Joint Hurricane Testbed Project Final Report
December 2009

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 Information

Under previous JHT support a new program for estimating the probability of occurrence of 34, 50 and 64 kt winds was developed. A Monte Carlo (MC) method was utilized to combine the uncertainty in the track, intensity and wind structure forecasts. The operational version of the MC model is described in detail in DeMaria et al. (2009a).

In the current proposal, three improvements were proposed to the MC model, as follows:

Topic1: The MC wind probability estimates will be refined by making the underlying track error distributions a function of the forecast uncertainty. The current MC model uses basin-wide error statistics but recent research has shown that the spread of track forecasts from various models can provide information about the expected track error. J. Goerss from NRL developed a real-time tool to quantitatively estimate the track forecast uncertainty (the Goerss Predicted Consensus Error, GPCE), which will be incorporated into the MC model.

Topic 2: The timeliness of the MC model will be improved by optimizing and modifying the code.

Topic 3: The code that calculates the track and intensity error distributions for the MC model will be generalized to update the "stand-alone" intensity probability product utilized by NHC. This product is provided in real time as the "wind speed probability table" on the NHC web site, and was developed from data from 1988-1997. The current version of this product only extends to 72 h even though the NHC official forecasts were extended to 120 h in 2003.

The timeline and deliverables for Year 2 of this project are listed below in the Appendix.

2. Accomplishments

Topic 2 was completed in Year 1 of the project and resulted in a speed-up by a factor of six of the MC model code. The optimized code provides identical results and was already implemented on the NCEP IBM.

Topic 3 was completed early in Year 2 of the project. After successful testing and evaluation, the final task was to provide NHC with a modified version of the code that returned all of the information for the wind speed probability table. This code was provided to C. Lauer from NHC prior to the start of the 2008 hurricane season, and was run for the 2008 and 2009 seasons. D. Brown (NHC) and R. Knabb (Honolulu WFO) are performing a verification of the table values.

The emphasis in Year 2 was on Topic 1 above. A method to stratify the Atlantic NHC track errors by the GPCE parameter was developed in 2008 and it was confirmed that the distributions have a well-behaved dependence, with wider distributions for the larger GPCE values. This initial analysis was performed with the 2002-2006 sample used in the 2007 MC model. For the 2008 testing, the track error distributions for the 2003-2007 sample were stratified in a similar manner, with similar results.

Given that the NHC track errors generally are larger when the GPCE values are larger, the MC probability model was modified to use track error distributions that depend on the GPCE value. The NHC track errors from the previous five years were divided into terciles (low, medium and high) based on the corresponding GPCE values. In real-time the track realizations are constructed by randomly sampling from the error distributions based on the real-time values of the GPCE input. The code accounts for case where the GPCE tercile changes with time along the forecast track. For example, the GPCE value might be in the low category for the early part of the forecast, but transitions to medium or high at later times. The method to account for serial correlation in the errors provides smooth tracks even when the GPCE category changes along the track. The code also accounts for the case when the GPCE value is missing. In that case, the track error distributions for the full 5-year sample are used.

When the GPCE category is high, the wind probability distributions tend to spread out, with lower probabilities along the track and higher probabilities farther from the track. The opposite is true when the GPCE category is low. For medium GPCE values, the probabilities vary only slight relative to the version where the full distributions are used. Figure 1 shows an example of the difference between the GPCE and operational versions of the MC model for the case were nearly all the GPCE values were in the low tercile. The probabilities along the track increased by more than 10% and the values away from the track decreased by more than 7%.

As a first test of the new model, about 160 cases from the 2008 Atlantic season that were within about 1000 km of the U.S. coast were re-run using the GPCE and operational versions of the MC model. For the qualitative evaluation, a web site was created to display the probabilities over a large domain, similar to that used in the graphical products on the NHC web page (see http://rammb.cira.colostate.edu/research/tropical_cyclones/tc_wind_prob/gpce.asp). The page shows the probabilities for both the operational and GPCE versions of the model and their differences. To facilitate the comparison, the MC model was run on a 0.25 degree lat/lon grid, rather than the 0.5 degree grid used for the NHC products.

