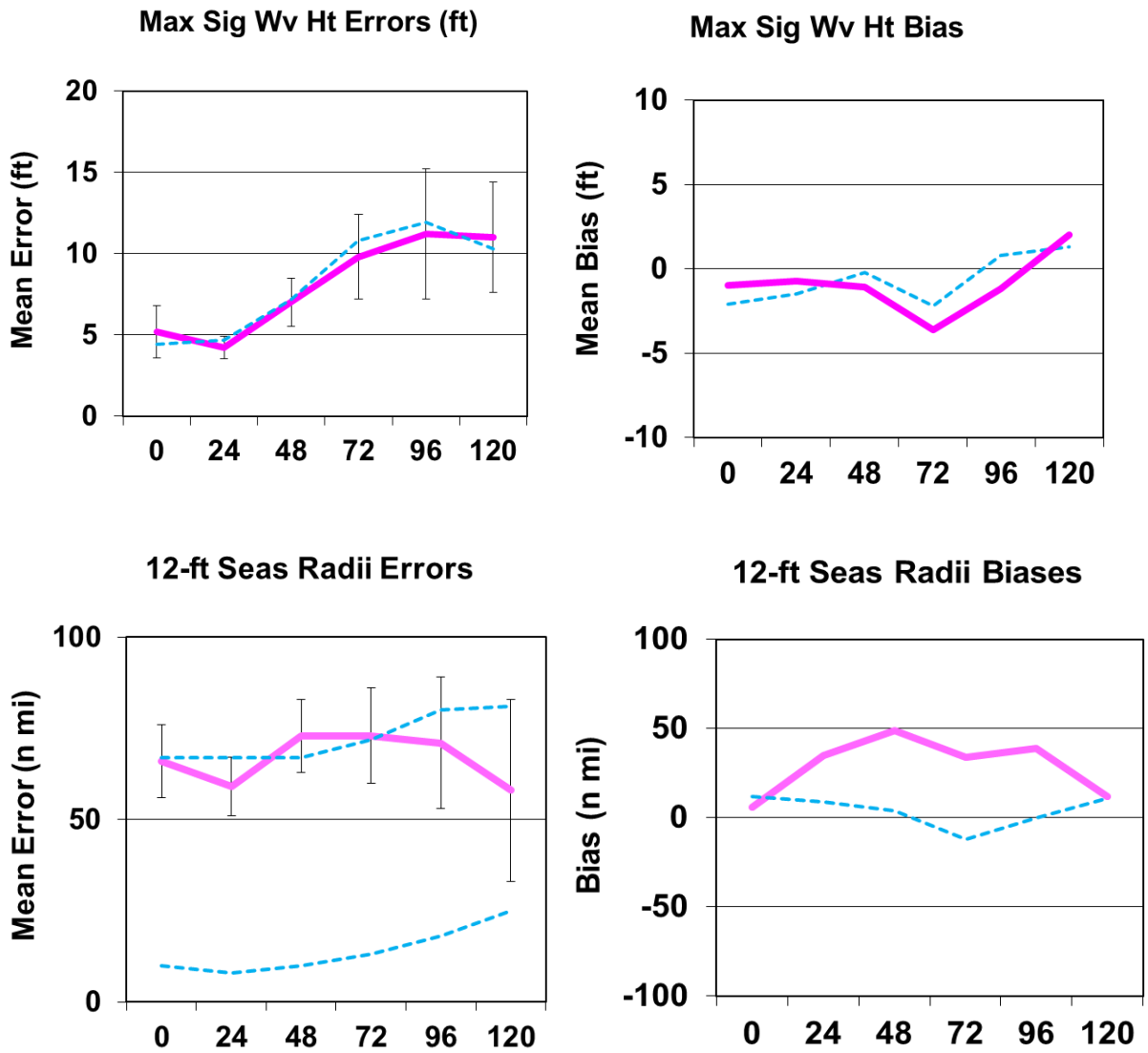

Figure 4. Maximum significant wave height (top) and radius of 12-ft seas (bottom) mean errors and biases for the Atlantic 2013-2015 seasons. WW3_TC_OFCL ensemble mean of 128 realizations (solid line) and WW3_TC_OFCL deterministic (dashed) real-time runs. The forecast period (h) is shown on the x-axis. Number of cases for the top two graphs is 53, 38, 23, 14, 9, and 5 for 0, 24, 48, 72, 96, and 120 h, respectively. Number of cases for the bottom two is 109, 120, 83, 57, 43, and 19 for 0, 24, 48, 72, 96, and 120 h, respectively, and standard error bars shown on ensemble means.

