Joint Hurricane Testbed Project Final Report Supplement
February 5, 2010

Project title: *An Improved Wind Probability Estimation Program*
Principal Investigators: Stan Kidder, Mark DeMaria and Pat Haar
Affiliation: Kidder (CIRA), DeMaria (NESDIS) and Haar (NPS)
Project dates: August 2007-July 2009
TPC Point of Contacts: Dan Brown (replacing Rick Knabb) and Chris Lauer

1. Background

The final report for this project was provided to the JHT in December 2009. At that time, only the NHC and JTWC working best track datasets were available, so the 2009 experimental runs of the Monte Carlo (MC) wind probability with the GPCE uncertainty information included were not verified. As of Feb. 1st, all of the Atlantic best tracks were completed and the eastern North Pacific best tracks were nearly complete (three storms needed final approval, but all others were finalized). After coordination with JTWC, the western North Pacific best tracks will not be available until April, so the probability verification was performed with the best tracks available on Feb. 1st. To determine the sensitivity to final versus preliminary best track information, the verification results for the eastern North Pacific were compared with the working best tracks and the nearly complete set of final best tracks, and the changes were very small. Thus, the verification results presented here should be representative of the results with the final best tracks.

2. Verification Procedure

As described in the final report, the verification for the 2008 Atlantic model runs was performed on a series of coastal points similar to that used in the NHC text product. This provided a large sample because of the large number of storms near the U.S. coast that year. However, the Atlantic storms in 2009 were generally not near the U.S. coast, so that method was inadequate. In addition, the verification in 2009 includes the eastern and western North Pacific. For these reasons, the verifications for 2009 were performed on basin-wide evenly spaced 1 degree lat/lon grids. Figure 1 shows the cyclone tracks and verification domains for each basin. The domains for the east and west Pacific are slightly larger than described in the final report because of the inclusion of late season cases in the west Pacific and to include the full length of the best track positions.

The verification from 2008 described in the final report showed that the Brier score and threat score were good metrics for comparing the GPCE and operational versions of the MC model at the coastal points. However, for the domains shown in Fig. 1, there are large regions where the probabilities are zero and where there were no winds of the specified speeds (34, 50 or 64 kt). These large "null" regions make the Brier score artificially small, which dampens the differences between the two versions of the model. In contrast, the threat score measures the percent overlap between the area enclosed within a specified probability threshold for a specified wind speed and the area where that wind was observed. The TS was calculated for every probability threshold from 1 to 99%

and then averaged. The TS is not affected by the large null regions of the domain, and so is a better statistic for comparing the two versions of the MC model. It emphasizes the larger probabilities that are more relevant for applications of the probability products.

Both versions of the MC model were run for every forecast case from 2009 for all three basins. This sample included the full life cycles of 11 Atlantic storms, 20 east Pacific storms and 28 west Pacific storms. These include 149, 333 and 519 forecast cases with at least a 12 h verification, respectively.

Unfortunately, it was not possible to perform the comparison for the three storms that began in the central Pacific because the GPCE values were not generated for these storms. The sample does include the portions of storms that began in the eastern Pacific, but moved into the central Pacific. If the GPCE version of the MC model was implemented in operations, GPCE values would be needed for the central Pacific storms. One option would be to use the east Pacific algorithm.

The verification grids were constructed from the storm positions and wind radii in the NHC and JTWC best track files. The 6 hr values were linearly interpolated to 1 hr intervals to determine which points in the domain were impacted by each wind threshold (34, 50 and 64 kt). Because the underlying track and intensity error distributions used by the MC model were determined by a verification of the full best track, including the extratropical, subtropical and dissipating stages, these stages were also included in the verification of the probabilities. Thus, any time period with an NHC or JTWC forecast and corresponding best track information was included in the MC model verification. The MC model runs and the verification grids were constructed on a storm by storm basis.

