Performance of the ocean wave ensemble forecast system at NCEP [1]

Degui Cao[2,3], Hendrik L. Tolman, Hsuan S.Chen, Arun Chawla[2] and Vera M. Gerald

NOAA /National Centers for Environmental Prediction
Environmental Modeling Center
Marine Modeling and Analysis Branch
Camp Springs, Maryland, USA

1. Introduction

An ensemble forecast system for ocean (wind) waves has been developed and implemented operationally at the National Centers for Environmental Prediction (NCEP). The Global Ensemble Ocean Wave Forecast System (GEOWaFS) was initially implemented in April 2004, and was upgraded on May 30, 2008. Henceforth, we will denote the two versions of GEOWaFS as the old and new system, respectively. The old system was based on the NOAA WAVEWATCH III (NWW3) wave models (Tolman et al., 2002, Chen, 2006). It consisted of eleven ensemble members including the operational deterministic NWW3 model as the control member (Chen 2006; Cao et al., 2007). It made 126 hour forecasts for the 00, 06 12 and 18z NCEP model cycles each day. Its spatial resolution was 1^0 in latitude and 1.25^0 in longitude from 78^0S to 78^0N globally. The wind forcing for the 10 ensemble members other than the control run were obtained from NOAA/NCEP Global Ensemble Forecast System (GEFS) 10m wind fields and were updated every three hours. For performance studies, Chen (2006) evaluated two cases of the storms in May through July, 2004, and Chen (2006) and Cao et al.(2007) used the significant wave height (H_S) and the wind speed at 10m height (U_{10}) from nearly 30 deep water buoys for a systematic validation of the ensemble forecasts. They indicated that the ensemble spread increased with the forecast hour. The ensemble forecasts of U_{10} and H_S hit most observation data. The old GEOWaFS was more reasonable and realistic than the deterministic wave forecast. Chen (2006) and Cao et al.(2007) also indicated the system was a better tool for forecasting and decision making than using single deterministic model.

Starting in 2007, the NOAA multi-grid WAVEWATCH III (NMWW3) wave model has been implemented incrementally in the NCEP operational model suite (Chawla et al., 2007, Tolman, 2008). The new multi-grid approach features two-way nesting where grids

[1] MMAB contribution Nr. 279
[2] Science Applications International Corporation (SAIC)
[3] Corresponding author, Degui.Cao@NOAA.gov

with various resolutions become a single wave model, added shallow water physics, but no major updates to deep water model physics. GEOWaFS has transitioned to this new model software, but using only a single global grid with 1°×1° spatial resolution still ranging from 78^0S to 78^0N. The grid modifications were introduced to unify the NCEP wave ensemble with the FNMOC wave ensemble to foster future joining of the two ensembles. Whereas the new GEOWaFS system has transitioned to the new wave model software, this upgrade is expected to have minor impacts on its results. More important are the upgrades to ensemble system itself, which include

1) Extending the forecast horizon to 10 days, conforming to the FNMOC ensemble.
2) Going from 10 to 20 ensemble members conforming to the GEFS. GEFS wind fields are generated using the Ensemble Transform (ET) technique (Wei et al., 2006).
3) Adding a control run forced by the GFS at the ensemble resolution. This is necessary because the high-resolution deterministic global NMWW3 wave model has a forecast horizon of only 7.5 days.
4) Cycle initial conditions of each ensemble member from the previous cycle run of the same ensemble member to generate a natural history of perturbations of swell. In the old system, each member started from the same (deterministic) initial conditions, effectively removing swell perturbations.
5) Use bias-corrected GEFS winds instead of 'raw' GEFS winds.

The new GEOWaFS was implemented operationally on June 1, 2008. The main outputs of the system (other than output of the individual members) include the ensemble mean, spread, spaghetti diagram and probability at different thresholds of wind speed, wave height and peak period.

The main purposes of the present study are to evaluate the improvement of the new GEOWaFS comparing to the old system, and to conduct statistical analyses on the performance of the new system.

2. Buoy data and analysis methodology

2.1 Buoy data

A total of 37 buoys were used in the evaluation study of the old GEOWaFS (Fig. 1). For the new system wind and wave data from 145 buoys are used for the period from July 1, 2008 to June 30, 2009 are used (Fig. 1). The selected buoys are mainly located in the northern hemisphere except for one buoy located on the equator. The hourly and quality controlled buoy data are compared with the ensemble output each hour. Wind and wave climatological data are generated using the data from January 1, 2002 to July, 2009. These data are averaged at each buoy location without smoothing at larger time scale.

