### **NCEP NOTES**

# **Performance of NCEP Regional Wave Models in Predicting Peak Sea States during the 2005 North Atlantic Hurricane Season<sup>1</sup>**

Yung Y. Chao<sup>2</sup> and Hendrik L. Tolman<sup>3</sup>

Marine Modeling and Analysis Branch, NOAA/NCEP/Environmental Modeling Center

Camp Springs, Maryland

<sup>&</sup>lt;sup>1</sup> MMAB Contribution Nr. nnn<br><sup>2</sup> E-Mail: Yung.chao@noaa.gov.<br><sup>3</sup> E-Mail: Hendrik.tolman@noaa.gov.

### ABSTRACT

Unprecedented numbers of tropical cyclones occurred in the North Atlantic Ocean and the Gulf of Mexico in 2005. This provides a unique opportunity to evaluate the performance of two operational regional wave forecasting models at model predictions of the tropical cyclone generated maximum significant wave height, simultaneous spectral peak wave period and the time of occurrence against available buoy measurements from the National Data Buoy Center (NDBC). The models used are third-generation operational wave models WNA and NAH. These two models have identical model physics, spatial resolution and domain, with the latter model using specialized hurricane wind forcing. Both models provided consistent estimates of the maximum wave height and period, with error typically less than 20% and 30%, respectively, and timing errors of typically less than 5h. Compared to these random errors, systematic model biases are negligible, with a typical negative model bias of 5%. It appears that higher wave model resolutions are needed to fully utilize the specialized hurricane wind forcing, and it is shown that present routine wave observations are inadequate to accurately validate hurricane wave models. the National Center for Environmental Prediction (NCEP). This study validates

### **1. Introduction**

The Atlantic hurricane season of 2005 was extraordinary not only for its early beginning and late ending (May to December) but also for the number and the intensity of tropical cyclones. According to the National Climatic Data Center report on the Climate of 2005 Atlantic Hurricane Season (NCDC, 2006), there were a record of 27 named tropical cyclones, of which 15 were hurricanes. Among them, 7 were major hurricanes of category 3 or higher (i.e., hurricanes Dennis, Emily, Katrina, Maria, Rita, Wilma, and Beta). Four of them reached category 5 (Emily, Katrina, Rita, and Wilma), in which Hurricane Katrina was the most intense and destructive land-falling hurricane on record for the Atlantic basin. Many of these tropical cyclones have created high waves disastrous to the coastal areas and offshore marine activities (in particular, oil exploration and production). Extensive measurements of wind and wave conditions made by the National Data Buoy Center (NDBC) provide an excellent opportunity to validate the NCEP operational regional wave models.

There are two operational regional wave models which forecast sea states over the Western North Atlantic Ocean domain at the National Center for Environmental Prediction (NCEP). These are the Western North 2005). They are part of the National Oceanic and Atmosphere Administration's global wave forecasting suite, NOAA WAVEWATCH III<sup>TM</sup> (NWW3) (Tolman 2002a, Tolman et al., 2002). The performance of the forecast guidance produced by the WNA and NAH models for sea states generated by Hurricane Isabel has been reviewed by, for instance, Tolman et al. (2005). The main purpose of the present study is to assess the accuracy of these two wave models regarding the maximum significant wave height, the associated spectral peak wave period and the time of occurrence for each storm event at a given buoy location. The model results used here are taken from the monthly hindcast data produced by NCEP. They represent operational hindcast model data consistent with the operational real-time products of NCEP and do not include any additional tuning or model modifications. Atlantic wave model (WNA) and the North Atlantic Hurricane wave model (NAH) (Chao et al. 2003 a, b,

#### **2. Models and data**

Wave model data are generated by two operational regional wave models – WNA and NAH. These two models have identical model physics, spatial resolutions and domains. The domain covers an area of latitude  $0^{\circ}$  $-50^{\circ}$  N, and longitude  $30^{\circ} - 98^{\circ}$  W, involving the North Atlantic basin, Gulf of Mexico and the Caribbean Sea. The grid resolution is  $0.25^{\circ}$  by  $0.25^{\circ}$  in latitude and longitude. Both models obtain boundary data from NCEP's global wave model, which has a resolution of  $1.00^{\circ}$  by  $1.25^{\circ}$  in latitude and longitude. The model physics consist of the default model settings of WAVEWATCH III<sup>TM</sup> version 2.22, as described in detail in Tolman (2002b). The difference between the two models lies in their input winds. The WNA model is driven solely with wind obtained from NCEP Global Forecast System (GFS) atmospheric model, previously known as the Medium- Range Forecast (MRF) or Aviation (AVN) model (Caplan et al., 1997). For the NAH model, high-resolution wind fields generated hourly at NCEP by the Geophysical Fluid Dynamics Laboratory (GFDL) hurricane model are blended into the GFS wind field. For the 2005 hurricane season, the GFS model horizontal resolution is T382, approximately 30 km, and the vertical resolution is 64-layers (see History of upgrades to GFS at http://www.emc.ncep.noaa.gov/modelinfo/). The lowest atmospheric level is at a pressure of 997.3 hPa. Since the GFDL hurricane prediction model became operational at NCEP in 1995, it underwent substantial modifications and improvements (Bender et al., 2007). The 2005 GFDL model has a movable three-nest grid configuration. The horizontal resolution of the outermost nest is  $1/2^{\circ}$ , covering an area of 75<sup>°</sup> by 75<sup>°</sup> in latitudelongitude. The size of the inner finer mesh is  $1/6^{\circ}$ , covering an area of  $11^{\circ}$  by  $11^{\circ}$  in latitude-longitude. The finest center core mesh size is  $1/12^{\circ}$ , covering an area of  $5^{\circ}$  by  $5^{\circ}$  aligned to a single storm center. The number of vertical levels in the GFDL model is 42. The lowest sigma level is 996 hPa.