3. Verification Results

Figure 2 shows the percent improvement of the Brier and threat scores for the GPCE version of the MC model, relative to the operational version, for the Atlantic cases. The verification was for the cumulative probabilities at 12 hr intervals. The results for the incremental probabilities were generally similar, but noisier. The GPCE version improved the verification results for every wind radii at every forecast period. The improvements were generally higher for the 64 kt wind radii, and for the longer range forecasts. Also, as expected, the percentage improvements were larger for the threat score than the Brier score because these are less impacted by the large null regions in the domain.

Figure 3 shows the percent improvement for the eastern North Pacific cases. The Brier score improvements were smaller than for the Atlantic for the 50 and 64 kt wind probabilities and were slightly negative for the 34 kt wind probabilities. The threat scores, which are more representative of the higher probability regions, showed improvement for every wind threshold at every forecast period, although the percent improvements were about half of those for the Atlantic. Similar to the Atlantic, the improvements were largest for the 64 kt probabilities, and tended to be larger for the longer forecast intervals.

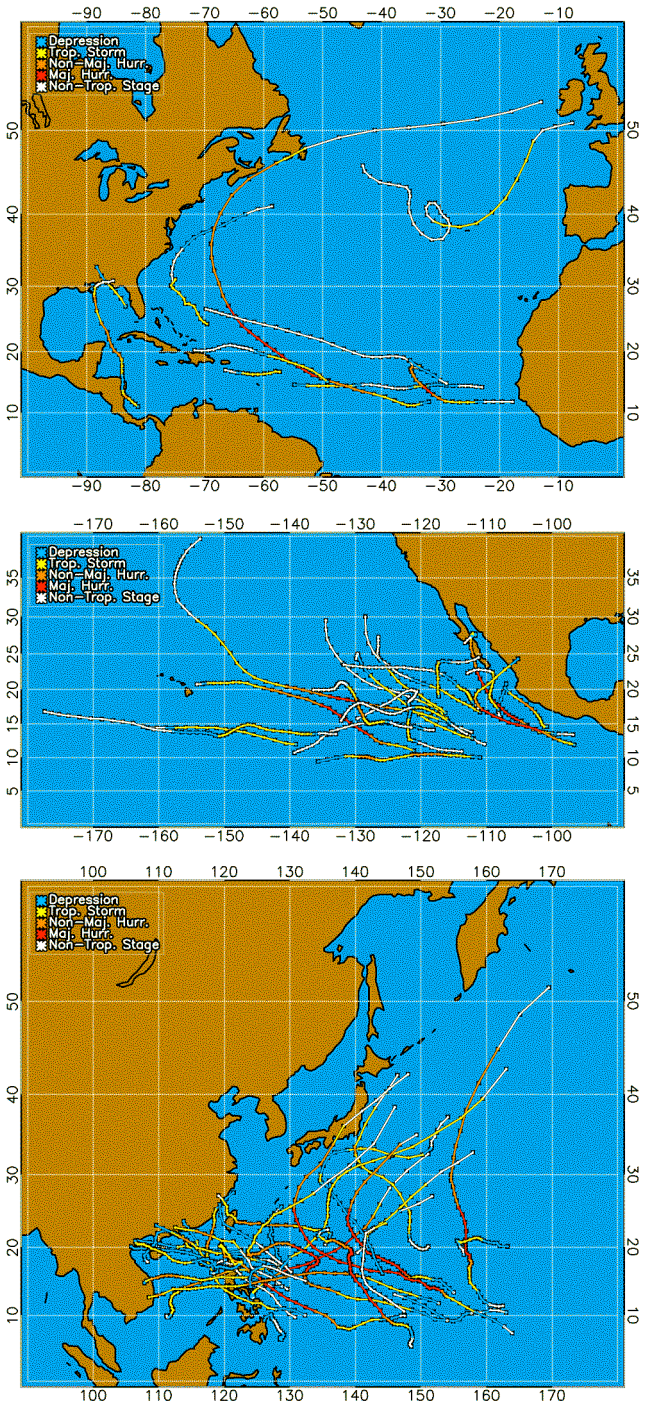

Figure 1. The tracks of all tropical cyclones included in the 2009 evaluation of the MC model for the Atlantic (top), eastern North Pacific (middle), and western North Pacific (bottom). The computational domain for the model runs is the same as the domains in each plot.

