Improving the Validation and Prediction of Tropical Cyclone Rainfall

Joint Hurricane Testbed
Midterm progress report February 2004

Principal Investigators

Timothy Marchok	NOAA / GFDL	(timothy.marchok@noaa.gov)
Robert Rogers	NOAA / AOML / HRD	(robert.rogers@noaa.gov)
Robert Tuleya	SAIC at NCEP / EMC	(robert.tuleya@noaa.gov)

Project summary and timeline

Through funding from the Joint Hurricane Testbed (JHT), this project is working to improve the validation and prediction of tropical cyclone rainfall. This will be achieved by accomplishing several goals: 1) Developing new rainfall validation schemes that provide a baseline of comparison for different forecast systems; 2) Producing rainfall forecast error statistics for historic United States landfalling storms using traditional and new validation techniques for the operational GFDL, Eta and GFS models, and the benchmark Rainfall CLIPER (R-CLIPER) model; and 3) Designing a new forecasting tool based on the R-CLIPER model that incorporates information related to vertical shear and storm track

With these goals in mind, the following tasks were proposed to be completed by the end of January 2004:

- Acquire National Precipitation Validation Unit (NPVU) and other historical rain datasets
- Assess availability of data for all historic landfalling cases back to 1995 for the GFDL, Eta and GFS models.
- Validate current & historic cases from the operational GFDL, Eta, GFS and R-CLIPER models

This document will report on the progress reached up to this point.

Summary of Progress

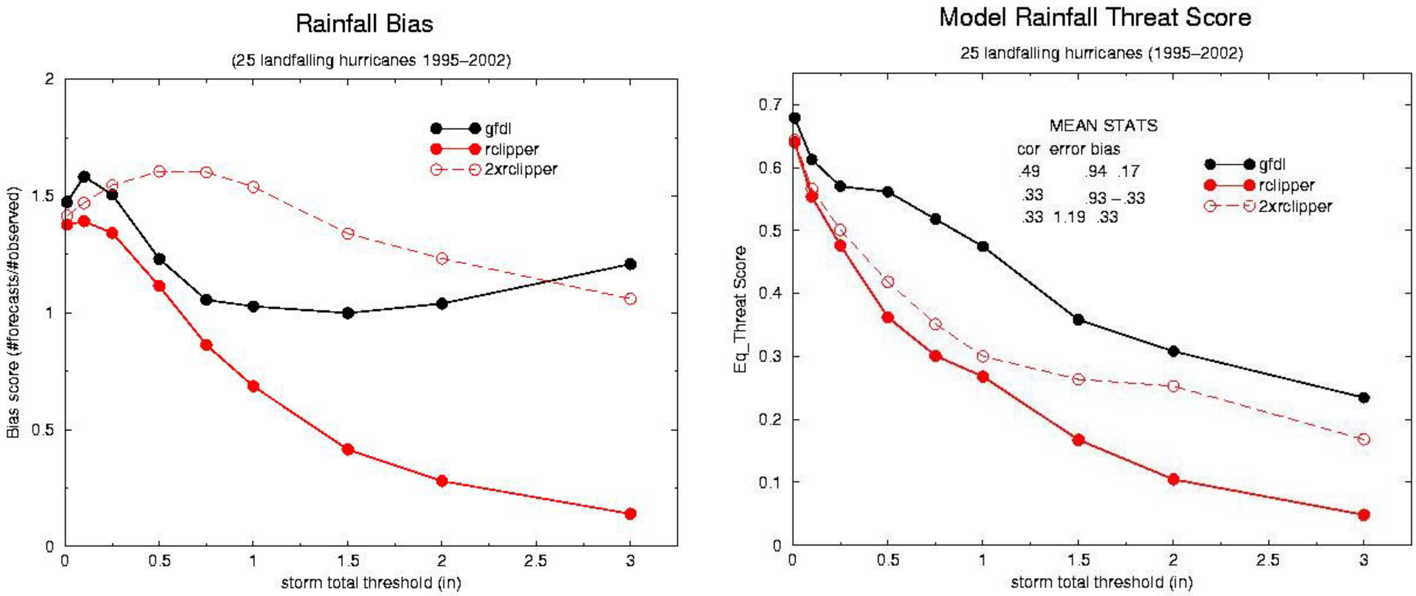
The first task involved the acquisition of the NPVU dataset and an assessment of the availability of model output back to 1995. Table 1 shows an inventory of datasets available for rainfall validation. As can be seen from the table, we have 4-km gridded NPVU data back to 1996 for all landfalling storms. We also have model output back to 1995 for the GFDL model (1/6 degree highest resolution), 1997 for the GFS model (½ degree highest resolution), and 1998 for the Eta model (12 km highest resolution). In addition, we can run two versions of the R-CLIPER, one developed from TRMM climatologies and one developed from raingage climatologies, for all of the cases. With these datasets several years worth of observations and forecasts can be compared to validate the models' rainfall predictions.