 The 2005 GFS model provides forecast wind fields at 3-hr intervals for the first 180 hours and then at lower spatial and temporal resolutions for up to 16 days for each operational cycle run. The GFDL model, on the other hand, provides forecast wind fields hourly up to 126 hours only. In order to blend with GFDL wind fields, hourly GFS wind fields are generated by interpolation. The required wind field to be used in the wave models is at a 10 m height. Thus, the lowest sigma level wind given by GFS and GFDL models are converted to 10 meter height before blending and interpolating to a uniform  $0.25^{\circ}$  by  $0.25^{\circ}$  wave model grid. The blending scheme is described in detail in Chao et al. (2005). The wave models run oper ationally four cycles per day. Each cycle generates 6-hour hindcast that precedes the actually forecasts. The for ecasts extend up to 180 and 126 hours for the WNA and NAH models, respectively. In this study only hindcast wa ve data are used. It should be noted that the term "hindcast" used in this paper has slightly different connotation from the conventional (engineering) definition. Wave "hindcasts" in the WNA model driven exclusively with GFS winds are generated using 3-hourly analyses from GFS's Global Data Assimilation S ystem (GDAS, see e.g., initial conditions for the wave model real time forecast. Unlike the GFS, the GFDL m odel does not include a data assimilation system for initialization of the model forecast. Thus the NAH mode l hindcast are generated using the GFS analysis winds blended with GFDL forecast winds for the 0-4 hr range fro m the previous cycle (- Caplan et. Al., 1997) for a 6-hr period preceding the current cycle's UTC time stamp and are used to provide 6 to -2 hr range in the current cycle). Wind input for NAH at -1 hr time of the current cycle is obtained by interpolating the -2 hr winds with the blended GFS-GFDL 0-hr nowcast. Although this may seemingly lead to lower quality winds being used for the NAH model hindcast, the higher resolution winds available from the GFDL short-range forecast (0-6 hr) may compensate for deficiencies in the lower-resolution GFS/GDAS analyses.

 Quality controlled wave data for model validation were obtained from the web site of NDBC (the National Data Buoy Center, http://www.ndbc.noaa.gov/historical\_data.shtm). Figure 1 shows the locations of all operational NDBC buoys that provide measured data used in the present study. The results of predictions made by the NAH and WNA wave models on the grid points surrounding these locations are interpolated to these locations for validation. In the present study, hourly data obtained from buoy measurements and model output,

including the wind speed at 10 m above the mean sea level, the wind direction, the significant wave height and the spectral peak wave period are used. The spectral peak wave period is the wave pe riod that corresponds to the frequency bin of maximum wave energy in the wave spectrum. In addition, the sig nificant wave steepness fields are calculated from the NAH model for the significant wave heights greater than 2 m eter. The significant wave steepness is defined here a s the ratio of the significant wave height to the wavelength associated with the spectral peak wave period. Since only a limited amount of wave data obtained from altimeters is available for the present study, they are not included.

### 3. Identification of the peak significant wave height associated with a tropical cyclone

In this study, we assume that waves appeared at a buoy location must have the significant wave height in a continuous record peaking up to greater than 2 meters in order to be considered as being caused by a tropical cyclone. Furthermore, we assume that the submarine bottom effects on wave height, such as wave refraction and bottom friction, can be ignored in the water of depth greater than 200 meter. Consequently, we may use wave data obtained for buoy stations from such locations for storm identification. The procedures to identify a storm that causes the significant wave height to peak up to a maximum (hear after will be called "the peak significant wave height") at a given buoy location at a specific time are best described by example. We use data for the buoy station 41002 off the Atlantic coast in deep water at a depth of 3316 m for September, 2005 for illustration. The example is particularly interesting because of three hurricanes co-existing over the North Atlantic Ocean at one time.

 Figure 2 shows hourly time series of observed and predicted wave and wind conditions for September 2005. The plots include the spectral peak wave period, the significant wave height, the wind speed at 10 m above the mean sea level and the wind direction. There are two significant wave height peaks shown in the second panel

from the bottom of Figure 2. For buoy measurement, the first peak appears at 1300UTC September 6 (090613) and the second peak appears at 091023. For the NAH model, the first peak appears at 090616 and the second peak appears at 091100. And for the WNA model, the first peak is at 090615 and the se cond peak is at 091023. For this example, the significant wave height peaks predicted by WNA model occur an hour earlier than the peaks predicted by NAH model. It should be noted that we are interested in the quan tity of the spectral peak wave period of the wave spectrum from which the calculated significant wave height ap pears to be a peak in the time series. These two quantities are "simultaneous" in time. We are not interested in co-re lating the significant wave height maximum of the significant wave height time series with the maximum value of the spectral peak wave period time series. The occurrence of a peak on the spectral peak wave period time series is not necessarily associated with (or related to) the considered peak on the significant wave height time series.