Figure 4 shows the percent improvement for the western North Pacific cases. The results are about halfway between those for the Atlantic and east Pacific. The Brier scores

improvements are small and mostly positive. The threat score improvements are much larger and all positive.

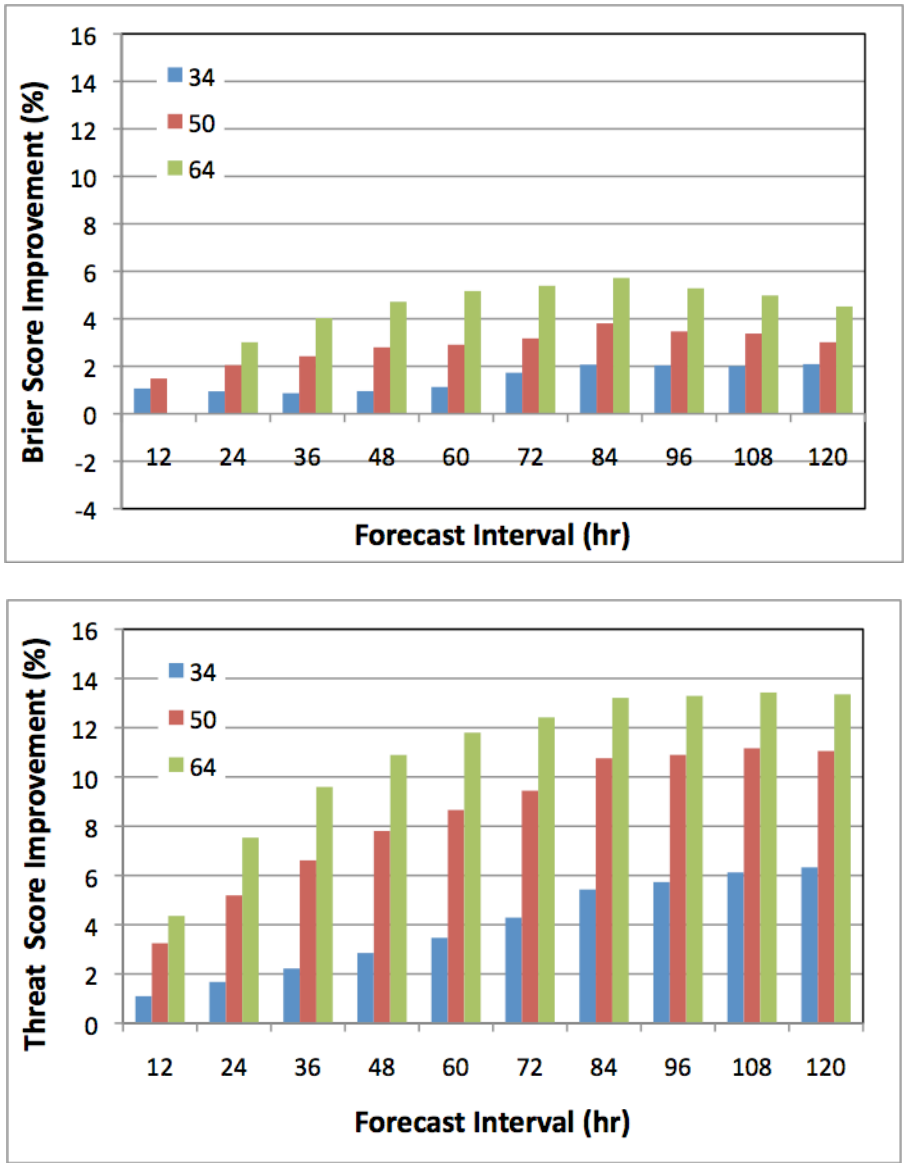

Figure 2. The percent improvement in the Brier score (top) and average threat score (bottom) for the 2009 Atlantic MC model forecasts with the GPCE input relative to the operational version.

