6A.4 EVALUATION OF RAINFALL FORECASTS FROM THE OPERATIONAL GFDL HURRICANE MODEL

R. Tuleya, SAIC/NOAA/NCEP/EMC
M. DeMaria, NOAA/NESDIS/ORA
R. Kuligowski, NOAA/NESDIS/ORA

1. INTRODUCTION

The devastating rains and resulting floods of landfalling tropical systems have been quite evident from the recent cases of Floyd (1999) and Allison (2001). Until recently, the primary rainfall guidance products of the GFDL model were graphical images that did not get much circulation beyond the NHC. To date, little objective verification has been made for hurricane specific landfalling cases from any numerical guidance. A preliminary evaluation of low resolution 1^o gridded output for US landfalling cases of 1995-1999 indicated that the GFDL model exhibited some degree of skill in forecasting storm total precipitation, and area-averaged rainfall. The GFDL model has been re-evaluated to include higher resolution output for these 16 cases and the verifications have now been expanded to include nine additional cases spanning from 1995-2002.

2. EXPERIMENTAL DESIGN

One drawback to evaluation of rainfall skill for hurricanes at landfall is the problematic quantitative validation of results. Traditional methods may be ineffective and incomplete due to the copious rainfall over relatively small areas controlled to a large extent by the hurricane track. Therefore a straight forward method used so far has been the comparison of model data directly with gauge data at the observation location. All cases were initialized at the same 12UTC times as the high density daily River Forecast Center data set.

The GFDL model was evaluated at 12UTC utilizing the last model forecast initialized within ~ one day of landfall. The gauge data within 800km of the storm track was summed over time to give a storm total. The storm total period was generally a 3 day total, but was less if the system dissipated and/or became extra-tropical according to NHC best track data. The cases evaluated are shown in Table 1. Recently a baseline definition of skill has been developed, a rainfall "Cliper" (i.e. RCliper) which uses climatological rainfall rates moved along a forecasted or observed track. The results of this baseline rainfall model will be compared with the operational version of GFDL model used at that time in order to evaluate it's forecast usefulness.

The statistics used to evaluate skill were the mean error and mean bias averaged over all the gauge sites. A estimate of quality of the forecasted pattern distribution was obtained through the correlation coefficient obtained by comparing the observed and predicted rainfall over all gauge sites. More traditional precipitation verification scores were also calculated including equitable threat and bias scores.

Storm/Year	Landfall Time	Model start Time
Allison/1995	06051400	06041200
Erin/1995	08020615	08011200
Opal/1995	10042200	10041200
Bertha/1996	07122000	07121200
Fran/1996	09060030	09051200
Joseiphine/1996	10080330	10071200
Danny/1997	07191800	07191200
Bonnie/1998	08270400	08261200
Charley/1998	08221000	08211200
Earl/1998	09030600	09021200
Frances/1998	09110600	09101200
Georges/1998	09281130	09271200
Hermine/1998	09200500	09191200
Bret/1999	08230000	08221200
Dennis/1999	09042100	09041200
Floyd/1999	09160630	09151200
Harvey/1999	09211700	09211200
Irene/1999	10152000	10151200
Gordon/2000	09180300	09171200
Barry/2001	08060500	08051200
Gabrielle/2001	09141200	09131200
Fay/2002	09070830	09061200
Hanna/2002	09141430	09141200
Isidore/2002	09260630	09251200
Lili/2002	10031400	10021200

Table 1. Model Cases Evaluated

3. RESULTS

Results indicate an overall correlation coefficient of ~0.5 between RFC gauges and GFDL model amounts for the storm total rainfall (primarily 72h) for these 25 cases. A version of rainfall "Cliper" run up the GFDL forecast tracks yields a correlation