Five named hurricanes appeared one after another in the western North Atlantic Basin during September 2005. Three of them existed when peaks of the significant wave height occurred at the buoy station 41002. They are Hurricane Maria (category 3) during September 1-10, Hurricane Nate (category 1) during September 5-10, and Hurricane Ophelia (category 1) during September 6-18. Figure 3 shows the track positions of the GFDL hurricane model at 6-hour intervals. The best tracks (the verified tracks) for these hurricanes are also plotted at 6- hour intervals based on data available from the National Hurricane Center (NHC) Archive of 2005 Atlantic Hurricane Season. The date of the best track position at 00 UTC is indicated along the path. The development of storm intensity along the track is indicated by segments of different colors and line types. They might involve tropical low/wave (LO/WV), subtropical depression (SD), subtropical storm (SS), extra-tropical system(EX), tropical depression (TD), tropical storm (TS) and hurricane (HU). It can be observed from Figure 3 that the GFDL hurricane model tracks are virtually the same as the best tracks. This is because data for the initialization of GFDL model is derived from the result of data assimilation (involving the use of observed data) for the GFS model initialization processes (i.e., in the hindcast mod).

In order to determine which one of these hurricanes cause the significant wave height to reach a maximum, the following steps have been taken. We begin with the construction of the wind fields and the wave steepness fields covering the life cycle of tropical cyclone under study. Figures 4 is an example showing the patterns of wind field (lower panel) and wave steepness field (upper panel) when three hurricanes co-exist over the N orth Atlantic Basin. The shaded wave steepness contours are given only for the region wh ere the significant wave height is greater than 2 m. Within the shaded area, the hurricane wind bars and the direction of spectral peak wave period are presented. The direction of spectral peak period is considered to be the representative wave direction.

We then visually examine sequential plots of the vector wind field and model derived significant wave steepness pattern. We first observe the pattern orientation and the extent of the wind field and wave steepness field to see if they are moving toward the buoy location (The animation of the significant wave steepness fields at three-hour intervals is very helpful.) If these fields indeed move toward and eventually cover the buoy location, we then examine whether the directional variation of wind and wave inside the shaded wave steepness areas is consistent with the time series of wind and wave direction at the buoy location as shown in Figure 2. It is tedious, time-consuming, trial and error processes. But in this manner, the storm that causes the wave height to peak up to a maximum at the given location and time can be identified eventually. For the case of buoy station 41002 in September 2005, it is found that the first wave height peak shown in Figure 2 is identified to be caused by Hurricane Nate and the second wave height peak is identified to be caused by Hurricane Ophelia.

 The same procedure is applied to all tropical cyclones that occurred in 2005 for all available deep water buoy stations. Table 1 list the names of tropical cyclones and the deep water buoy locations where the significant wave height peaks to more that 2 m height are observed and/or modeled. A total of 14 storms (among 28 for the whole 2005 hurricane season) are identified to have the peak significant greater than 2 m at one or more than one of 14 deep water buoys. For buoy stations in shoaling waters (Buoy stations in the water depths less than or equal to 200 meter), peak conditions associated with a specific storm event is inferred from nearby deep

water buoy stations. Detailed one-to-one comparison of the significant wave height, p eak period and the time of occurrence between buoy measurements and model predictions are given in Appendix -A and Appendix-B for the Atlantic b asin and the Gulf of Mexico, respectively. In these appendixes, additional 16 shoaling water buoy sites are included.

### **4. Model performance**

As previously mentioned, our main objective in this study is to evaluate the performance of WNA and NAH models for the western North Atlantic basin and the Gulf of Mexico-Caribbean Sea, respectively, in predicting the tropical cyclone generated maximum significant wave height, simultaneous spectral peak wave period and the time of occurrence during the 2005 hurricane season. These two regions are considered separately because the Gulf of Mexico-Caribbean Sea is a semi-enclosed basin while the Atlantic basin is an open ocean. The accuracy of prediction for the two regions might differ due to different geographic constrains on the characteristics of tropical cyclone induced wind waves. The section is divided to two sub-sections. Section 4.1 evaluates the wind speed and the significant wave height predictions for each tropical storm against available buoy observations around the time evolution of the peak significant wave height. Section 4.2 then evaluate specifically the performance of models in predicting the peak significant wave height, the simultaneous wave period and the time of occurrence.

### **4.1 Five-day statistics around the significant wave height peak**

 We begin with an evaluation of modeled wind speeds and significant wave heights against buoy measurements for four selected storms over five days time span around the significant wave height peaks. The selected storms are three Category 5 hurricanes, namely Katrina, Rita and Wilma and a Category 1 hurricane Ophelia. Hurricane Ophelia never made landfall but because of its slow movement along the East Coast

coastline, it produced sustained high waves for several days (see Figure 3 for the track of Hurricane Ophelia). Each buoy selected represents the site where the maximum significant wave height pe ak of the corresponding hurricane was ever recorded among all buoys. Although rather subjective, the selected 5-day time span is assumed to be sufficient to see the rise and fall of the significant wave height around the peak. For each selected storm and buoy site, a total of 120 hourly data points are involved. Figure 5 exhi bits the tracks of three category 5 hurricanes. Figures 6a to 6d present the time histories and scatter plots of the wind speed at 10 m height above the mean sea water level  $(U_{10})$  and the significant height  $(Hs)$  caused by Hurricanes Katrina at Buoy 42040 and Rita at Buoy 42001 in the Gulf of Mexico, Wilma at Buoy 42056 in the Caribbean Sea and Ophelia at Buoy 41002 in the North Atlantic Basin. Also shown in these figures are the root mean square error (RMSE), mean bias (BIAS), correlation coefficient (COR), scatter index (SI) and the lin ear trend, including the the mean observation. It can be seen from the time series plots shown in Figure 6a that for Hurricane Katrina at Buoy 42040 during the time period of 08/2700 through 08/3123, *U10* of WNA and NAH are both over-predicted for most of time, especially WNA near the peak However, WNA make a much better o ver all prediction of *Hs* and error statistics for NAH and WNA models show that the bias in  $U_{10}$  of WNA and NAH are both positive, yet both are negative in *Hs*, the bias is almost negligible for WNA, however. For Hurricane Ophelia at Buoy 41002, as shown in Figure 6b, the slowly moving feature is shown by relatively long du ration of *U10* around 20 slope and the intersection with an axis. The scatter index is defined as the root mean square error normalized by than NAH, particularly near the peak. The NAH predicted *Hs* are much lower than measured. The scatter plots m/s and *Hs* around 6 m for almost two days. Again, *U10* is over-predicted, and *Hs* are under-predicted for both models. Note that there are 6 missing data points in model predictions. Figure 6c shows the "worst" wind input for WNA and NAH wave models. As shown in the time evolution of Hurricane Rita, there is a steep drop of the wind speed. As shown in Figure 5, the center of Hurricane Rita is in the proximity of the 42001 buoy site at about 09/2300 and the modeled winds tend to indicate the condition near the eye of the hurricane coincided with a rapid change in the wind direction; the wind blows counterclockwise from NNE to NW, to W then to S within

