

## Evaluation of GFDL and Simple Statistical Model Rainfall Forecasts for U.S. Landfalling Tropical Storms

ROBERT E. TULEYA

*SAIC and NOAA/NCEP/EMC, Norfolk, Virginia*

MARK DEMARIA

*NOAA/NESDIS/ORA, Fort Collins, Colorado*

ROBERT J. KULIGOWSKI

*NOAA/NESDIS/ORA, Camp Springs, Maryland*

(Manuscript received 21 April 2005, in final form 13 April 2006)

### ABSTRACT

To date, little objective verification has been performed for rainfall predictions from numerical forecasts of landfalling tropical cyclones. Until 2001, digital output from the operational version of the Geophysical Fluid Dynamics Laboratory (GFDL) hurricane forecast model was available only on a  $1^\circ$  grid. The GFDL model was rerun or reanalyzed for 25 U.S. landfalling tropical cyclones from 1995 to 2002 to obtain higher resolution ( $1/3^\circ$ ) output. Several measures of forecast quality were used to evaluate the predicted rainfall from these runs, using daily rain gauge data as ground truth. The overall quality was measured by the mean error and bias averaged over all the gauge sites. An estimate of the quality of the forecasted pattern was obtained through the correlation coefficient of the model versus gauge values. In addition, more traditional precipitation verification scores were calculated including equitable threat and bias scores. To evaluate the skill of the rainfall forecasts, a simple rainfall climatology and persistence (R-CLIPER) model was developed, where a climatological rainfall rate is accumulated along either the forecasted or observed storm track. Results show that the R-CLIPER and GFDL forecasts had comparable mean absolute errors of  $\sim 0.9$  in. (23 mm) for the 25 cases. The GFDL model exhibited a higher pattern correlation with observations than R-CLIPER, but still only explained  $\sim 30\%$  of the spatial variance. The GFDL model also had higher equitable threat scores than R-CLIPER, partially because of a low bias of R-CLIPER for rainfall amounts larger than 0.5 in. (13 mm). A large case-to-case variability was found that was dependent on both synoptic conditions and track error.

### 1. Introduction

The devastating rains and resulting floods from landfalling tropical systems have been quite evident from the recent cases of Mitch (1998), Floyd (1999), Allison (2001), and Gaston (2004). Rainfall from tropical systems is characterized by copious amounts of precipitation, but is strongly controlled by the mesoscale dynamic forcing of the swirling wind field (Riehl 1954). Until recently, the primary rainfall guidance products

of the operational version of the Geophysical Fluid Dynamics Laboratory (GFDL) hurricane model were graphical images that were provided only to the National Hurricane Center (NHC). To date, little objective verification has been performed for tropical storm specific landfalling cases from any numerical guidance. The GFDL model rainfall forecasts have not been utilized much by National Weather Service (NWS) forecast centers mainly due to the lack of verification statistics and because of the use of other operational NWS rainfall guidance on a daily basis.

A preliminary evaluation of low-resolution ( $1^\circ$ ) output for 16 U.S. landfalling cases from 1995 to 1999 indicated that the GFDL model exhibited some degree of skill in forecasting storm total precipitation and area-

---

*Corresponding author address:* Robert Tuleya, SAIC and NOAA/NCEP/EMC, CCPO/ODU, 768 W. 52d St., Norfolk, VA 23529.

E-mail: robert.tuleya@noaa.gov

averaged rainfall (DeMaria and Tuleya 2001, hereafter DT). All hourly rain gauges within 800 km of the storm track were included in the verification. On average, there were 211 stations for each storm and the GFDL model storm total rainfall amounts were interpolated to the rain gauge locations. Three measures were used to compare the model and gauge rainfall totals as follows: direct comparison at each gauge location, the average rainfall over all the gauges for each storm, and the maximum storm total rainfall for each storm. Results from this study indicated that the GFDL model can forecast the maximum storm total rainfall to within about 35%, and the average rainfall forecasts have only a small high bias. The rainfall forecast at individual gauge locations was accurate to within a factor of 2.