2.2 Analysis methodology

Talagrande histograms are constructed from the unmodified ensemble data and are used to assess bias and dispersion characteristics of the ensemble system. It is constructed from the notion that an ideal ensemble system will correspond to a verification analysis that is equally distributed between any two ordered adjacent ensemble members, including the cases when the analysis will be outside the ensemble range on either side of the distribution. Therefore the Talagrande diagram should be flat in the ideal ensemble systems. In reality the distribution is slightly U-shape. The U-shape indicates the ensemble does not spread out sufficiently due to over-representation. Sometimes the Talagrande histograms look like J-shape which indicates the ensemble system has a bias.

Spaghetti plots, spread, the root mean square error (rmse), bias and mean have been used in the analyses. Their definitions can be found in many mathematical books.

3 Ensemble analysis and comparison

3.1 Comparison of the old and new GEOWaFS systems

Figure 2 presents the spreads of the old and new GEOWaFS systems and rmse of the control run in the current operational system from March 1, 2008 to June 30, 2008. Only the 37 buoys used in the old system (Fig. 1) to retain consistency between system and analysis reliability. The spread of the old system is only available up to the 120 hour forecast. For the ensemble to realistically capture the uncertainty in the deterministic forecast, the spread of the ensemble should be equal to the rmse of the deterministic forecast. For the GEFS winds, the ensemble spread is somewhat smaller that the model error, but grows similarly as a function of forecast hour. This behavior is typical for many ensemble systems. For the old GEOWaFS system, the spread of the ensemble is severely underestimated. This is clearly related to the zero spread imposed at the 0^{th} forecast hour. For the new system, the spread is greatly improved, with the wave height spread showing similar behavior as the wind speed spread. This behavior can be attributed to the cycling of initial conditions as introduced in the new system. Figure 3 indicates that the introduction of the cycling of initial conditions represents a massive improvement of the GEOWaFS system.

As an example, Fig 3 presents the spread and mean of the wave heights for the old and new systems for the 00, 48 and 120 forecast hour on March 28, 2008 (06z model cycle). For the old system, the spread at the nowcast hour is by definition zero. For the new system, initial wave fields of the twenty members of wave forecast are obtained from a restart file from the same member of the ensemble at the 6h forecast of the previous model cycle. For both systems the spread increases with the forecast time. The spread in storm track areas is largest and generally coincides with the largest mean wave heights. The wave height spread in the new system at the 00h forecast is larger than the spread of the old system at the 48h forecast. This indicates that the new system has a `memory' of at least 2 to 3 days with respect to the building up of spread in the system. Since the model is cycled every 6h, the building up of the wave model spread indicates that there is

high consistency between wind perturbations between model cycles. This appears essential to generate consistent wave height perturbation of appropriate magnitude. Note that in swell dominated areas such as the equatorial Pacific Ocean, the old system at the 120h forecast still has less spread than the new system at the 0h forecast. This could be expected because swell travel times from storm tracks to the Equator can be as much as 7 days. Because the swell perturbations in the old system were generated exclusively during the forecast cycle, no significant swell perturbation can reach the Equator even after five days of running the model. Unlike the spreads, the global means are nearly the same in the two systems.

This example is representative for the ensemble behavior. Combined with the spread as discussed in the previous section, it clearly shows the superiority of the new system. The example also indicates that the impact on swell is even bigger than the impact on the dominant wind seas. This makes it interesting to separately assess the spread of wind seas and swell in future validation of GEOWaFS.

3.2 Spaghetti diagram

An example of a spaghetti diagram for buoy 42001 on Feb. 21, 2009 (00z cycle) is presented in Fig. 4. It displays 20 ensemble members, the control run, the ensemble mean and the observed data up to the 240 hour forecast. The ensemble mean is close to the observation for all forecast hours. The wind speed (U_{10}) ensemble diversity increases after the 96 forecast hour, indicating an increase in wind speed uncertainty. The diversity of the significant wave height (H_s) also increases, but with a small but notable time lag. The time lag is likely due to the fact that the wave model can be considered as an integrator over space and time of the wind forcing.. The control run (U_{10} or H_s) is closer to the observation in the first 5 days of the forecast, but has a larger bias compared to the observation later in the forecast.