4. Summary

The 2009 verification results showed that the GPCE input resulted in an improvement in the threat score in all basins at all forecast times for all wind thresholds. The improvements were largest in the Atlantic and smallest in the east Pacific. The results were a mixed for the Brier scores, although in the majority of cases, there was a small positive improvement. Based on these results, and the 2008 verification results for the Atlantic coastal points, the GPCE information improves the MC model probabilities.

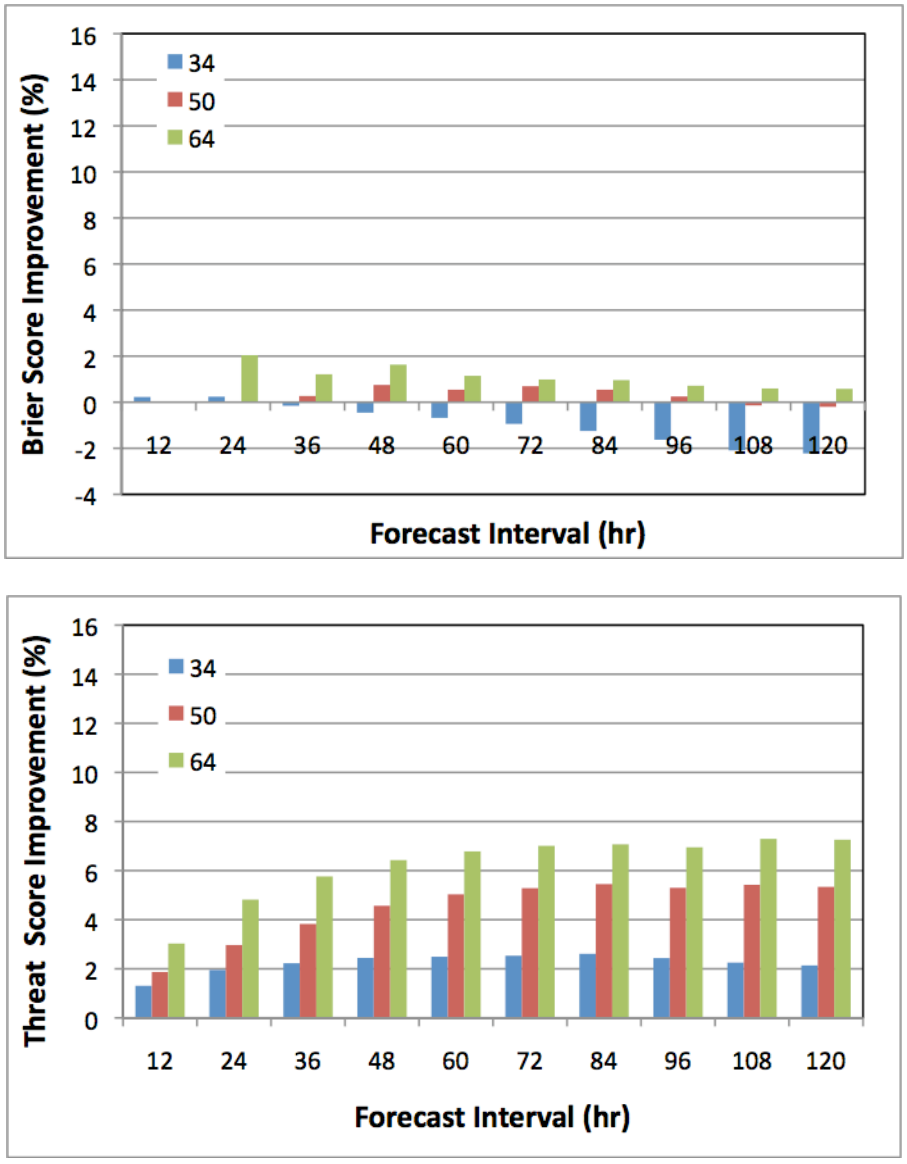

Figure 3. The percent improvement in the Brier score (top) and average threat score (bottom) for the 2009 east Pacific MC model forecasts with the GPCE input relative to the operational version.