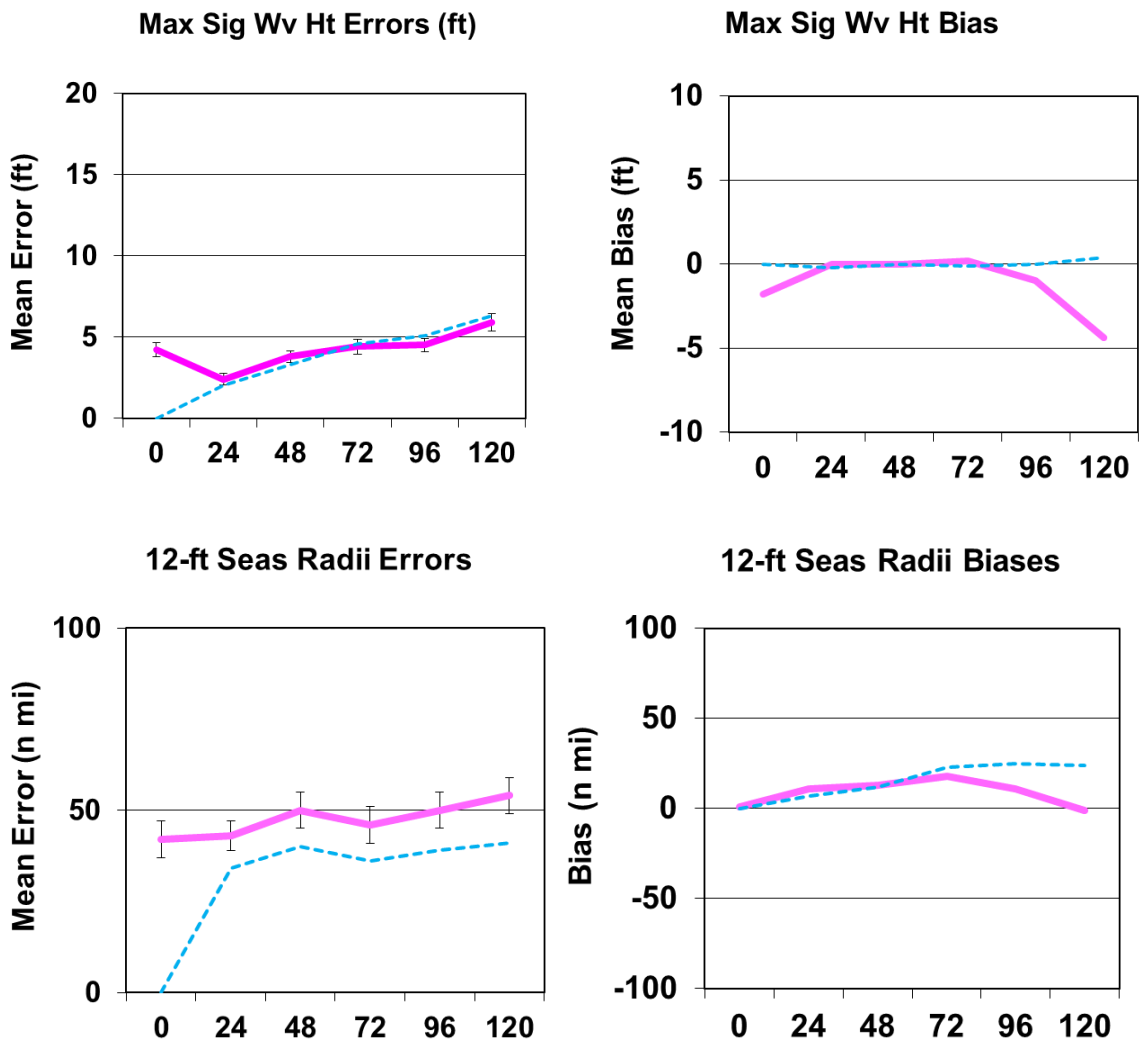

Figure 5. Same as in Figure 4, but for the western North Pacific 2014-2015 seasons and ground truth being the WW3_TC_OFCL deterministic real-time run. The forecast period (h) is shown on the x-axis. Number of cases for maximum significant wave height is 189, 149, 122, 96, 76, and 60 for 0, 24, 48, 72, 96, and 120 h, respectively. Number of cases for 12-ft seas radii is 627, 528, 466, 363, 289, and 224 for 0, 24, 48, 72, 96, and 120 h, respectively, and standard error bars shown on ensemble means.