five hour period (The time history of wind directional variations is not shown.) In spite of the substantial discrepancy in the modeled wind speed in comparison with buoy measured, the predict ed *Hs* seems to behave fairly well underlying the modeled *Hs* respond to wind speed variations much slowly and less dramatics. The results of NAH and WNA predictions for Hurricane Wilma at Buoy 42056, which is located in the Caribbean Sea are shown in Figure 6d. As shown in Figure 6d the WNA model over predicts  $U_{10}$  and *Hs* around the peak but predicts quite consistent with observations in the ascending and descending stages. O n the other hand, NAH model predicts the peak *Hs* near the same height as measured but the time of occurrence is much earlier than measured even though modeled maximum wind speed is consistent with measurem ent in time and in magnitude. During the descending stage both *U10* and *Hs* are considerably u nder-predicted. The scatter plots linear regression equations and have negative BIAS, large RMSE, large SI and low COR while the linear trend for WNA  $U_{10}$  and  $Hs$  indicate both are over-predicted but are fairly good in the statistical quantities. shown in Figure 6d reveals quantitatively that NAH modeled *U10* and *Hs* have substantially low slops on the

The statistical evaluation of NAH and WNA for selected hurricanes at selected buoys described above has extended to all tropical cyclones occurred in the 2005 hurricane season at all available deep-water buoys based on the procedure described in a previous section. Figures 7a and 7b summarize the results for the significant wave height  $(Hs)$  and the surface wind  $(U_{10})$ , respectively. They are constructed based on data given as Appendix-C and Appendix-D of this paper. The vertical axis for each panel in these figures represents one of statistical quantities described previously, i.e., RMSE, BIAS, COR, SI, and "a" (slope) and "b" (intersection) terms of the linear trend. The horizontal axis represents the case number, which is the event number assigned to the combination of a deep water buoy and a tropical storm been identified. The case number is assigned in the ascending order for buoy ID number and the alphabetic order of storm names for the Atlantic and the Gulf. As shown in Appendix-C and Appendix-D, the case number 1 to 19 are for the Atlantic Basin, e.g., No. 1-3 represent 41001 for Maria, Ophelia, and Wilma; and No.17-19 represent 44004 for Maria, Ophelia and Wilma. Cases 20 – 55 are for the Gulf of Mexico and the Caribbean Sea, e.g. No. 20–27 represent 42001 for Arlene,

Cindy, Dennis, Emily, Katrina, Rita, Stan, and Wilma, and No. 54-55 represent 420 57 for Wilma and Beta (Hurricane Beta is an exception. It does not follow the alphabetic order). In each pan el, values correspond to NAH and WNA modeled are shown by color blue and red, respectively. The dash lines indicate the mean. Also given in the panel are the mode and the standard deviation (std) of the data set. The m ode in statistics is not necessarily unique, but if it is considered in conjunction with the mean, it can capture important information Based on graphs shown in about what is the value that is most likely to be expected in a discrete data set. Figures 6 and 7 the following observations might be made:

- 19 of the horizontal axis) and the Gulf of Mexico regime (case No.20-55) from NAH or WNA modeled (1) There are hardly distinct differences visually in the resulting statistics for the Atlantic Basin (case No.1- *Hs* or  $U_{10}$ .
- (2) In considering the mean and the mode values given for each statistical quantities, both NAH and WNA models provide the following results:
	- (a) For  $H_s$ , the RMSE is about  $0.5$  m, the BIAS is  $0.1$  m or less, the COR is higher than  $0.9$ , and SI is less than  $0.2$  (20%). The slope of the linear regression line ("a" term) is less than  $0.95$  and the intersection ("b" term) is around 0.1 m, indicating that the models tends to under-predict Hs slightly ;
	- (b) For  $U_{10}$ , the RMSE is around 2 m/s, the BIAS is near zero but with opposite sign on values for the mode, the COR is only slightly above 0.80, and the SI is around 20%. The slope of the linear regression line ("a" term) is nearly 1.0 and the intersection ("b" term) is closed to 1.0 m/s. indicating the tendency of slight over-prediction of the wind speed.
- (3) WNA performs comparably to better than NAH in over-all statistical results.
- (4) There is a substantial number of outlying points that deviate beyond one "std" from the mean in each statistical quantity. No attempt is made to get rid of those extreme values in this paper.

