

**NOAA Joint Hurricane Testbed (JHT)  
Annual Project Progress Report, End of Year 1**

Date: September 30, 2014  
Reporting Period: September 1, 2013 – August 31, 2014  
Project Title: Guidance on Intensity Guidance  
Principal Investigators: David S. Nolan, RSMAS, University of Miami, and  
Andrea Schumacher, CIRA, Colorado State University  
Award Period: September 1, 2013 – August 31, 2015

***1. Long-term Objectives and Specific Plans to Achieve Them:***

This goal of this project is to develop a system for real-time prediction of the expected errors of individual hurricane intensity forecast models and to use this information to improve operational forecasts. In the first year of the project, we have built on the recent results of Bhatia and Nolan (2013) to construct a model that predicts the expected error of each intensity forecast model at each forecast interval based on real-time synoptic and climatological information, such as wind shear, current intensity, and latitude. Error prediction models have been developed for each of the “early” intensity forecast models that are available to forecasters: DSHP, LGEM, GHMI, and HWFI. Our goal by the end of year 1 was to have a prototype of this prediction system running in real-time during the 2014 hurricane season. As noted below, this goal was not quite met by the end of year 1. In year 2, we plan to build a corrected consensus model that will weight each of the four intensity models based on their relative expected errors at each time.

***2. Year 1 Accomplishments:***

*a. Development of model error and predictor databases*

In the first year of this project, we developed a comprehensive database of intensity forecasts, intensity forecast errors, and synoptic and environmental information from the 2007-2013 hurricane seasons. All information such as storm intensity, wind shear, maximum potential intensity, ocean heat content, etc., comes from the SHIPS database (stext files), information that is available in real-time during operational forecasts.

From this database, a large number of candidate predictors of error have been selected. These can be divided into synoptic predictors (which include information about the storm itself, such as its current intensity and location) and “proxy” predictors that are indicative of the stability of the atmospheric flow or the uncertainty of the initial condition. For each forecast, the synoptic predictors are computed at the analysis time (zero hour) and for the average of the forecast period (e.g., the 48 hour average wind shear magnitude during a 48 hour forecast).

Since all the models used are updated almost every year, it should be most effective to use forecasts and errors based on the versions of the model that are used in the present year.

Fortunately, the staff at NCEP/EMC has generously provided us with the results of retrospective forecasts from the GHMI, HWFI, DSHP and LGEM for four years using the 2014 versions of each model. These forecasts and their errors have also been tested as the training data for the multilinear regressions.

*b. Predictor selection, adjustment, and results*

The methodology for the development of the error prediction models is very similar to that used for SHIPS (DeMaria and Kaplan 1994). Multiple linear regression models have been derived using the synoptic and proxy predictors to predict both the absolute error (AE) and the actual error (bias) of DSHP, LGEM, GHMI, and HWFI every 12 hours from 24 to 120 h (the 12 h error forecast will be included in future iterations). While we did not originally intend to predict bias (the positive or negative error value), this has since been included (see below) in response to feedback from NHC staff.

The standard “cross-validation” approach is used, whereby all but one of the years from 2007-2013 are used as the training data, and then the excluded year is used for validation; this is repeated for all years. As in SHIPS, a backward stepping stepwise regression procedure was used to select the predictors. The regression equation starts with all of the predictors and then the least significant predictor is removed. This process is repeated until the weighting coefficients associated with the predictors are all different from 0 at the 95% confidence level. For each model, the same set of predictors is used at all forecast times (this greatly simplifies implementation), but the weighting coefficients can be different for each model. The top 10 predictors for each predictand are listed in Table 1.

AE	BIAS
Intensity deviation from ensemble mean	Intensity deviation from ensemble mean
Intensity forecast standard deviation	Forecast intensity
Average latitude	0 hour intensity
0 hour latitude (2 <sup>nd</sup> order Gaussian fit)	0 hour land distance ( <sup>^2</sup> fit)
0 hour latitude	0 hour RH ( <sup>^3</sup> fit)
Forecast intensity	0 hour MPI
Absolute value of forecasted intensity change	0 hour heat content
Average heat content	Absolute value of forecasted intensity change
Average MPI	Average MPI
RH (Gaussian fit)	Average heat content

Table 1: The top 10 predictors for absolute error and bias.

An important but challenging intermediate step is the transformation of some of the predictor inputs into functions that do not vary linearly. For example, our previous work showed that “medium” levels of humidity are an indicator of higher forecast error, since storms in such environments can either weaken or intensify. Therefore, relative humidity and some other predictors are modified to generate maximum values for an intermediate value and minimum values for their extremes, using either a Gaussian or a polynomial function.

To date, results are moderately favorable, with some cases of fairly high correlations ( $R$  values  $\sim 0.6$  or more). An example of very good correlation of predicted AE versus true AE is shown in Fig. 1. The general trend is for better predictions of forecast errors for the longer intervals (96h, 120h). This may be due to multiple factors, such as the accumulated effect of physical processes over time (e.g., large ocean heat content over several days), or the fact that errors are simply larger over longer forecast periods.