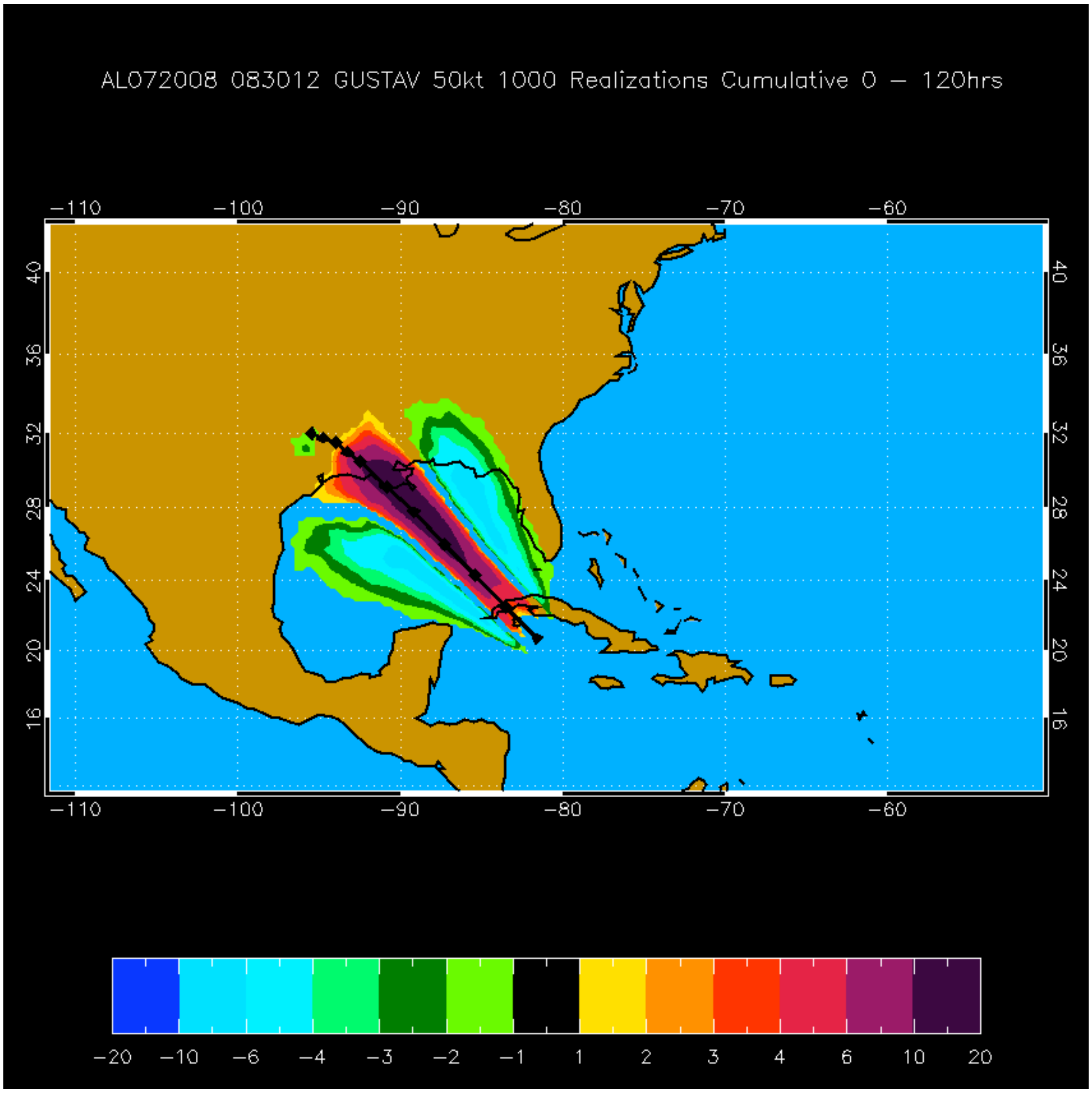

Figure 1. The difference between the GPCE and operational 0-120 h cumulative 50 kt wind probabilities for Hurricane Gustav beginning on 30 Aug 2009 at 12 UTC. The GPCE values were mostly in the low tercile for this case.