(5) The present study clearly shows the complexity of the hurricane wind field a nd wave field that the validation of model performance for one storm event at limited buoys sites is not n ecessarily applicable to another storm event. An in-depth investigation of the model performance o n each storm scenario regarding the causes of success or failure is important for the improvement of modeling methodology but is beyond the scope of the present study.

### **4.2 Statistics for the peak significant wave height and the associated wave period**

A major concern in an operational wave forecasting system is the ability to forecast the possible maximum wave height and the time of occurrence at a given location associated with a given tropical cyclone. The remaining part of the present section evaluates the deviations of the NAH and WNA modeled peak significant wave height (hereafter, the peak *Hs*) and the associated spectral peak wave period (hereafter, *Tp*) and the time of occurrence against buoy measurements.

Figure 8a, containing four panels, depicts the scatter plots of the NAH and WNA modeled peak *Hs* and *Tp* for plotted with different symbols. In addition, the over-all error statistics including Root Mean Square Error, Bias, Correlation Coefficient and Scatter Index along with the linear trend for all storms and buoys involved are also presented. Figure 8b shows the normalized bias (difference) between model predictions and buoy measurements of the peak *Hs* and *Tp* as a function of the time lag (difference) in occurrence. The bias is normalized with the buoy measurements and is expressed in percentile on the vertical axis. The time lag is expressed in hour on the horizontal axis: negative (or positive) time lag means that the predictions are earlier (or later) than actually observed. The central line on each graph represents the mean value of the labeled quantity while the outer two lines represent one standard deviation from the mean. Similar graphs for the Gulf of Mexico-Caribbean Sea are depicted in Figures 9a and 9b based on the dataset given in Appendix-D. all North Atlantic tropical cyclones as shown by the asterisk symbol. The data base is given in Appendix-A. In addition, the peak *Hs* and *Tp* associated with the hurricanes of particular interest at various buoy sites are

The scatter plots of the peak *Hs* for the NAH and WNA models are shown in the top rows of Figure 8a for the Atlantic Basin and in Figure 9a for the Gulf of Mexico. The plots indicate that the W NA model prediction is slightly better than the NAH model prediction in either the Atlantic Basin or the Gulf of Mexico-Caribbean Sea (based on the slope of regression line). Both models under-predict the peak significant wave height for the Atlantic Basin, but predict reasonably well for the Gulf regimes. However, the regres sion slope for the NAH and WNA for the Atlantic Basin are mostly suppressed due to the underestimation of o nly a small subset of the most extreme wave conditions. Furthermore, for both models, the correlation between observations and m odel predictions in the Gulf region is better than in the Atlantic basin. The scatter plots of the *Tp* for the NAH and WNA models (bottom rows of Figure 8a for the Atlan tic and Figure 9a for the Gulf regions) show results similar to those for the peak  $H_s$ . However, in comparison with the scatter plots of the peak  $H_s$ , the  $Tp$  has greater bias and lesser correlation coefficient. This is consistent with typical wave model validation results for the spectral peak wave period (Bidlot et. al, 2002; Tolman et. al., 2005).

The time-lag plots of the peak *Hs* for the NAH and WNA models are shown in the top rows of Figure 8b for the Atlantic Basin and in Figure 9b for the Gulf of Mexico-Caribbean Sea. As can be observed from the top rows of Figure 8b and Figure 9b, the normalized bias of the peak  $H_s$  is mainly within  $\pm 20$  % but may reach  $\pm 30$ % of the observed value for the Atlantic Basin and  $\pm 50$ % for the Gulf of Mexico-Caribbean Sea. The plus sign indicates over-prediction while the minus sign indicates under-prediction. The mean bias of the peak *Hs* for both models is approximately -5% indicating that the models tend to under-predict. The result is consistent with the findings from the scatter plots described previously. The time lag of model predicted peak *Hs* spread considerably although mostly clustered within  $\pm 5$  hours of the observed peak. On average, the NAH model is slightly late (in the order of one hour) in predicting the peak *Hs* in the Gulf of Mexico-Caribbean Sea; otherwise both models are slightly early (in the order of 1-2 hour). Note that the observation accuracy of the timing of the peaks is known to be  $\pm 1$  hour. The time-lag of the *Tp* for both models (bottoms rows of Figure 8b and Figure 9b) shows similar result as the validation of the wave height. Errors are dominated by the model uncertainty,

with the bias comparatively small (less than  $\pm 5\%$ ).

### **5. Discussion**

generated waves, with a potential accuracy of peak significant wave heights that deviate from observations by roughly 30% within 5 hours of the observed time of these maxima. The associated mean biases are much smaller (typically 5%), in comparison to the corresponding random model error. We consider wave model "hindcast" only in this study. Hence, it should be emphasized that the present results merely identify the potential accuracy of wave model prediction in the framework of the real-time operational environment. It is anticipated that the accuracy of hurricane associated extreme wave forecasts will be similar depending strongly on the results of the track and wind intensity forecast of the tropical cyclone that might occurred. For instance, forecast errors for wave models for hurricane Isabel in 2003 are discussed in detail in Tolman et al. (2005). Both the WNA and NAH wave models are capable of providing useful forecast guidance for hurricane