Inventory of datasets

Type of data	**Title of data**	**Time period**	**Dataset specifics**
• observation	• NCEP Stage IV precip data	• 1996-present	• hourly rainfall from gage-corrected radar rainfall • GRIB format • 4 km resolution, polar stereographic projection
• model	• GFDL model	• 1995-present	• operational model for each year • 1 degree entire domain for each year • 1/3 degree entire domain starting in 2000 • 1/6 degree inner mesh (11 x 11 degrees) starting 2002, pre-2002 1/6 degree inner mesh (5 x 5 degrees) • 6-hour output time resolution (some cases re-run are 1-h or less) • cylindrical equidistant map projection (lat-lon grid)
• model	• ETA model	• 1998-present	• operational model for each year • 12-km horizontal grid length • 6-hour output time resolution, but 3-h resolution starting a few years ago
• model	• GFS model	• 1997-present	• operational model for each year • spectral model; equivalent resolution on Gaussian grid of ½ x ½ degree resolution 2000-present; pre-2000 0.7 x 0.7 degree • 6-hour output time resolution • spectral model, Gaussian grid conversion for calculations
• model	• R-CLIPER1	• all times	• Marks-DeMaria version of R-CLIPER • based on climatology of rain rates generated from TRMM satellite using 10-km range rings • 1/6 deg resolution grid produced • can be run on any set of positions and intensities
• model	• R-CLIPER2	• all times	• DeMaria-Tuleya version of R-CLIPER • based on climatology of rain gage measurements • can be run on any set of positions and intensities

Table 1. Inventory of observational and modeling datasets to be used in rainfall validations.

In addition to organizing the database shown above, we have made progress in performing evaluations using standard evaluation criteria commonly used at EMC such as the bias score, which measures, for various rainfall thresholds, the ratio of forecast occurrences to observed occurrences, and the equitable threat score, which evaluates the ratio of correct forecasts to the total number of forecasts and observations for given thresholds. These evaluations have been performed on a variety of 25 landfalling cases from the past several years. An example of these evaluations is shown in Figure 1. From Figure 1a, it can be seen that both the R-CLIPER and GFDL models exhibit an over-forecast bias for small rainfall amounts (<= 0.5 inch). While the GFDL model has a weak bias for larger rainfall amounts, the R-CLIPER exhibits a strong under-forecast bias for amounts >= 1.5 inches. Multiplying the R-CLIPER values by 2 lessens this under-forecast bias. The threat scores shown in Figure 1b indicate that,



(a) (b)