For the quantitative comparison of the operational and GPCE versions of the model, the 2008 cases were also run for the contiguous U.S. anc Caribbean coastal breakpoints, similar to those used in the NHC operational text product. This allowed an evaluation for those cases most relevant to U.S. watches and warnings, and over a more focused region. The verification code was adapted to the case where the probabilities are on an irregularly spaced set of lat/lon points (the coastal breakpoints) rather than on a regular grid.

The Brier Score and Threat Score were used to evaluate the new model. Both measures showed that the GPCE version of the model is an improvement over the operational version. For example, Fig. 2 shows the percent improvement of the Brier Score for the cumulative probabilities. The Brier Score improved for all radii at all forecast intervals.

The results for the incremental probabilities were similar. Threat Scores for the GPCE version of the model were also improved for the cumulative probabilities at nearly all time intervals, and more than half of the forecast intervals for the incremental probabilities. Further details of the 2008 verification can be found in DeMaria et al. (2009b).

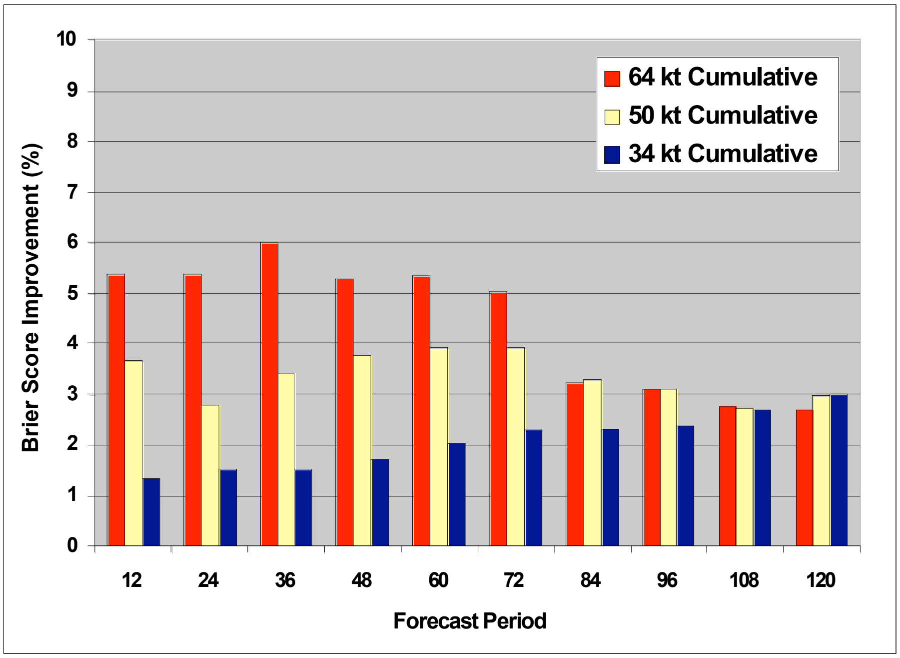

Figure 2. The improvement in the Brier Score of the 34, 50 and 64 kt wind probabilities of the GPCE version of the MC model relative to the operational version. The sample includes 169 Atlantic cases from 2008 that were within 1000 km from land, and the probabilities were calculated at 257 breakpoints along the U.S. Gulf and east coasts and the Caribbean.

Following the successful test in 2008, GPCE versions of the MC model were developed for the 2009 season for the Atlantic, eastern North Pacific and western North Pacific. The track error distributions for each basin from the previous 5 years (2004-2008) were stratified by the GPCE values and the results were very similar to the 2008 Atlantic model. Figure 3 shows the boundary values between the 1st and 2nd GPCE terciles and between the 2nd and 3rd terciles for each basin. The Atlantic values are generally the largest and the eastern North Pacific values are the smallest.

To test the 2009 GPCE version of the MC model, all available cases from each basin were run with the operational and experimental versions of the model. A web page was created that compares the 0-120 h cumulative probabilities for the GPCE and operational version of the model (see http://rammb.cira.colostate.edu/research/tropical_cyclones/tc_wind_prob_2009). Most cases on the web page were run after the events. However, an automated system for running the MC model was implemented near the end of the 2009 season, and the

Atlantic cases for Tropical Storm Henri and Hurricane Ida, the east Pacific cases from Tropical Storm Patricia and Hurricane Rick, and west Pacific cases for Typhoon Nida were available in near real time (normally by about 15 minutes after advisory time) for forecaster evaluation. Some forecaster feedback (from James Franklin and Dave Roberts) was obtained on the real time runs, mostly regarding interpretation of the results.