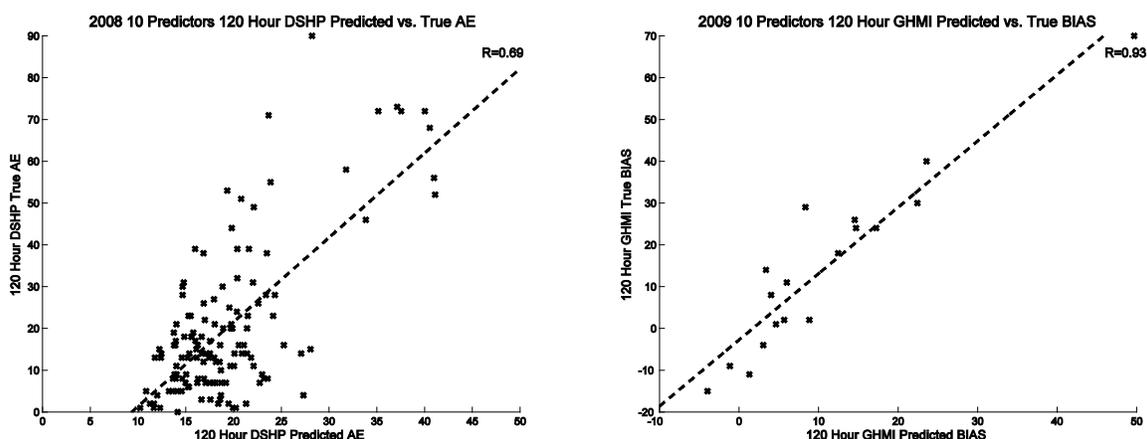


Fig 1. Predicted absolute intensity error (AE) versus true absolute error for 120 hour DSHP forecasts (left) and for bias for 120 hour GHMI forecasts (right). The dashed line indicates the least squares regression line and the  $R$  values are shown in the upper-right of each plot.

An unexpected result is that many of the predictions for bias are more accurate than the predictions of AE. The right panel of Fig. 1 shows such a case, for the 120 hour GHMI forecasts. Similarly, the skill scores (AE of the error prediction compared to the AE of using average error [from climatology] as the error prediction) are generally higher for forecasting model bias rather than for forecast error. Tables showing skill scores by hour for AE and bias are shown below.

### c. Implementation for 2014

One of our goals for the first year was the implementation of a real-time (or quasi-real time) prediction system to begin assessments of the operability of the system. However, this implementation is currently underway and we expect to have the system operating in a few

weeks. Of course, the part of the season that we missed will be evaluated using retrospective data.

**3. Future Year 2 Efforts:**

October 2014: Implementation of a real-time system that predicts absolute error (AE) and error (bias) for each of the 4 intensity forecast models, and retrospective calculations for the part of the hurricane season that has already passed.

November 2014 - December 2014: Assessment of the 2014 hurricane season results and exploration of communication methods to the forecasters (e.g., forecasts of low, medium, or high errors rather than numerical values).

January - March 2015: Development of the weighted consensus model, whereby each intensity forecast model is weighted by the inverse of its expected AE.

April – September 2015: Further refinements of the error forecasts and the weighted ensemble, implementation for the East Pacific, and delivery of the operational system.

Hours	DSHP	LGEM	HWFI	GHMI
24	5%	4%	5%	4%
36	4%	5%	6%	4%
48	7%	5%	7%	7%
60	8%	6%	4%	6%
72	6%	5%	7%	8%
84	7%	6%	8%	8%
96	4%	5%	8%	11%
108	6%	5%	5%	11%
120	6%	5%	7%	15%

Table 2: Skill scores for predictions of AE by forecast hour, using cross-validated data from 2007-2013.

Hour	DSHP	LGEM	HWFI	GHMI
24	9%	11%	11%	14%
36	10%	12%	13%	19%
48	10%	12%	14%	18%
60	11%	12%	17%	17%
72	10%	9%	20%	18%
84	11%	8%	22%	16%
96	13%	8%	24%	15%
108	14%	10%	28%	15%
120	17%	14%	32%	21%

Table 3: Skill scores for predictions of bias by forecast hour, using cross-validated data from 2007-2013.

#### **4. References**

Bhatia, K. T., and D. S. Nolan, 2013: Relating the Skill of Tropical Cyclone Intensity Forecasts to the Synoptic Environment. *Wea. Forecasting*, **28**, 961–980.

Bhatia, K. T., and D. S. Nolan, 2014: Prediction of tropical cyclone intensity forecast error. 31<sup>st</sup> Conference on Hurricanes and Tropical Meteorology, American Meteorological Society, San Diego, California. Available for download from:  
<https://ams.confex.com/ams/31Hurr/webprogram/Paper244417.html>

DeMaria, M., and J. Kaplan, 1994: A statistical hurricane intensity prediction scheme (SHIPS) for the Atlantic basin. *Wea. Forecasting*, **9**, 209-220.