3.3 Bias, spread and rmse

Figure 5 presents the monthly U_{10} (upper) and H_s (lower) bias (model-observation) as a function of the forecast hour. The good performance of the wave ensemble system is evident in the modest H_s bias for the entire forecast period. All H_s biases are within -0.1m to 0.1m except for the H_s bias in Jan. (from 0.15m to 0.2m), in July around 0.2m from 0 forecast hour to 120 hour and in Aug. at about -0.2 m from 168 forecast hour. The U_{10} bias fluctuates from -0.7 m/s to 0.6m/s. The spread of the U_{10} bias increases after 72 forecast hour, but the spread of the H_s bias does not show obvious increases. The smaller spread in biases for the wave height again may tentatively be attributed to the fact that the wave model can be considered as a large scale (space and time) integrator of the wind fields, which is less sensitive to sampling of large variabilities at single observation points than the local wind.

Figure 6 presents the monthly U_{10} (upper) and H_s (lower) spread as a function of the forecast hour. Small spread indicates low predictability uncertainty, while large spread indicating high predictability uncertainty. At the same time spread indicates how far into the forecast the ensemble forecast can carry informative events related to the forecasting. U_{10} and H_s spreads increase with the forecast hour. The U_{10} (H_s) spread in the boreal winter is larger than that in the boreal summer because there are larger winds and waves in the winter over the open ocean. The monthly U_{10} spread does not have a direct and close relationship with the monthly H_s spread.

Figure 7 presents the monthly U_{10} (upper) and H_s (lower) rms errors (rmse). The rmse is of the standard deviation for the difference between the model forecast and its observation. It is a measure of forecast accuracy. The U_{10} and H_s rmses increase over the time. The rmse in the winter is larger than in the summer because U_{10} and H_s are larger in the summer.

3.4 Talagrande histogram

Figure 8 presents the Talagrande histograms of H_s. Because there are 20 members in the current GEOWaFS, there are 21 intervals including the intervals at the edges of the distribution (x-axis). The y-axis indicates the number of events (relative to the total number of events). The analyses from the H_s are found between two ordered adjacent members on the x-axis. For the example forecast hours, the histograms show an U-shap due to over representation of events when the verification falls outside the ensemble and under representation of events when it falls inside the ensemble range of predictions. This indicates that the ensemble spread is smaller than the actual model uncertainty, consistent with the model errors presented in Fig. 2.

4. Discussion and conclusion

The new GEOWaFS ensemble forecast system shows a massive improvement over the old system. This improvement becomes evident in the much more realistic ensemble spread as indicated in Fig.2. Tentatively, this improvement can be attributed to the fact that cycling of initial conditions with consistent forcing perturbations between cycles naturally generates swell perturbations. Implicitly, it is thus assumed that the swell perturbations are dominated by uncertainty in their generation, and not in the physics of swell propagation. Whereas initial results are encouraging, more detailed analyses are needed. Particularly, we intend to assess swell and wind sea perturbations and uncertainties separately in detail in this ongoing study.

The new GEOWaFS system shows encouraging performance. The monthly bias is generally between -0.1m to 0.1m. The wave height spread increases with the (seasonal) wave conditions as would be expected. Example spaghetti diagrams show a significant

increase of information content for the ensemble compared to a deterministic model run. A major challenge still is how to use such data in the day-to-day forecast practice.

As mentioned above, the intent is to combine the NCEP wave ensemble to the Navy's (FNMOC) ensemble. Most of the technical work for combining the ensemble has been done, particularly, unification of model grids. FNMOC is already receiving real time data from NCEP, NCEP is in the process of establishing the corresponding data stream from FNMOC. Combining the two ensembles will pose some interesting scientific questions, particularly relating to generating initial model perturbations (internal to the modeling system at NCEP, explicit to ensemble to be developed at FNMOC), and the effects of consistency of wind field perturbations from cycle to cycle (maximized at NCEP, minimized at FNMOC). In the combination of the ensembles, the individual representations of wind sea an swell perturbations is also expected to be essential to obtain a solid understanding of the two systems individually and combined.

5. References

Cao, D., H. S. Chen and H. L. Tolman, 2007: Verification of ocean wave ensemble forecast at NCEP. 10th International Workshop on Wave Hindcasting and Forecasting & Coastal Hazards Symposium, Turtle Bay, Oahu, Paper G1.

Chawla, A, D. Cao, V. M. Gerald, T. Spindler, and H. L. Tolman, 2007: Operational implementation of a multi-grid wave forecasting system. 10th International Workshop on Wave Hindcasting and Forecasting & Coastal Hazards Symposium, Turtle Bay, Oahu, Paper B3.

Chen H.S, 2006: Ensemble prediction of ocean waves at NCEP. Proceedings of 28th Ocean Engineering Conference in Taiwan, NSYSU.

Tolman H . L., B. Balasubramanian, 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.