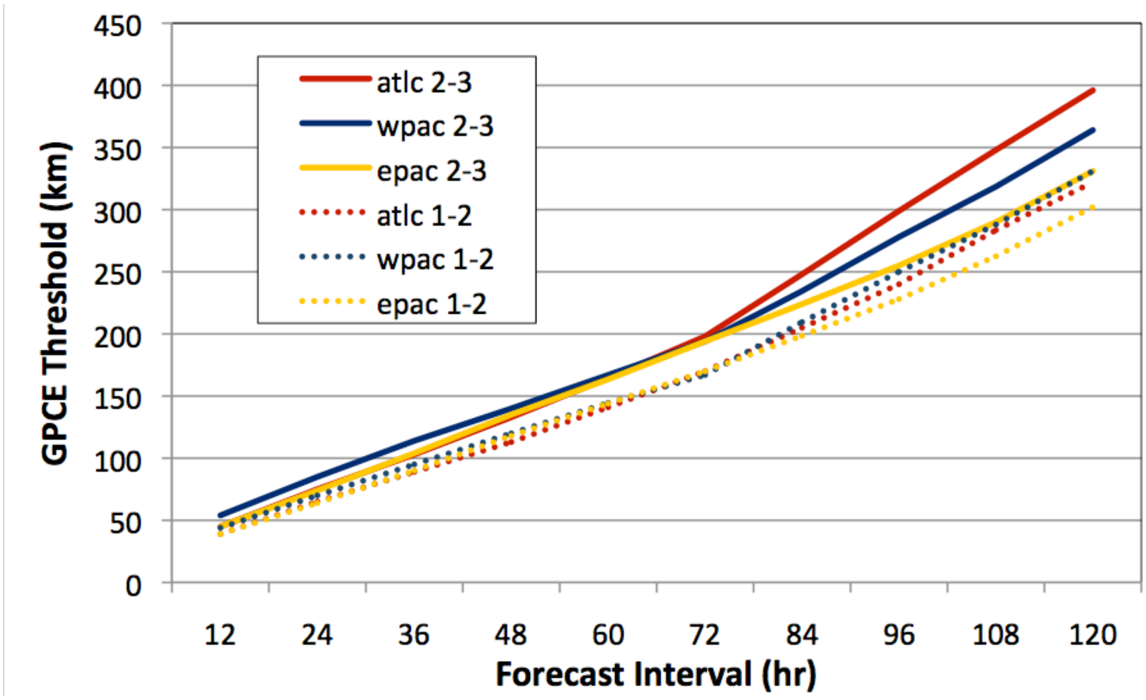

Figure 3. The values(km) that separate the GPCE values into terciles (1 to 2 and 2 to 3) for the Atlantic, eastern North Pacific and western North Pacific for the 2004-2008 sample.

The GPCE versions of the MC model for all basins in 2009 behaved very similar to that for the Atlantic in 2008. The probability distributions spread out when the GPCE values are large and become more confined to along the track when the GPCE values were small. For example, Fig. 4 shows the difference between the GPCE and operational 0-120 hr cumulative 34 kt wind probabilities for Typhoon Linfa for a case where the GPCE values were in the high tercile for most of the forecast. This figure shows that the probabilities decreased along the track and increased away from the track.

The 2009 hurricane and typhoon season provided a variety of cases for testing the GPCE version of the model. Figure 5 shows the preliminary best tracks for all storms in each basin (including the central Pacific). For the Atlantic the sample size was much smaller than in 2008 and very few cases were near land. Thus, the verification procedure used for the 2008, where the probabilities were evaluated at breakpoints along the U.S. coast and Caribbean, is not appropriate. Therefore, an alternate procedure was developed. Both versions of the MC model for all individual storms were run on a large domain, roughly the size of that shown in Fig. 5. The data on this large domain will be verified using the same metrics as in the 2008 evaluation (Brier Score and Threat Score) as soon as the final best tracks are available. The verification code was modified to eliminate points where the probabilities are zero from the MC model and no winds were observed, because those "no-no" points dominate the statistics on the large domain. Those points do not impact the Threat Score, but make the Brier Score artificially small (favorable). Similarly for the eastern North Pacific, very few cases are near land, so a large-domain verification will be performed. The coastal breakpoint verification could be used for the western North Pacific since many of those storms are near land. However, for consistency with the other basins, the large-domain verification will also be used.