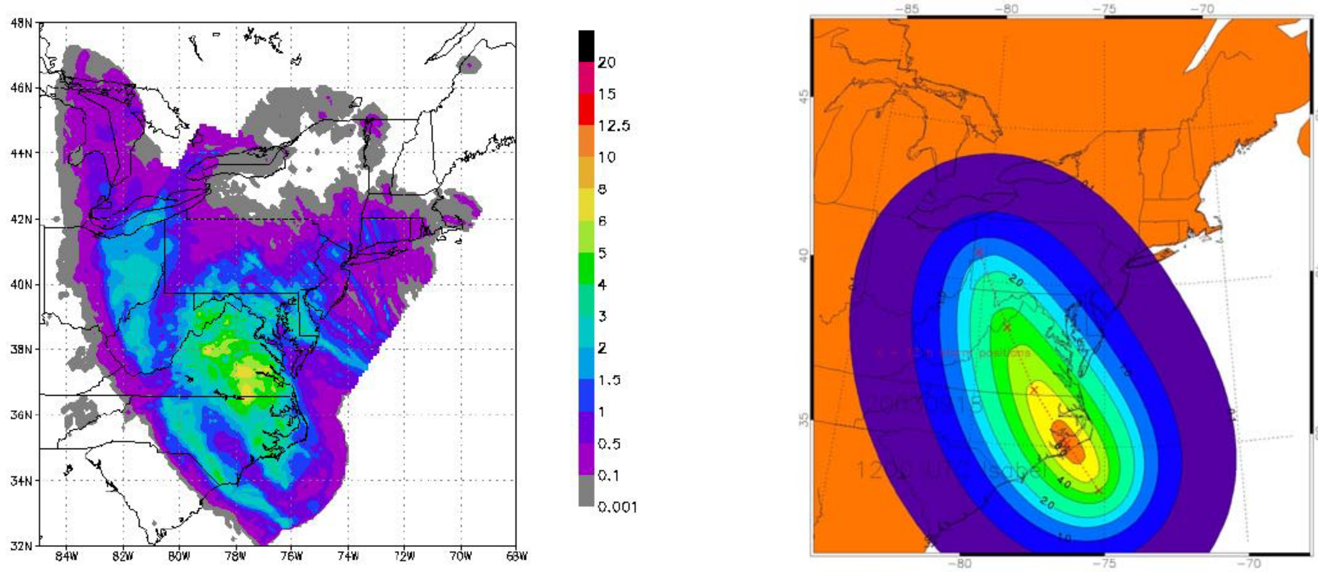
Figure 1. Plot of (a) Bias score and (b) Equitable Threat score for comparisons of GFDL and R-CLIPER for 25 landfalling cases from 1995 to 2002. Also shown is a comparison for a version of R-CLIPER2 where the values produced from the forecast are multiplied by 2.

for all thresholds, the GFDL model has produced a greater ratio of forecast hits than has the R-CLIPER model. Further evaluations using the other dynamical forecast models are ongoing, and they will be reported at the upcoming Interdepartmental Hurricane Conference in March 2004.

While the validation techniques shown above yield valuable information regarding the biases of the models, there are limitations in what these techniques can reveal. For example, some standard validation techniques do not account for the significant error that can arise simply from having an incorrect forecasted storm track. Furthermore, a great deal of useful information can be obtained from considering the performance of the forecasts for the entire distribution of rainfall, not just peak rainfall amounts or point comparisons with specific rain gauges. As a result of these limitations, work has begun in developing new validation techniques that better account for such factors as track error, sampling size discrepancies, and comparing the entire distribution of rainfall rather than peak rainfall amounts and point comparisons.

The development of these techniques has begun on a single storm, Hurricane Isabel of 2003. Once the techniques are developed and refined, they can be applied to all of the landfalling cases in our database. Figure 2 shows a plot of 24-hour accumulated rainfall for Hurricane Isabel from 12 UTC 18 to 12 UTC 19 September for the NPVU dataset and a run of the R-CLIPER model for the same time period. The R-CLIPER reproduces the distribution of rainfall fairly well compared with the NPVU data, with a broad swath of rain greater than 2 inches covering roughly equivalent area and maximum rainfall amounts of 5-6 inches in a 24-hr time period.

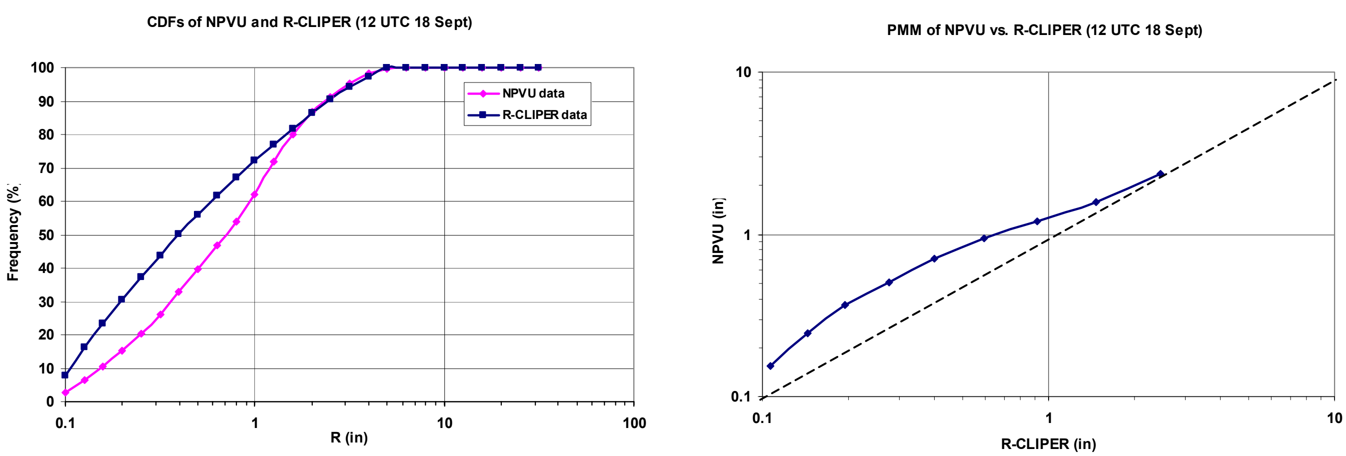
From these datasets characteristics of the entire rainfall distributions can be compared. Figure 3a shows a plot of the cumulative distribution function (CDF) for the NPVU and R-CLIPER data from the time period shown in Fig. 2. The CDFs show that the R-CLIPER is