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

Wei M., Z. Toth, R. Wobus and Y. Zhu, 2006, Initial perturbations based on the ensemble transform (ET) technique in the NCEP Global forecast system, U.S. Department of Commerce, NOAA/NCEP office note 453, 33pp.

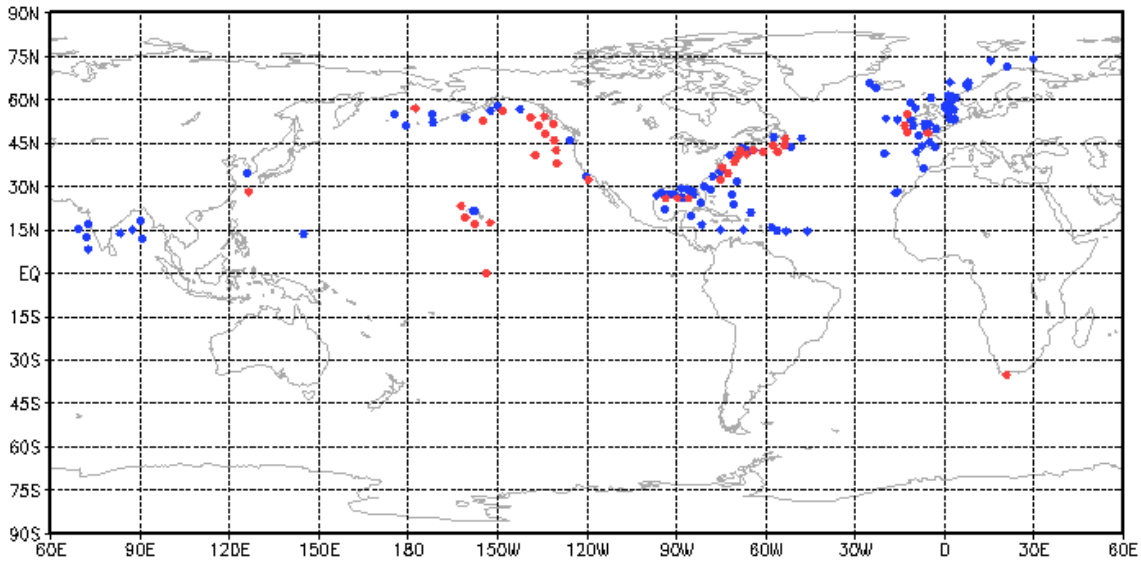

Fig. 1 Locations of buoy data used for validating GEOWaFS. Red: buoys used for ofd and new system. Blue: buoys used for new system only'

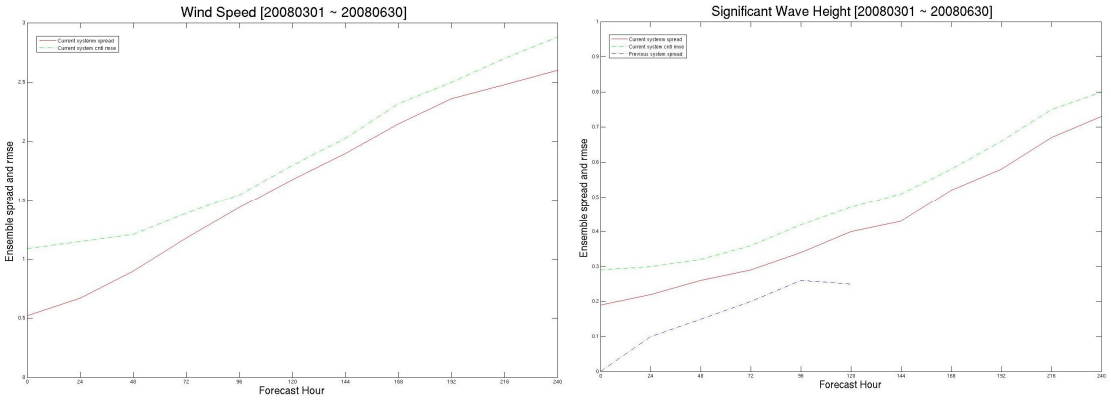

Fig. 2 rms error of control run (green lines) and ensemble spread of the new (red lines) and old ensemble systems for wind speeds (left) and wave heights (right).