Considering the problems of providing accurate hurricane wind nowcast and forecasts, the method of blending GFS and GFDL model wind fields for the NAH model becomes a subject of concern. In previous hurricane seasons, the NAH model in general has outperformed the WNA model (see Chao et al., 2005). However, for the 2005 season, the models behaved similarly, with arguably better behavior for the WNA model. In this context, it is important to realize that the wind blending algorithm was developed almost a decade ago. At that time GFS, previously known as Medium-Range (MRF) and Aviation (AVN) models, had a grid resolution of about 50 km which was too coarse to resolve the wind field structure associated with a relatively small hurricane vortex. Thus, the blending algorithm was initiated to incorporate the GFDL hurricane model and take advantage of its high resolution inner mesh of about 15 km (Chao and Tolman, 2000; Chao and Tolman, 2001). Since then, GFS underwent various improvements; among them was the change of grid resolution to about 30 km in 2005. As a result, GFS was able to provide improved wind forecast near the hurricane core. More importantly, the resolution of the GFS is now comparable to the resolution of the wave

models. Conversely, the resolution of the GFDL model winds are much higher than th e resolution of the wave models, and hence the wave models no longer make optimal use of the resolution of the hurricane wind models. It therefore appears to be necessary to increase the spatial resolution of the (hurr icane) wave models to the hurricane wave model to utilize hurricane winds at or near the native resolution of the hurricane wind fields. effectively use the increased resolution of the hurricane wind models. For this reason, it is necessary to upgrade

Another reason for the apparently comparable model behavior of the WNA and NAH wave models may be history of wind and wave data at the buoy station 42040 near the track are shown in Figure 5 and Figure 6a, respectively). The upper panels in Figure 10 show the wind fields of the WNA and NAH models. Both models have near identical tracks, with the centers of maximum wind shifted by 10 to 20 km. The NAH winds are more intense with reasonable spatial scales, but are shifted too much to the shallow waters (west). The WNA winds have a lower speed but larger spatial scale. This produces good wind results at the only relevant observation location (buoy 42040), although the wind fields as a whole are less realistic than the NAH wind fields (Chao et al., 2005; Tolman et al., 2005). The corresponding wave height fields (lower panels in Figure 10) are also shifted between the models, due to the similar track but different spatial scales of the wind fields. If only buoy data at buoy 42040 are considered, one could easily come to the conclusion that the WNA model is far superior (Figure 6a). With only the buoy in the view of Figure 10, there is clearly insufficient information to rigorously validate hurricane wave models, unless the hurricane track is close to the buoys (see Chao et al., 2005; Tolman et al., 2005 for case studies). It therefore appears essential to have routine on-demand wave observations in hurricanes, as was available for hurricane Bonnie from a Scanning Radar Altimeter (Alves et. al., 2004; Wright et al., 2001), to systematically address the accuracy of hurricane wave models. the sparsity and a corresponding lack of representativeness of the validation data. This is illustrated in Figure 10 with results for Hurricane Katrina near landfall on September 29, 1200UTC (The hurricane track and the time

 Note that the model resolution in 2005 was insufficient to resolve this coastline, and therefore results at buoy 42007 cannot be expected to be very accurate. Furthermore, wave heights in the shallow waters behind the

Chandeleur Islands are obviously unrealistic due to the lack of shallow water physics i n the model and due to the fact that the spatial resolution is too poor to introduce these islands as obstructions. For the 2007 model implementation, the coastal resolution in this area is g reatly improved, and surf-zone physics (depth-induced breaking) were added to the model (Chawla et al. 2007. Tolman 2008).

### **6. Conclusions**

In this study, we validate NCEP's operational Western North Atlantic regional wave model (WNA) and North Atlantic Hurricane wave model (NAH) against NDBC buoy measurements for more than 20 tropical cyclones (including three category 5 hurricanes) for the 2005 hurricane season. The parameters evaluated include the maximum significant wave height, corresponding spectral peak period and the time of occurrence induced by each individual tropical cyclone. The results show that the deviation of model predicted wave heights and periods from buoy measurements is essentially within 20% and 30%, respectively, and that the time lag (behind or ahead of observation) on the occurrence of peak wave height is within the 5 hour range for both models. Both models show similar behavior, with model uncertainty dominating the mean model bias, which is typically approximately 5%. Considering that these are operational model results produced in near real time with no case specific tuning of the wave model or the wind fields, the biases of both models can be considered to be rather good. Clearly, the model presents useful results for real time forecasting, but also leaves room for improvement. The similar behavior of the WNA and NAH models suggests that the hurricane wave model, NAH, no longer optimally uses the higher resolution of the hurricane wind model, suggesting that the spatial resolution of the hurricane wave model needs to be increased to be comparable to that of the hurricane wind model. Note that generally, better validation of hurricane wave models is greatly hampered by the lack of wave observations with suitable spatial coverage.

 The NAH and WNA as many other existing third generation (3G) models are ess entially developed and validated on extra-tropical wind forcing regimes characterized with slowly varying w ind fields in space and time. The application of such a model in the real time operational environment for tropical cyclones which are characterized with the fast varying extreme surface wind fields along the moving storm track faces various obstacles and uncertainties. The sparsity of measured data is just one of them. We like to stress that the models are intended as operational models for real-time forecasting. Even if there is insufficie nt data to do a rigorous statistical analysis of bias versus uncertainty, it appears obvious to us from the present study that a hum an forecaster using these model data to do his or her work, will have to expect model uncertainty to be the main problem with the guidance, and that adding a systematic bias correction to model guidance is a minor correction compared to this uncertainty. Hence, we cannot, based on the sparse data, do an in-depth statistical analysis, but, from the perspective of these being operational forecast models, feel confident to say that the biases of the model are small compared to the general uncertainty.

*Acknowledgments.* The authors would like to thank Janna O'Connor, Arun Chawla, Robert Grumbine and the anonymous reviewers, for their valuable comments and suggestions on our drafts of this manuscript.