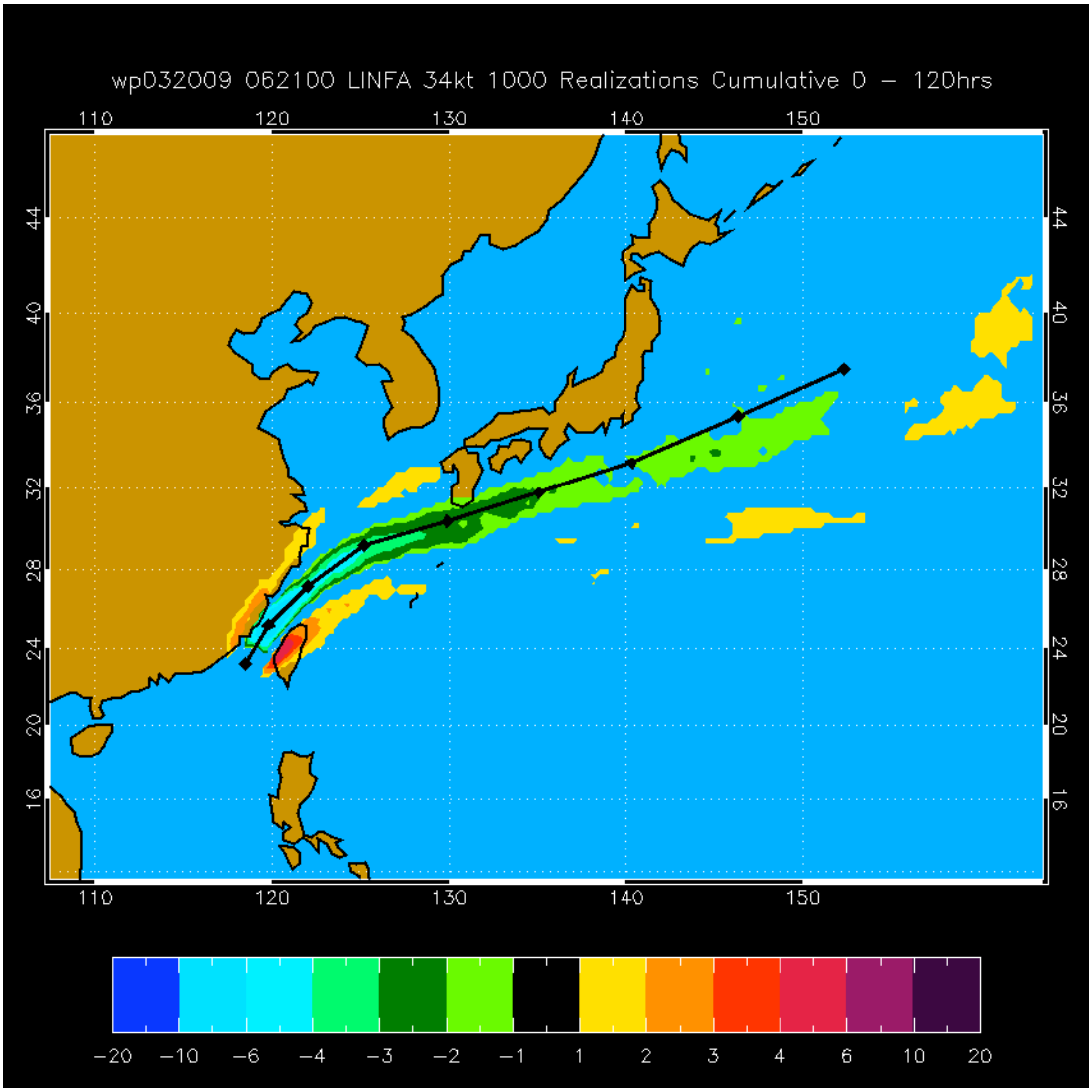

Figure 4. The difference between the GPCE and operational 0-120 h cumulative 34 kt wind probabilities for Typhoon Linfa beginning on 21 June 2009 at 00 UTC. The GPCE values were mostly in the high tercile for this case.