In this study the results of DT will be generalized by adding nine new cases from 2000 to 2002 (for a total of 25), and resampling the GFDL model output on a higher-resolution grid ( $1/3^\circ$ ). For some of the older cases from 1995 to 1999, the resampling required re-running the model since the high-resolution output files were not saved during the operational runs. In addition, the hourly gauge data used by DT for the ground truth are replaced by 24-h rain gauge totals from a combination of daily rain gauge amounts from the daily cooperative observer network via the River Forecast Centers (RFCs), the daily automated gauges from the National Climate Data Center (NCDC) climate network, and 24-h accumulations of hourly rain gauge data from the Hourly Precipitation Dataset (HPD) that are combined and quality checked at the Climate Prediction Center. The density of the daily data ( $\sim 1300$  per storm) is about five times that of the hourly data ( $\sim 200$  per storm; see Fig. 7 for an example of each, and Fig. 1 for an objective analysis of the rain gauge data). The use of gauge data for the verification has some disadvantages because of the nonuniform nature of the rain gauge network and the effects of the different temporal and spatial scales of gauges and models (e.g., Scofield and Kuligowski 2003). However, in this initial study we wished to avoid the additional sources of uncertainty that would result if rainfall estimates from radar and satellite were used for ground truth. In addition to the evaluation measures used in DT, additional verification statistics, including the equitable threat and bias scores, are used to evaluate the GFDL rainfall forecasts.

The NHC performs a comprehensive annual evaluation of all of their guidance models for tropical cyclone track and intensity forecasting. To evaluate the skill of these models, the forecasts are compared with those from simple models based upon climatology and persistence. The Climatology and Persistence (CLIPER) model (Aberson 1998) is the baseline for track skill and

the Statistical Hurricane Intensity Forecast (SHIFOR) model (Knaff et al. 2003) is the baseline for intensity skill. To provide a similar benchmark for the evaluation of the skill of the GFDL model rainfall forecasts, a rainfall CLIPER (R-CLIPER) is developed here. For this model the climatological rainfall rate along the storm track was determined from hourly rain gauge data for 120 U.S. landfalling hurricanes and tropical storms from 1948 to 2000. Analysis of these data revealed that the primary reason for the decrease in average rainfall rate as a storm moved inland was the decay of the storm intensity. Based upon this observation, the rainfall rates as a function of storm intensity and radius from the Tropical Rainfall Measuring Mission (TRMM) determined by Lonfat et al. (2004) were used to refine the R-CLIPER model (Marks et al. 2002). The accumulated rainfall can then be calculated by integrating the rainfall rate along the storm track, given the intensity.

The GFDL and R-CLIPER models are described in section 2. In section 3, the statistical evaluation procedures are presented, followed by the verification results in section 4. A summary and discussion appear in section 5.

## 2. Deterministic models of tropical storm rainfall

### a. GFDL model

The GFDL model was introduced operationally by the National Oceanic and Atmospheric Administration (NOAA) in 1995 and has been the primary mesoscale numerical guidance for Atlantic and east Pacific tropical cyclones to date (Kurihara et al. 1998). The main utility to date of the GFDL model has been for track forecasts, and has contributed (along with global models and other factors) to an increase in skill of NHC official track forecasts over the past decade (DeMaria and Gross 2003). Although capable of resolving hurricane structure to some extent, the GFDL model has proven less useful in intensity forecasting. Furthermore, rainfall forecasts from the model have not been routinely verified by the NHC.