*Corresponding author: Robert Tuleya, SAIC/EMC/NCEP/NOAA CCPO/ODU 768 W. 52nd St. Norfolk, VA 23529. Email: robert.tuleya@noaa.gov

coefficient of ~0.35, thus indicating some degree of relative skill for the GFDL model in predicting precipitation distribution. Both the GFDL and rainfall "Cliper" exhibited a rather large mean error of ~0.9in for the 32430 gauge observations. To assess the impact of track error, RCliper was evaluated by computing the rainfall along the best track rather than the GFDL forecast track. This reduced the mean error to ~.8in for the data set. The correlation coefficient increased from .35 in RCliper to .48 in the best-track version of RCliper. This indicates that the track error for these cases lead to a reduction of ~ 0.1in in mean error and and explains ~11% more variance. Note that the GFDL model correlation is roughly equal to that of the best track RCliper. The mean bias was also computed for these 25 cases. The GFDL mean bias was .16in and can be compared to the RCliper and best-track RCliper values of -.39in and -.47in. The under- prediction of rainfall amounts greater than ~.5in contributes to this general negative bias in rainfall amounts.

Figure 1 illustrates the equitable threat score, a measure of the relative skill for a distribution of rainfall amounts, including greater than 9in to evaluate possible copious storm rainfall at hurricane landfall. For these cases the threat score peaks at ~.3 at values of ~one inch. Both RClipers also have the same tendency but have less skill. The Rclipers are handicapped by a rather severe low bias for amounts greater than one inch. This has been shown in previous results from RCliper. On the other hand, the GFDL scores at high amounts are effected by a relatively high bias; at low amounts GFDL predicts rain in almost all areas leading to low scores below ~0.5in. Interestingly, the Rclipers appear to have a similar tendency. It has been anticipated that track error may have detrimental impact of rainfall errors. This can be seen in the improvement of RCliper evaluated along the observed best track instead along the GFDL model track. For these cases it appears that the detrimental track effect has a major impact only for values less than ~1.5 inches. This indicates that for relatively good track forecasts made near landfall other factors such as topographical and synoptic forcing may play an important role for large rainfall amounts. The GFDL model has these effects included whereas the versions of RCliper do not at present.

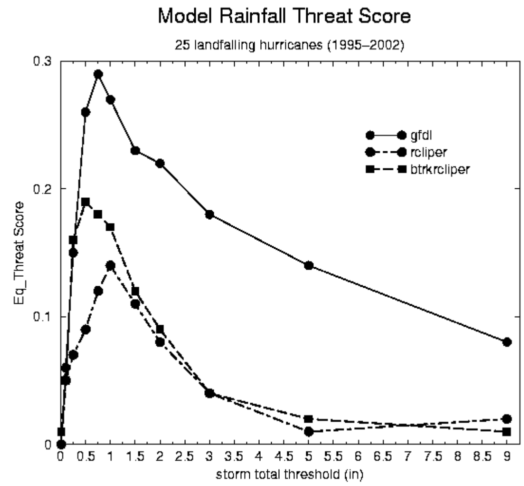

Figure 1. GFDL threat scores (solid), and Rainfall Cliper scores evaluated along the GFDL track(dash-dotted) and the NHC best track (long dashed) for all 25 landfalling cases listed in Table 1.

4. SUMMARY AND DISCUSSION

A thorough quantitative evaluation of landfallling hurricane rainfall has been performed for the 1995-2002 seasons for the operational GFDL hurricane model. The analysis utilized high resolution RFC daily rain gauges and $1/3^{O}$ resolution model output and emphasized storm total rainfall near the storm track. The GFDL model was compared to a baseline rainfall Cliper model to asses relative skill. Both RCliper and GFDL had comparable mean errors of ~0.9in for the 25 cases. The GFDL model exhibited a relatively higher pattern correlation than Rcliper, but still only explained ~25% of the spatial variance. The GFDL model also had higher equitable threat scores than RCliper, partially because of the known large low bias of RCliper for amounts larger than 0.5 in. A large case-to-case variability was found which was dependent on both synoptic conditions and track error. This will be discussed as well as the impact of the GFDL model upgrades made in 2003.