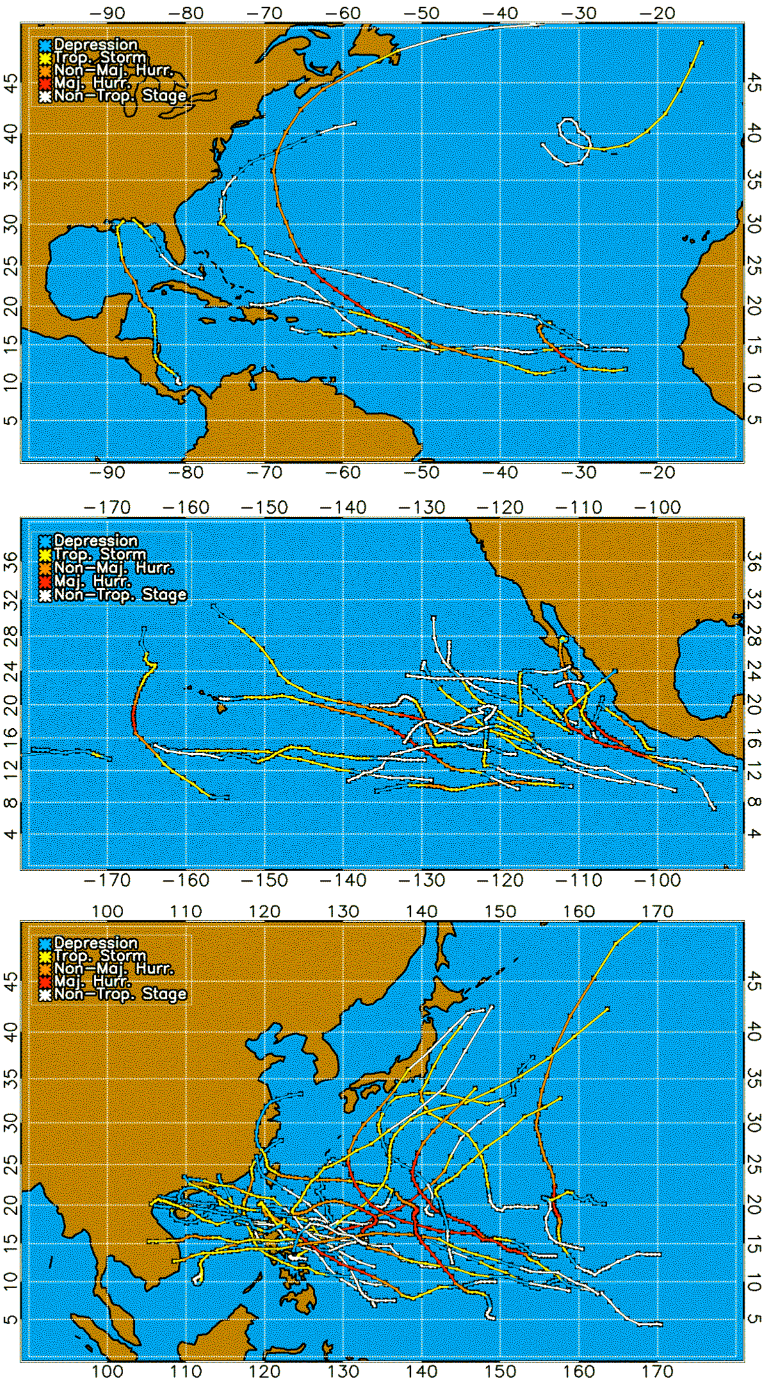

Figure 5. The tracks of all Atlantic (top), eastern/central North Pacific (middle) and western North Pacific (bottom) tropical cyclones from 2009 through Nov. 30 2009.