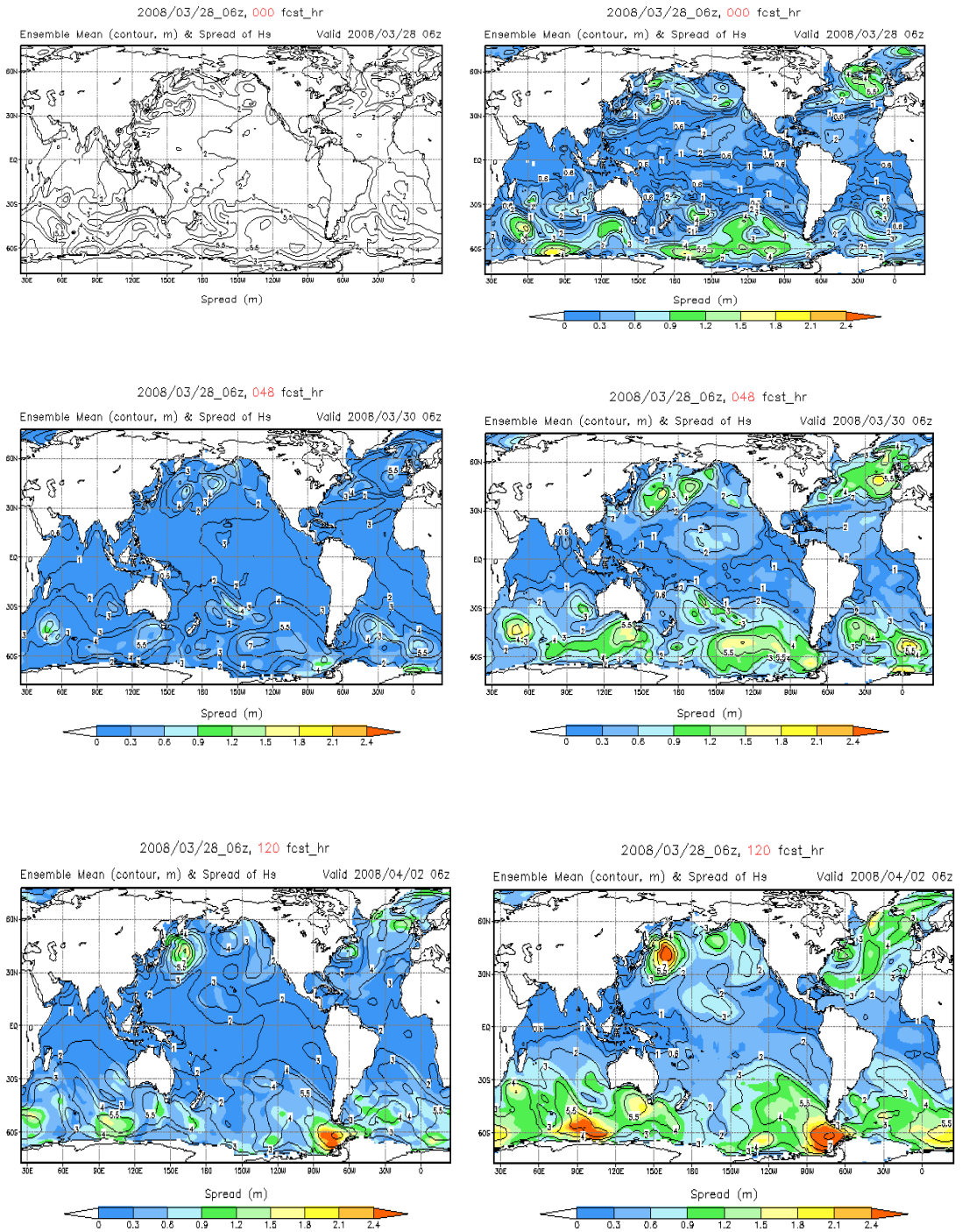

Fig.3 Global ensemble spread (shaded) and mean (contours) of the significant wave height for the old system (left panels) and the new system (right panels) for the nowcast (top panels), 48h forecast (center panels) and 120h forecast (bottom panels) for the March 28, 2008, 06z model cycle.