(a) (b)

Figure 2. Plots of 24-h accumulated rainfall (in) from 12 UTC 18 to 12 UTC 19 September for Hurricane Isabel from (a) NPVU observations; and (b) R-CLIPER forecast

producing more of the lighter rainfall amounts than the NPVU data is showing. For example, the R-CLIPER has more points producing rain less than 1 inch than the NPVU data shows. About 40% of the raining area in the NPVU data is receiving rain amounts up to 0.5 inches, while about 60% is receiving that much rain in the R-CLIPER. A comparison of the CDFs is also shown using the probability matching method (PMM; Fig. 3b). The PMM finds the set of pairs of NPVU and R-CLIPER CDFs at which the cumulative probabilities of the two are equal, assuming that the area covered by that cumulative probability rain amount is equivalent for both



(a) (b)

Figure 3. (a) Cumulative probability distributions of NPVU and R-CLIPER 24-h rainfall (inches) for Hurricane Isabel (2003). (b) Probability-matched 24-h rain estimates from NPVU data and R-CLIPER for Hurricane Isabel. Each point represents the probability-matched value at 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, and 90% from left to right, respectively. The dashed line denotes the point at which the NPVU and R-CLIPER CDFs are equal.

the NPVU data and the R-CLIPER. This comparison shows the same bias toward smaller rain amounts in the R-CLIPER at all points in the rainfall distribution less than 1 inch seen in Fig. 3a. The R-CLIPER bias is nearly half of the rain in the NPVU data for those rainfall amounts, indicating the need to double the predicted amounts of rainfall from the R-CLIPER for the lighter rain amounts. This is consistent with the bias seen in the previous JHT work of Marks and DeMaria.

Future work

Future work will involve expanding the efforts shown above. In the short term (i.e., within the next 1-3 months), we will continue the rainfall validation using traditional methods of validation for the landfalling cases between 1995 and the present. We will evaluate the performance of the new 2003 version of the GFDL model, the Eta model, the GFS model, and the GFDL model with a different land-surface parameterization scheme. We will also continue to work toward developing new validation techniques such as that shown in Fig. 3. As a part of this development, we will perform the same validations on Isabel using the GFDL, Eta, GFS, and R-CLIPER forecasts of the storm. Other validation techniques will also be performed. For example, we will evaluate the areally-averaged rainfall for Isabel. Such averages can cover the entire storm or target subregions of the storm to identify differences in the development of asymmetries in the rainfall fields from the different models.

On a longer time scale, we plan to apply these validation techniques to all of the landfalling cases in our database. Such comparisons will yield a comprehensive evaluation of the forecasts covering the entire distribution of rainfall, rather than just focusing on peak rain amounts, and will allow for a more reliable assessment of the models that may identify biases that can be incorporated into each of the models (such as the doubling of light rain amounts in the R-CLIPER). Finally, we plan to incorporate information regarding vertical shear and storm heading into the R-CLIPER. The R-CLIPER assumes a circularly symmetric rainfall distribution, so details related to variations in the storm-relative distribution of rainfall and its possible impact on the accumulated rainfall swaths are not included. Recent research has shown a link between the storm heading and vertical shear encountered by the storm. The inclusion of these parameters in the R-CLIPER will add more structure to the rainfall fields produced by the model.