The GFDL model is a multiple nested system that has had rather minor modifications in its dynamics and physics during the period covered by this study. For this time period (1995–2002), the model had a convective adjustment parameterization developed by Kurihara (1973) and a boundary layer parameterization following that of Mellor and Yamada (1974). Clouds were simply diagnosed and interacted with the radiation scheme, and the subsurface and surface layers described by Tuleya (1994). A diurnal cycle was included. The outer domain has spanned a tropical belt of  $75^\circ$

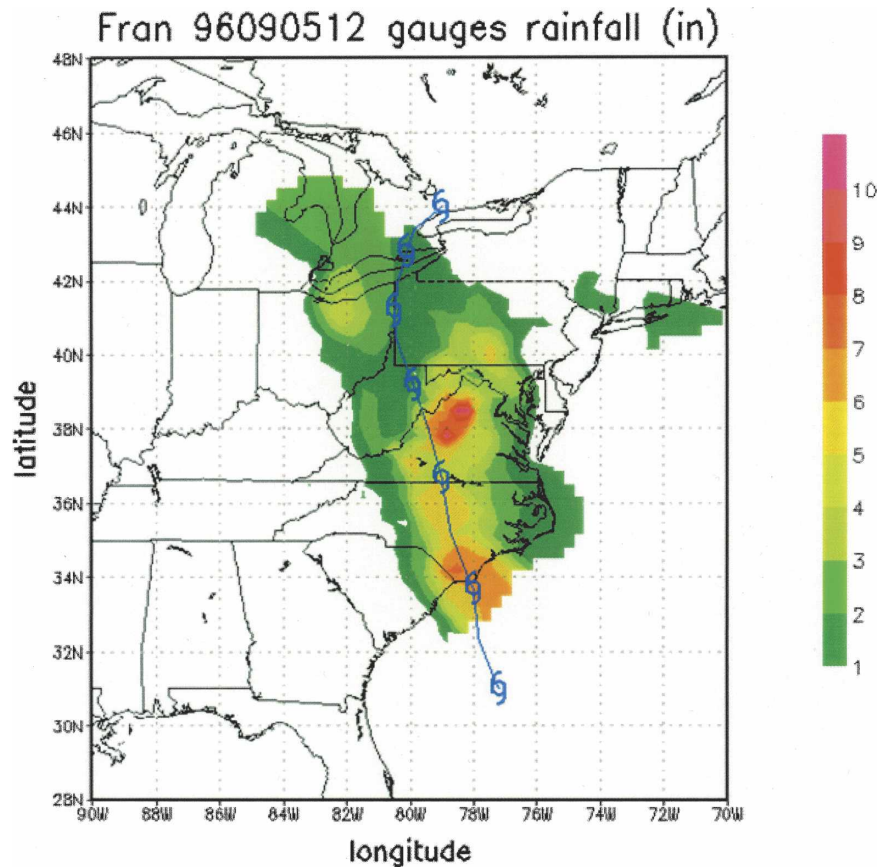


FIG. 1. Storm total rainfall for Hurricane Fran (1996) observed from rain gauges for the period from 1200 UTC 9 Sep to 1200 UTC 12 Sep. Observed track is shown every 12 h.

latitude by  $75^\circ$  longitude. The innermost nest of  $1/6^\circ$  grid spacing moved with the forecast position of the storm. The grid spacing on the outer domain was  $1^\circ$  latitude–longitude until 2002 when it was reduced to  $1/2^\circ$ . The number of nests was reduced from three to two for the 2002–04 seasons. In the operational version of the GFDL model used through 2002, rainfall was estimated as the sum of the convectively and large-scale adjusted mixing ratio with no evaporation and instantaneous rainout. This instantaneous rainfall was accumulated each time step and used for validation. Figures 1 and 2b show the storm total rainfall from the rain gauges and the GFDL model forecast for a case from Hurricane Fran (1996).