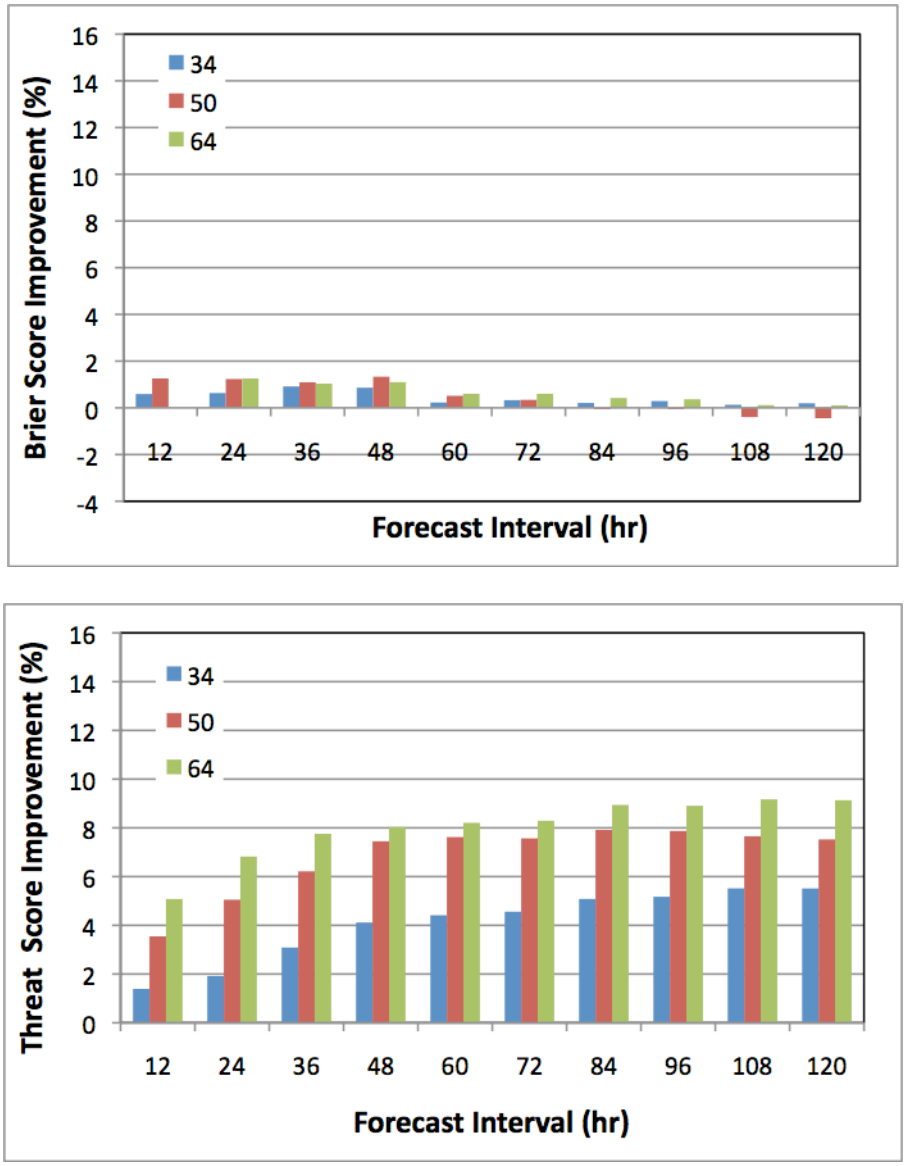

Figure 4. The percent improvement in the Brier score (top) and average threat score (bottom) for the 2009 west Pacific MC model forecasts with the GPCE input relative to the operational version.