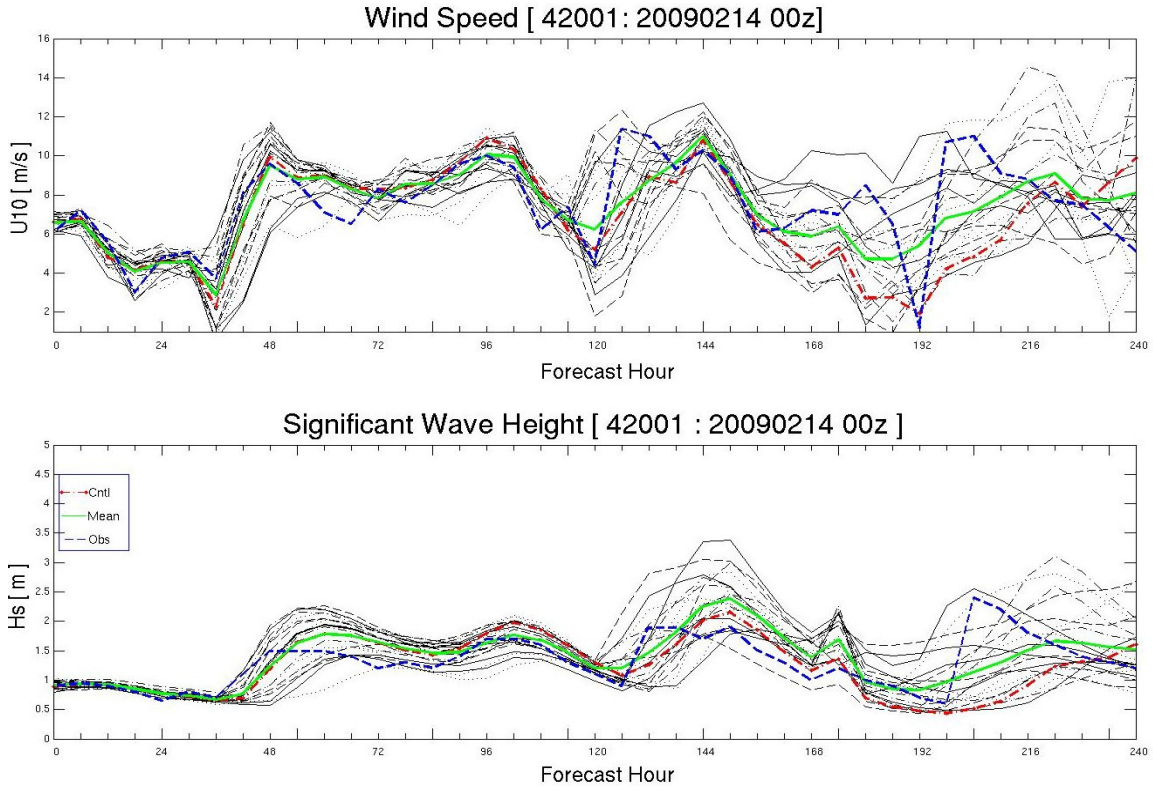

Fig. 4 Example of wind speed (upper panel) and wave heigh (lower panel) spaghetti diagrams for buoy 42001 in the Gulf of Mexico for the February 14, 2009 00z model cycle.