#### b. R-CLIPER model

As described in the introduction, predictions based upon climatology and persistence are often used to evaluate the skill of more general forecast models. For tropical cyclone rainfall, an R-CLIPER model was developed for this purpose. In the R-CLIPER model, a

climatological rainfall rate is determined and then integrated along the storm track. Because the primary interest in tropical cyclone rainfall is over land, the variation in rainfall rate after landfall needs to be taken into account. In addition, a number of studies have shown that the rainfall rate is a function of storm intensity, with a tendency for higher rain rates for stronger storms (e.g., Lonfat et al. 2004). This effect will also be taken into account in the R-CLIPER model. Other studies such as those by Corbosiero and Molinari (2003) and Rogers et al. (2003) have shown that there are significant azimuthal asymmetries in tropical cyclone rain rates. These asymmetries are caused by a number of factors, but the most significant factor appears to be the storm response to the environmental vertical wind shear. Because this effect depends on the particular synoptic environment, asymmetries will not be taken into account in the R-CLIPER model. A more general statistical tropical storm rainfall model could be developed to include this effect.

The starting point for the R-CLIPER development is the hourly rain gauge data from the primary and sec-

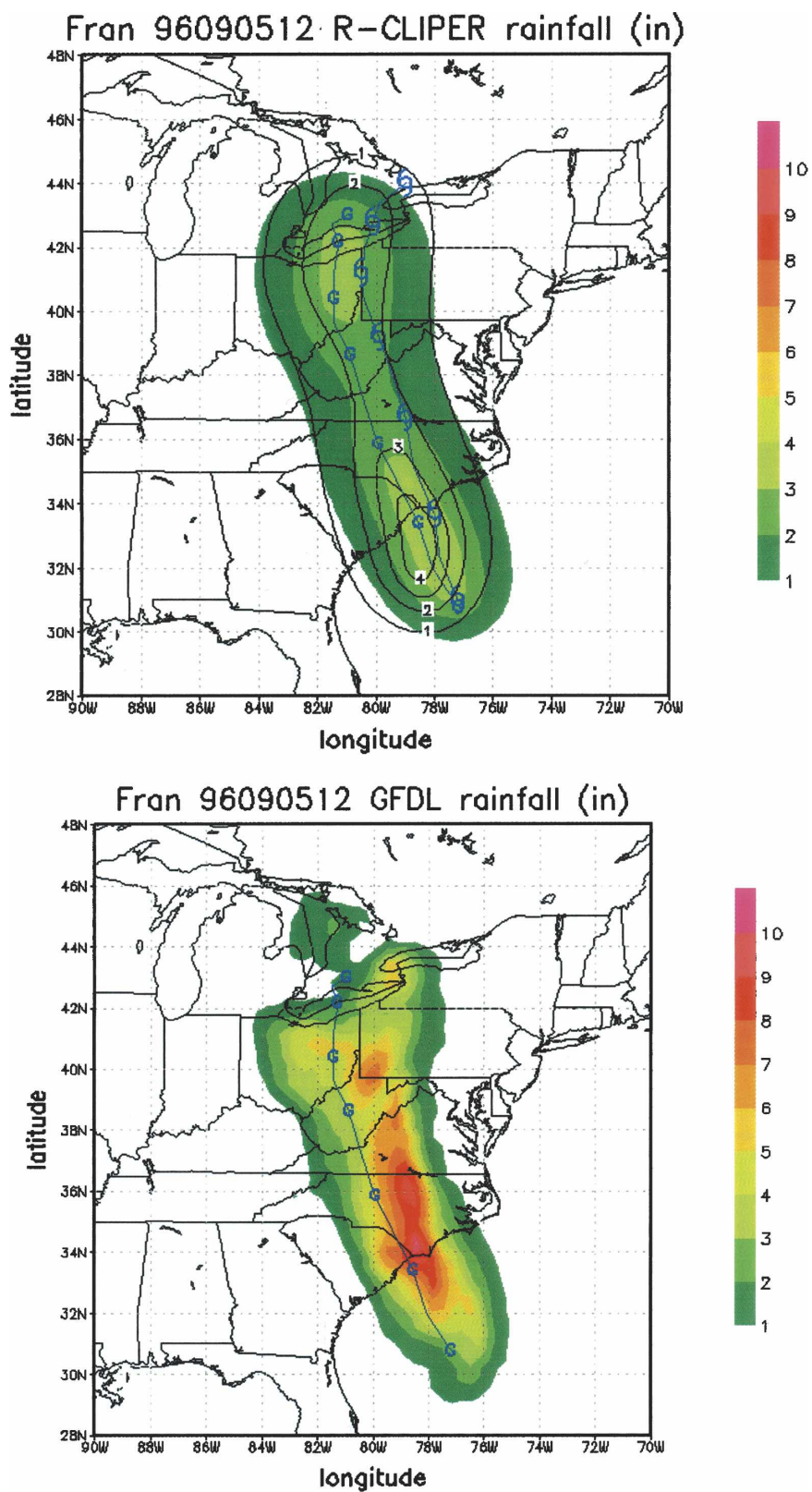


FIG. 2. (a) Storm total prediction using R-CLIPER model along the GFDL forecast track for Fran (1996) for the period from 1200 UTC 9 Sep to 1200 UTC 12 Sep. Contour lines denote rainfall obtained along the best track, i.e., best-track R-CLIPER. (b) Storm total rainfall prediction using operational GFDL model. The "G" and hurricane symbols denote GFDL-predicted and observed tracks, respectively.



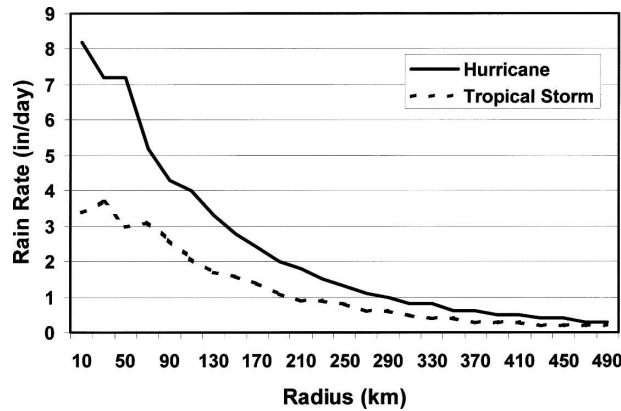