During the 2009 tests it was discovered that GPCE values are not generated for storms that originate in central North Pacific. This is unfortunate because the central Pacific was more active than normal in 2009 as can be seen in Fig. 5. To overcome this problem, Buck Sampson from NRL has agreed to re-run the GPCE code for the 2009 central Pacific storms (cp012009-cp032009) using the eastern Pacific version. This is the same procedure used for GPCE values for storms that begin in the eastern Pacific and later move west of the 140°W. A separate validation will be performed for all cases in the central Pacific region as soon as the GPCE values for the central Pacific storms and final best tracks are available. The central Pacific cases will also be added to the 2009 web page.

With the exception of the central Pacific cases discussed above, all the 2009 runs are completed, and the verifications can be run shortly after the best track files are ready. These verifications will be provided to the JHT as a supplement early in 2010 so that they can be used in the final decision regarding operational implementation of the GPCE version of the code.

3. Things not completed

All tasks in this project were completed except the verifications from the 2009 runs. These are still awaiting the final best track files. The real time demonstration of the parallel runs was more limited than originally planned, but some were available at the end of the 2009 season. All cases from 2009 in all basins are available on the project web page described above.

4. Things that did not succeed.

Based on the 2008 Atlantic results, the GPCE version of the MC model improved the Brier and Threat Scores. The confirmation of that result for the Atlantic in 2009 and the other basins is awaiting the verification of the 2009 cases as soon as best tracks are ready. Thus, it remains to be seen whether the GPCE version improves the Brier and Threat Scores in all basins.

5. Follow-on work

As described above, the central Pacific cases from 2009 still need to be run once the GPCE values are provided by NRL. After that, the verifications for all basins from 2009 will be run and the results will be sent to the JHT. It should be possible to complete the verifications within a few days after the best tracks are finalized. The JTWC best tracks sometimes take longer to complete. If so, the Atlantic, east and central Pacific verifications will be provided first, with a preliminary west Pacific verification using the JTWC working best track. However, given that the verification relies on the wind radii estimates, which must be consistent with the final intensity estimates, it is highly preferable to use the final best track for the verification. For example, if the final best track intensity were increased from 60 to 65 kt for a given 6 h period, the operational estimates of the 64 kt wind radii would all be zero, even though they should be non-zero. Because the differences between the GPCE and operational models are not that large (typically 10% or less), it is preferable to eliminate unneeded noise from the verification due to inconsistent intensity and radii estimates in the working best track.

The subroutine for the GPCE version of the MC model is similar to the operational version, but a few extra calling arguments are required (the GPCE values). If the GPCE version is implemented in operations, some modifications to the driver programs on the NCEP IBM and the ATCF will be needed. These will be coordinated with NHC and NRL.

As a final remark, the GPCE information was incorporated into the MC model in a very conservative way. The impact of the GPCE input could be amplified by further stratifying the track error distributions (for example, into quartiles or pentads). However, if the distributions are stratified too thinly, the increase in the average NHC track errors with increasing group number might no longer be monotonic, which would introduce noise in the calculation. As time permits, higher stratifications of the error distributions will be tested. The new MC model code was developed to include an option for stratifications into more than three groups. In the longer term, if reliable dynamical model ensemble forecasts become available, the statistically generated track, intensity and structure realizations could be replaced by those from an ensemble system.

References

DeMaria, M., S. Kidder, B. Sampson, J. Knaff, C. Lauer and C. Sisko, 2009b: An improved wind probability program: A Year 2 Joint Hurricane Testbed project update. Presented at the 63rd Interdepartmental Hurricane Conference, March 2-5, 2009, St. Petersburg, FL (available from http://www.ofcm.gov/ihc09/linking_file_ihc09.htm).

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

Appendix

Year-2 Project Timeline and Deliverables

Apr 2008 – Coordinate with TPC for implementation of new wind probability table
May 2008 – Begin monitoring of new wind probability table
Jul 2008 – Prepare final version of MC code for parallel runs during the 2008 season
Aug 2008 – Begin parallel runs during 2008 season and monitor results during the season
Dec 2008 – Perform preliminary verifications of parallel MC runs
Dec. 12, 2008 – Provide mid-year progress report for year 2
Feb 2009 – Perform final verifications of parallel MC runs (depends on timing of final best track)
Mar 2009 – Report results at IHC
Mar 2009 – Make final modifications to MC code based upon 2009 results and assist with final operational implementation if appropriate
July 31, 2009 – Final report due