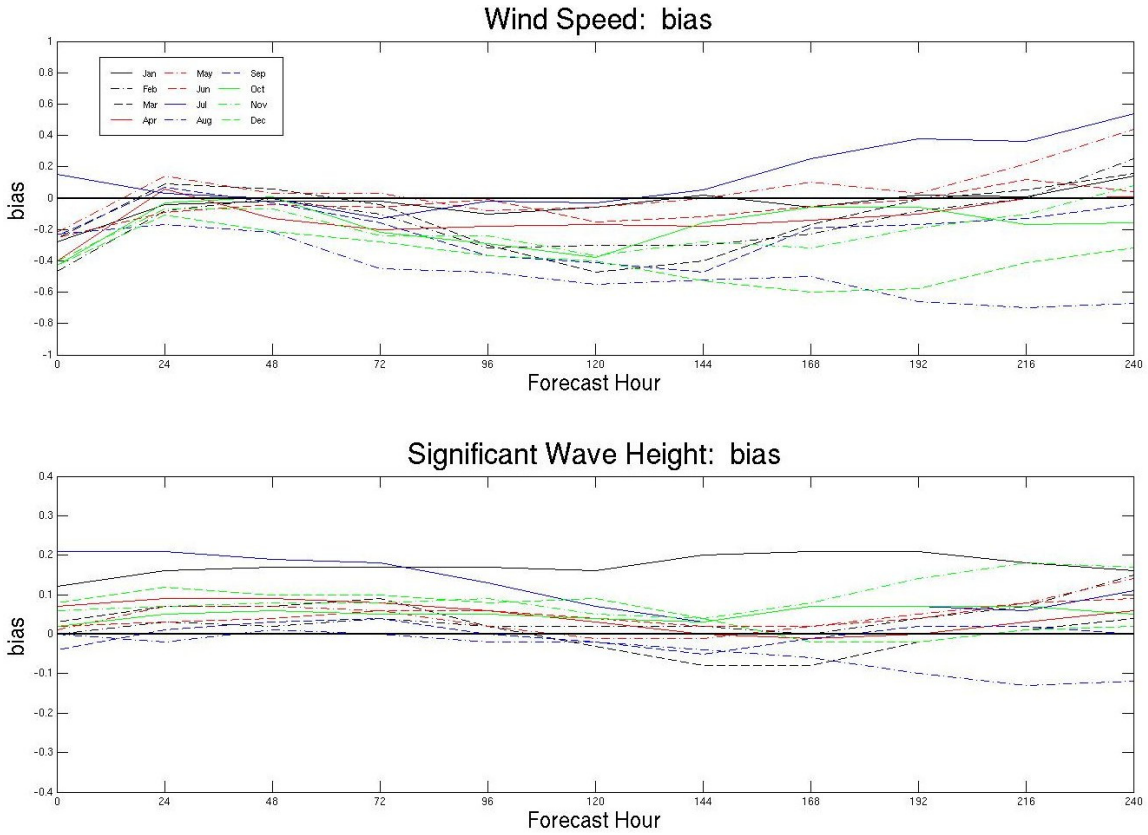

Fig. 5 Monthly wind speed and wave height biases.