FIG. 3. The average rain rate from the gauge data as a function of radius for all gauge observations when the storm center was over water or had been over land for less than 3 h. Profiles are for storms that were of tropical storm and hurricane intensity at landfall.

ondary stations in the United States available from the National Climatic Data Center (NCDC) archive. Data from the cooperative observer network were also available in this archive, for a total of about 2500 sites within the United States and its territories. These data were obtained for nearly all U.S. landfalling tropical storms and hurricanes from 1948 to 2000. This sample includes 120 storms, of which 63 were hurricanes and 57 were tropical storms just prior to landfall. The sample was restricted to storms that affected the contiguous United States. An example of the coverage of the hourly gauges is shown later (see Fig. 7).

The hourly gauge data were collected for all points within 1000 km from the center starting from the time when the storm was at least 500 km offshore or the first point in the NHC best track, and ending at the time of the last point in the NHC best track for which the storm was still classified as tropical. The hourly data for the 120 storms for the years 1968–2000 were stratified as a function of radius from the storm center and of the time from when the storm center moved inland. The radial interval for the stratification was 20 km and the temporal interval was 6 h (i.e., 3–9 h, 9–15 h, etc.). All gauge values for the time when the storm center was still offshore were assigned to the time interval of 0–3 h.

Figure 3 shows the rain rate as a function of radius for the first time interval (from when the storm center was offshore and up to 3 h inland) for the tropical storm and hurricane cases. This figure shows that the rain rates near the storm center for the hurricane cases are almost three times as large as they are for the tropical storm cases. The rain rates are nearly constant with radius for the first few tens of kilometers and then appear to decay exponentially with radii after that. By