### **REFERENCES**

- Alves, J.H.G.M., H.L. Tolman and Y.Y. Chao, 2004: Forecasting hurricane-generated wind waves at NOAA/NCEP. Preprints 8<sup>th</sup> International Workshop on Wave Hindcasting and Forecasting. Turtle Bay, Hawaii, Paper G3.
- Bender, M.A., I. Ginis, R. Tuleya, B. Thomas, and T. Marchok, 2007: The operational GFDL Coupled hurricane-ocean prediction system and a summary of its performance. Mon. Wea. Rev., 135, 3965-3989.
- Bidlot, J.R., D.J. Holmes, P.A. Whittman, R. Lalbeharry, and H.S. Chen: Intercomparison of the performance of operational ocean wave forecasting systems with buoy data. AMS, Apr. 2002, 287-310.
- Caplan, P., J. Derber, W. Gemmill, S.Y. Hong, H.L. Pan, and D. Parish, 1997: Changes to the 1995 NCEP operational medium-range forecast model analysis-forecast system. Wea. and Forecasting, 12, 581-594.
- Chao, Y. Y., and H. Tolman, 2000: Numerical experiments on predicting hurricane generated wind waves. Preprints, *Sixth Int. Workshop on Wave Hindcasting and Forecasting*, Monterey, CA, U.S. Army Engineer Research and Development Center, 167-179.
- Chao, Y.Y. and H. Tolman, 2001: Specification of hurricane wind fields for ocean wave prediction. Proc. *Fourth Int. Symp., on Waves: Ocean Wave Measurement and Analysis*, Vol. 1, San Francisco, CA, Office of Naval Research, 671-679.
- Chao, Y.Y., L.D. Burroughs, and H.L. Tolman, 2003a: The North Atlantic Hurricane wind wave forecasting system (NAH). Technical Procedure Bulletin 478, National Weather Service. [Available online at http://polar.ncep.noaa.gov/mmab/tpbs/operational.tpbs/tpb478/tpb478.htm.]
- Chao, Y.Y., L.D. Burroughs, and H.L. Tolman, 2003b: Wave forecasting for the western North Atlantic and adjacent waters. Technical Procedure Bulletin 495, National Weather Service. [Available online at http://polar.ncep.noaa.gov/mmab/tpbs/operational.tpbs/tpb495/tpb495.htm.]
- Chao, Y.Y., J.H.G.M. Alves, and H.L. Tolman, 2005: An operational system for predicting hurricane-generated wind waves in the North Atlantic Ocean. *Wea. and Forecasting,* **20,** 652-671.
- Chawla A., D. Cao, V. Gerald and H.L. Tolman, 2007: Operational implementation of a multi-grid wave forecasting system, 10th International Workshop on Wave Hindcasting and Forecasting, Turtle Bay, Hawaii.
- NCDC (National Climatic Data Center), 2006: Climate of 2005 Atlantic Hurricane Season. [Available online at http://www.ncdc.noaa.gov/oa/climate/research/2005/hurricanes05.html]
- *nt. Workshop on Wave*  Tolman, H.L., 2002a: The 2002 release of WAVEWATCH III. Preprints. *Seventh I Hindcasting and Forecasting*, Banff, AB, Canada, Environment Canada, 188-197.
- Tolman, H.L., 2002b: User manual and system documentation of WAVE WATCH III version 2.22. Tech. Note 222, OMB214, Ocean Modeling Branch, National Weather Service, 133pp.

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

- Tolman, H.L., B. Balasubramaniyan, L.D. Burroughs, D.V. Chahlikov, Y.Y. Chao, H.S. Chen, and V.M. Gerald, 2002: Development and implementation of wind-generated ocean surface wave models at NCEP, *Wea. Forecasting*, **17**, 311-333.
- Tolman, H.L., 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.
- Wright, C.W., E.J. Walsh, D. Vandemark, W.B. Krabill, A.W. Garcia, S.H. Houston, M.D. Powell, P.G. Black, and E.D. Marks, 2001: Hurricane directional wave spectrum spatial variation in the open ocean. *J. Phys. Oceanogr*., **31,** 2472-2488**.**

### Figures

Fig. 1 Locations of NDBC buoys used in model validation.

Fig. 2 Monthly time series of measured and predicted spectral peak period, significant wave height, wind speed and wind direction (from bottom to top) at Buoy 41002, September 2005.

Fig. 3. "Best tracks" and GFDL model tracks for Hurricanes Maria, Nate, and Ophelia.

Fig. 4. Wave steepness and blended wind fields while Hurricanes Maria, Nate and Ophelia co-existed. (The bottom figure shows from east to west, Hurricanes Maria, Nate and Ophelia.)

Fig. 5. "Best tracks" and GFDL model tracks for Hurricanes Katrina, Rita and Wilma.

Fig. 6a. The time history and error statistics of NAH and WNA predicted significant wave height and wind speed at 10m height  $(U_{10})$  for Hurricane Katrina at Buoy 42040.

Fig. 6b. Same as Fig.6a except for Hurricane Ophelia at Buoy 41002.

Fig. 6c. Same as Fig.6a except for Hurricane Rita at Buoy 42001.

Fig. 6d Same as Fig.6a, except for Hurricane Wilma at Buoy 42056.

Fig. 7a. Error statistics and linear trend of NAH (in blue) and WNA (in red ) predicted significant wave height for all tropical cyclones at all buoy sites (Dash lines show the mean values).

Fig. 7b. Same as Fig. 7a except for the wind speed.