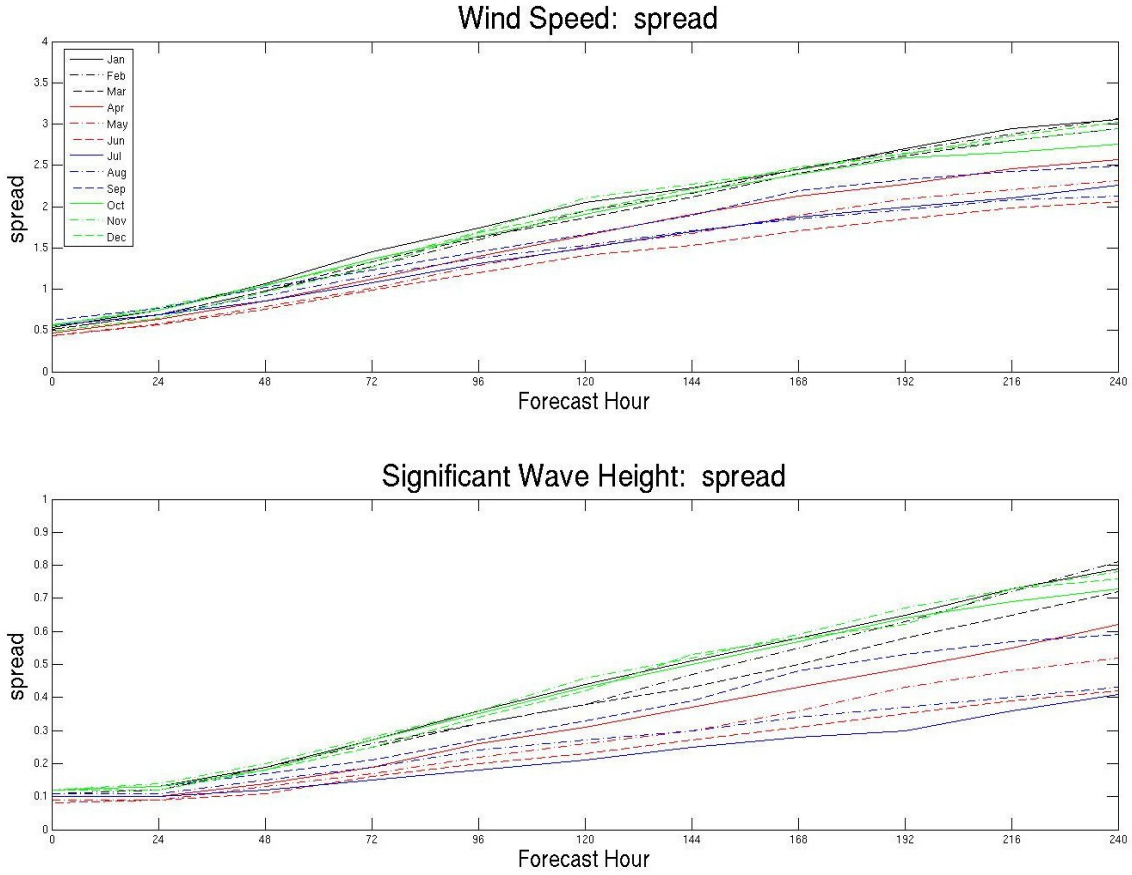

Fig. 6 Monthly wind speed and wave height spreads

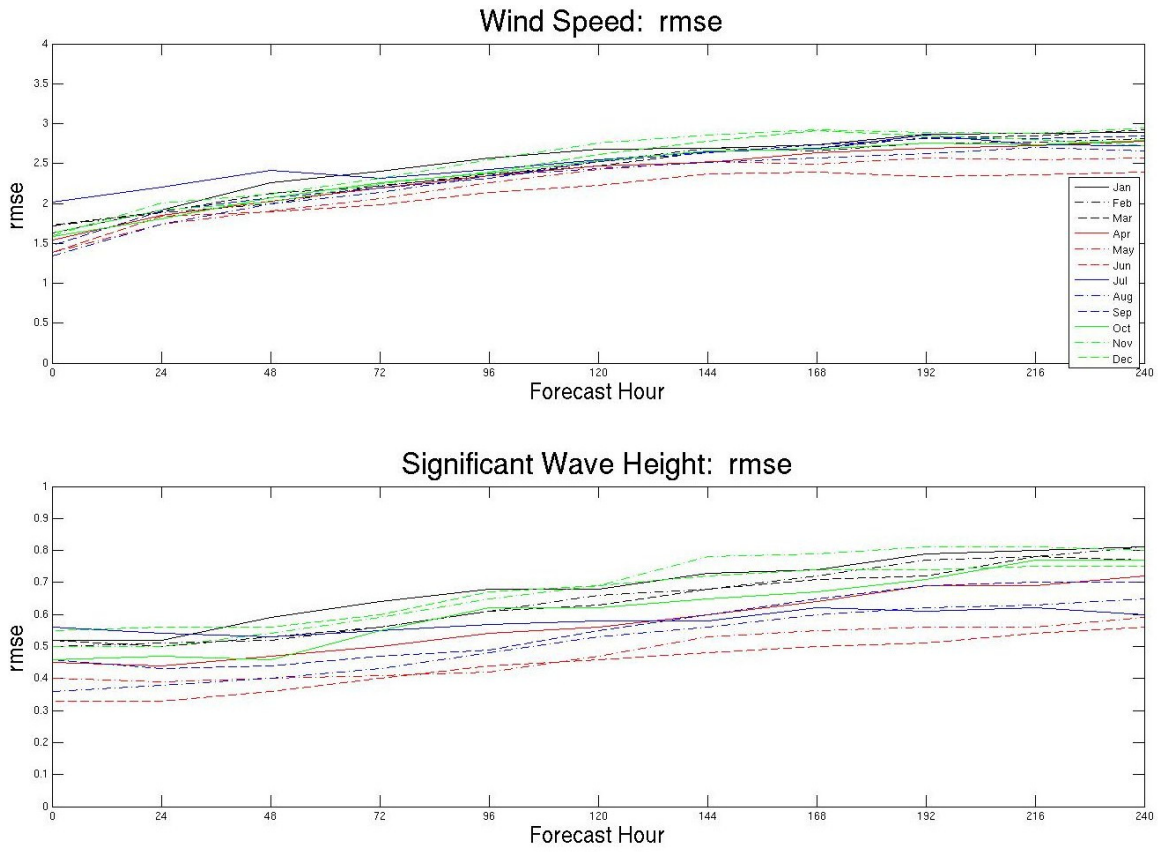

Fig. 7 Monthly wind speed and wave height rms errors.

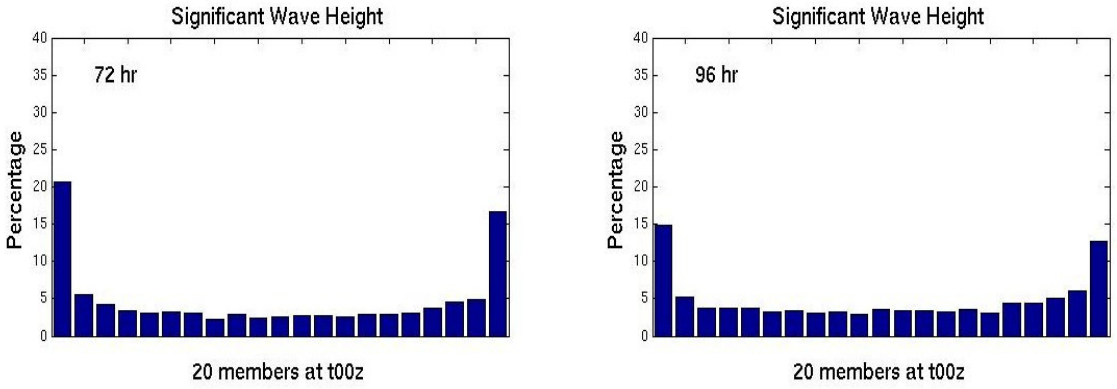

Fig. 8 Talagrande histogram for wave height for 72 and 96 hour forecast from 00z model cycle.