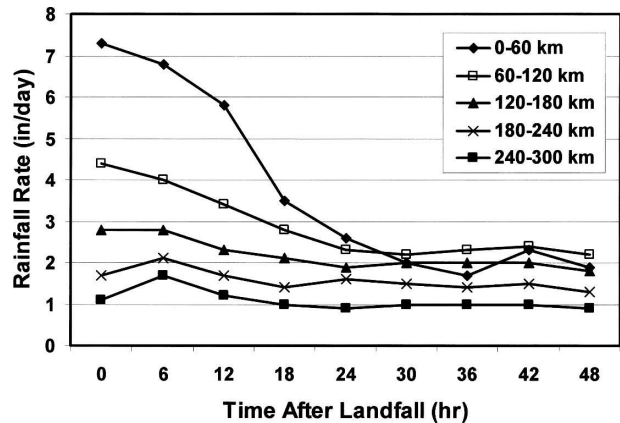


FIG. 4. The rainfall rate as a function of time after landfall for several radial intervals from the storm center. These profiles are for storms that were of hurricane intensity at landfall.

500-km radius, the rain rates become very small. For this reason, the development of the R-CLIPER model only used the rain rates out to 500-km radius. About  $\sim 10^6$  hourly gauge values were within 500 km of the storm centers.

Based upon the exponential appearance in Fig. 3, the gauge rain rate as a function of distance from the storm center  $r$  and time  $t$  subsequent to landfall— $GRR(r, t)$ —was modeled with

$$GRR(r, t) = A(t) \quad r < r_m \quad (1a)$$

and

$$GRR(r, t) = A(t) \exp[-(r - r_m)/r_e] \quad r \geq r_m, \quad (1b)$$

where  $A(t)$  is a function describing the temporal function of the rain rate (see below),  $r_m$  is the radial extent of the inner-core rain rate, and  $r_e$  is a measure of the radial extent of the tropical system rainfall.

Figure 4 shows the gauge rainfall rate in 60-km radial intervals out to 300 km as a function of time after landfall for the hurricane cases. For the innermost radial intervals, the rain rate varies slowly for the first 6 h and then appears to decay exponentially after that time. However, for each interval the rain rate does not decay to zero, but to some nonzero constant. Based upon this behavior, the temporal variation of the rain rate  $A(t)$  was modeled out to 500 km as

$$A(t) = ae^{-\alpha t} + b, \quad (2)$$

where  $a$ ,  $b$ , and  $\alpha$  are empirically determined constants.

The functional form of  $GRR$  in Eqs. (1) and (2) has five free parameters ( $r_m$ ,  $r_e$ ,  $a$ ,  $b$ , and  $\alpha$ ). These parameters were obtained from a least squares fit to the binned rain rate data as a function of  $r$  and  $t$ , where  $r$  ranged from 0 to 500 km and  $t$  ranged from 0 to 48 h.

TABLE 1. The constants from the fit of the hourly rain gauge data as a function of radius and time inland for the gauge R-CLIPER model for the cases where the storm is of tropical storm or hurricane intensity at landfall.

	Tropical storm	Hurricane
$r_m$ (km)	30	20
$r_e$ (km)	188	168
$a$ (in. day <sup>-1</sup> )	2.1	5.8
$b$ (in. day <sup>-1</sup> )	1.4	1.9
$\alpha$ (h <sup>-1</sup> )	0.042	0.045

Table 1 shows the values of these parameters for the hurricane and tropical storm cases, where units were chosen to provide the rain rate in units of inches per day. With the parameters in Table 1, the fits of Eqs. (1) and (2) explain 91% of the variance of the binned rainfall rates with a root-mean-square error (RMSE) of 0.45 in. day<sup>-1</sup> (11.4 mm day<sup>-1</sup>) for the hurricane cases; the corresponding numbers for the tropical storm cases are 88% explained variance and a RMSE of 0.27 in. day<sup>-1</sup> (6.9 mm day<sup>-1</sup>). When Eqs. (1) and (2) are used to estimate the rainfall rate along the storm track using the parameters from Table 1, the method is referred to as the gauge R-CLIPER. In the R-CLIPER model code, the track is interpolated to 0.5-h intervals for the integration of the storm total rainfall.