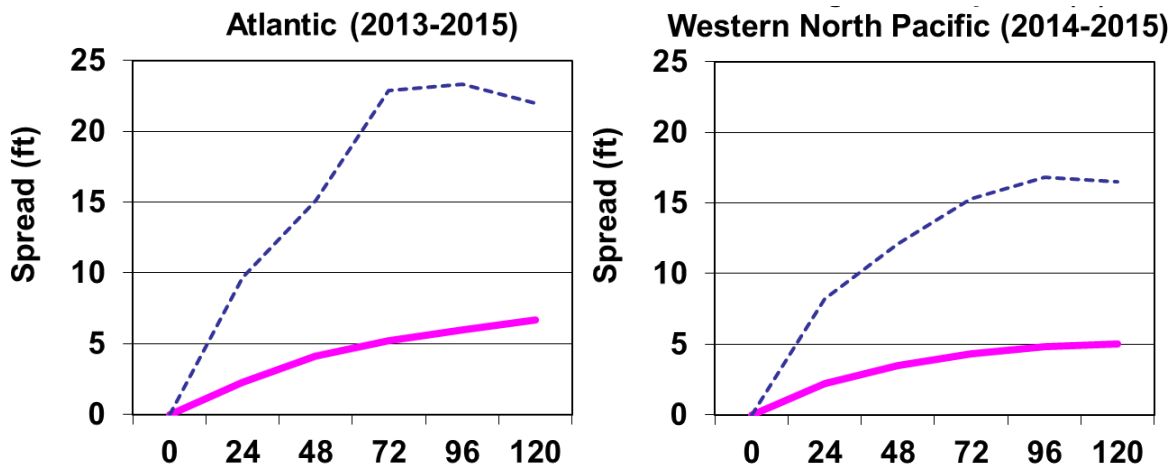

Figure 6. Mean (purple) and maximum (dashed blue) distance from the ensemble mean of maximum significant wave height for the WW3_TC_OFCL ensemble. The forecast period (h) is shown on the x-axis. Number of cases for the Atlantic is 42, 25, 15, 8, and 7 at 0, 24, 48, 72, 96, and 120 h, respectively. Number of cases for the western North Pacific is 142, 116, 99, 79, and 64 for 24, 48, 72, 96, and 120 h, respectively.

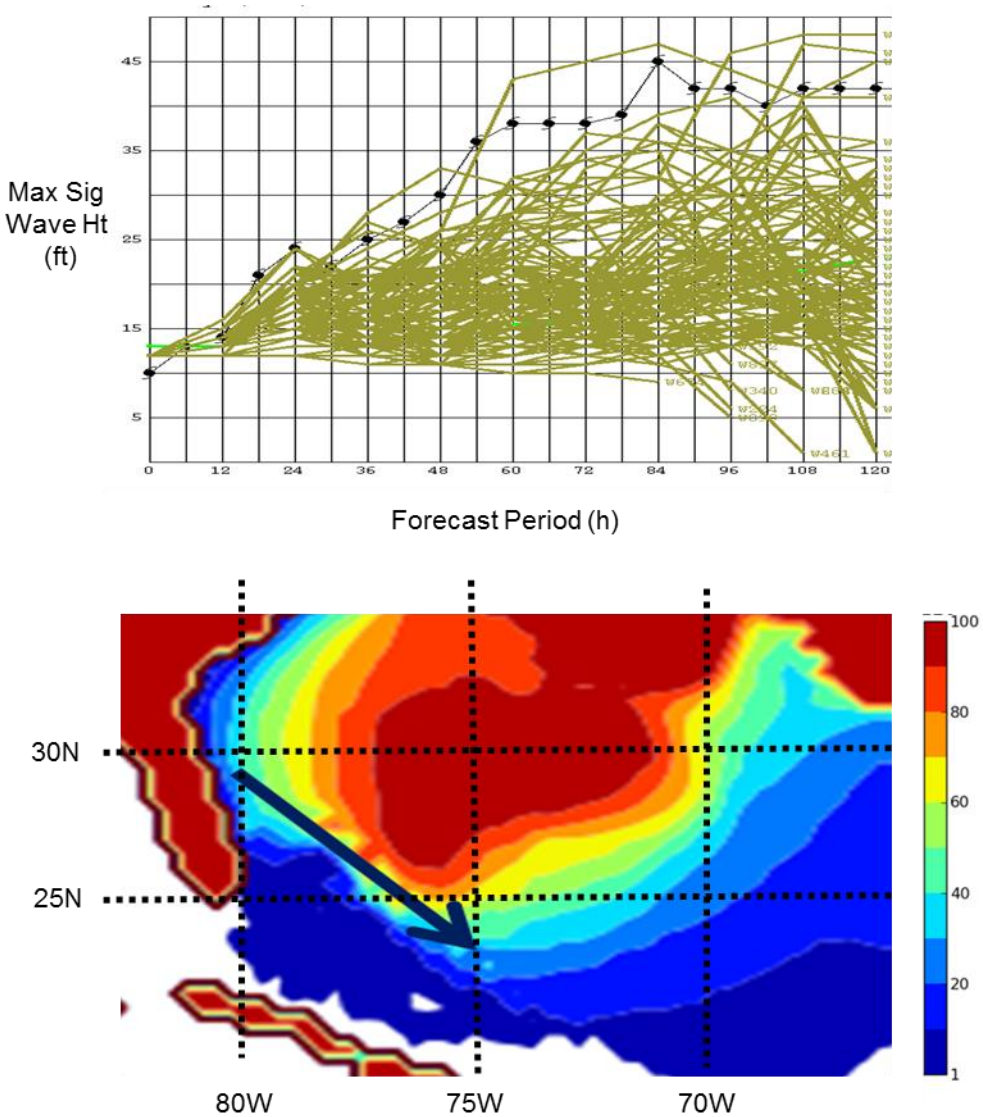

Figure 7. (Top) 120-h WW3_TC_OFCL ensemble forecasts of maximum significant wave height, and (bottom) 0-120 h cumulative probabilities of 12-ft seas with approximate El Faro route (blue arrow). Blue (red) shades indicate 0-40% (80-100%) probabilities. (Lower right) NHC best track and estimated 12-ft seas radii. Initial time in all panels is 00 UTC September 29, 2015.