Fig. 8a. Scatter plots of the peak significant wave height ( $H_s$ , top row), and the associated spectral peak period ( $Tp$ , bottom row) for NAH (left column) and WNA (right column) for the Atlantic Basin. Legend: W-Wilma; O-Ophelia; K-Katrina; R-Rita; Number-buoy I.D.

Fig. 8b. Time lag of normalized bias of the peak significant wave height (top raw) and the associated spectral peak period (bottom raw) predicted by the NAH (left column) and the WNA (right column) models for the Atlantic Basin. In each panel, center lines represent the mean and the outer lines represent the standard deviation. Marker's symbols and colors have the same legend as Fig. 8a.

Fig. 9a. Same as Fig. 8a except for the Gulf of Mexico-Caribbean Sea. Legend: W-Wilma, R-Rita, K-Katrina.

Fig. 9b. Same as Fig. 8b except for the Gulf of Mexico-Caribbean Sea. Legend: Same as Fig. 9a.

Fig. 10. A comparison of wind (top row) and wave (bottom row) fields predicted by WNA (left col.) and NAH (right col.) for Hurricane Katrina, September 29, 1200UTC.

# Table

Table 1. List of tropical cyclones for the wave model validation study

### **Appendices**

Appendix A The peak significant wave height, simultaneous spectral peak period, time of occurrence and associated cyclone name for the Atlantic basin.

Appendix B The peak significant wave height, simultaneous spectral peak period, time of occurrence and associated cyclone name for the Gulf of Mexico-Caribbean Sea.

Appendix-C 5-day error statistics for NAH and WNA modeled significant wave height for All available tropical cyclones at all available buoys.

Appendix-D 5-day error statistics for NAH and WNA modeled wind speed at 10 m height for all available tropical cyclones at all available buoys.



Fig. 1. Locations of NDBC buoys used in model validation



 Fig. 2. Monthly time series of measured and predicted spectral peak period, significant wave height, wind speed and wind direction (from bottom to top) at Buoy 41002.

.



Fig. 3. "Best tracks" and GFDL model tracks for Hurricanes Maria, Nate, and Ophelia.



 Fig. 4. Wave steepness and blended wind fields while Hurricanes Maria, Nate and Ophelia co-existed. (The bottom figure shows from east to west , Hurricanes Maria, Nate and Ophelia).



Fig. 5. "Best tracks" and GFDL model tracks for Hurricanes Katrina, Rita and Wilma.



Fig. 6a. Time history and error statistics of NAH and WNA predicted significant wave height (Hs) and wind speed at 10m height ( $U_{10}$ ) for Hurricane Katrina at Buoy 42040.



Fig. 6b. Same as Fig.6a except for Hurricane Ophelia at Buoy 41002.



Fig.6c Same as Fig.6a except for Hurricane Rita at Buoy 42001.



Fig. 6d. Same as Fig. 6a, except for Hurricane Wilma at Buoy 42056 .



 Fig. 7a Error statistics and linear trend of NAH (in blue ) and WNA (in red ) predicted significant wave height for all tropical cyclones at all buoy sites (Dash lines show the mean values.)



Fig. 7b. Same as Fig. 7a except for the wind speed.



Fig. 8a. Scatter plots of the peak significant wave height (*Hs*, top row), and the associated spectral peak period (*Tp*, bottom row) for NAH (left column) and WNA (right column) for the Atlantic Basin. Legend: W-Wilma; O-Ophelia; K-Katrina; R-Rita; Number-buoy I.D.



Fig. 8b Time lag of normalized bias of the peak significant wave height (top raw) and the associated spectral peak period (bottom raw) predicted by the NAH (left column) and the WNA (right column) models for the Atlantic Basin. In each panel, center lines represent the mean and the outer lines represent the standard deviation. Marker's symbols and colors have the same legend of Fig. 8a.



Fig. 9a. Same as Fig. 8a except for the Gulf of Mexico-Caribbean Sea. Legend: W-Wilma, R-Rita, K-Katrina.



Fig. 9b. Same as Fig. 8b except for the Gulf of Mexico-Caribbean Sea. Legend: Same as Fig. 9a.



 Fig. 10. A comparison of wind (top row) and wave (bottom row) fields predicted by WNA (left col.) and NAH (right col.) for Hurricane Katrina, September 29, 1200UTC.



### Table 1 List of tropical cyclones for the wave model validation study.

Note: Buoy number shown as 41xxx or 44xxx is in the North Atlantic and 42xxx is in the Gulf of Mexico or in the Caribbean Sea (see Fig. 1 for buoy locations).

# **Appendix A**

*Tp***), time of occurrence (month-day-hour as mmddh) and associated cyclone name for the Atlantic basin. The peak significant wave height (***Hs***), simultaneous spectral peak period (**



# **Appendix-A Cont'd**



# **Appendix B**

The peak significant wave height  $(H_s)$ , simultaneous spectral peak period  $(T_p)$ , time of and associated cyclone name for the Gulf of Mexico-Caribbean Sea. height  $(H$ **occurrence nd associated cyclone name for the Gulf of Mexico-C ibbe**



# **Appendix-B cont'd**



### **Appendix-C**

#### 5-day error statistics for NAH and WNA modeled significant wave height (*Hs*, *m*) for All available tropical cyclones at all available buoys. atistics for NAH and WNA modeled significant wave height (*H* **ll** available tropical cyclones at all available buoy



# **Appendix-C (Cont'd)**



# **Appendix-D**

**5-day error statistics for NAH and WNA modeled wind speed at 10 m height (***U10***, m/s) for all available tropical cyclones at all available buoys.** 



## **Appendix-D (Cont'd)**