Because the rainfall rate shown in Fig. 3 is a strong function of the storm intensity at landfall, it would be advantageous to further stratify the data into smaller intensity intervals. Unfortunately, the radial and temporal profiles became too noisy when further division was attempted due to the small sample sizes. However, Fig. 4 and the form of Eq. (2) suggested an alternate method for taking into account the effect of storm intensity on rainfall rate. Kaplan and DeMaria (1995) showed that the decay of tropical cyclone maximum winds after landfall can be modeled fairly accurately with an exponential decay equation very similar to Eq. (2). After an initial adjustment right at the coast due to roughness differences between ocean and land, the maximum winds decay exponentially from the value just after landfall to some background wind speed that can be maintained over land (~27 kt). The average maximum wind speed of the hurricane cases in the 53-yr rain gauge sample just before landfall was 91 kt, which is reduced by 10% in the wind decay model to 82 kt. Figure 5 shows the maximum wind from the inland wind decay model normalized by the intensity just after landfall. Also shown in Fig. 5 is the inner-core (0–60 km) average rainfall rate normalized by its initial value. This figure shows that the maximum winds and rainfall rates have similar time evolutions (the correlation of

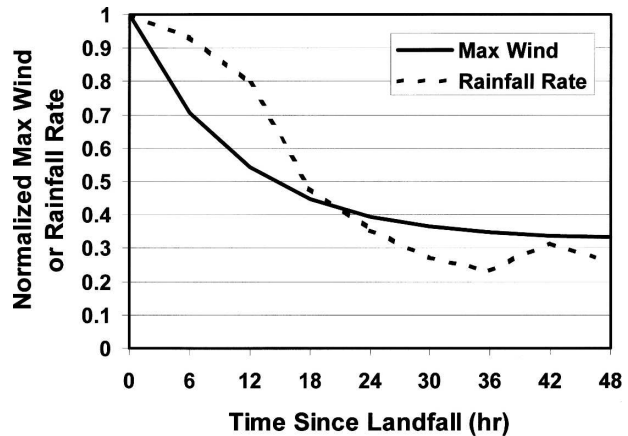


FIG. 5. The maximum wind from the Kaplan and DeMaria inland wind decay model normalized by the intensity just after landfall, and the inner core (0–60 km) gauge rainfall rate (hurricane cases only) normalized by the rainfall rate at landfall, as a function of time inland.

the two time series was 0.92). This similarity suggests that the average decay of the rainfall rate as the storm moves inland is related to the decay of the primary circulation of the storm. Thus, if an accurate estimate of the storm rainfall rate as a function of storm intensity can be obtained, then that relationship could be used over land or water, given the track and maximum wind.

Lonfat et al. (2004) presented a comprehensive satellite climatology of tropical cyclone rainfall rates determined from the Tropical Rainfall Measuring Mission (TRMM). This sample included 260 global tropical cyclones from 1998 to 2000 and was large enough to further stratify the cases of hurricane intensity (maximum winds >64 kt) into category 1–2 (maximum winds of 64–99 kt) and category 3–5 (maximum winds ≥ 100 kt) cases. Figure 6 shows the radial profiles of the TRMM rainfall rates for the three categories of intensity (tropical storm, categories 1–2, and categories 3–5). As with the gauge data, there is a strong relationship between TRMM rain rate and storm intensity. The rain-rate profiles from the gauge data are also shown in Fig. 6, where the data were resampled into 10-km radial bins to match the TRMM data. Although the rain gauge profiles are somewhat noisy with this smaller sampling interval, the agreement between the gauge and TRMM profiles is remarkable, given the difference in the way the rainfall was measured. A correlation of the TRMM and gauge radial profiles in Fig. 6 explains 96% of the variance for the tropical storms and 95% for the hurricanes (TRMM category 1 and 2 cases). The average absolute difference in the TRMM and gauges rain rates is only 0.24 in. day<sup>-1</sup> for the tropical storm radial profiles and 0.27 in. day<sup>-1</sup> for the hurricane profiles. This

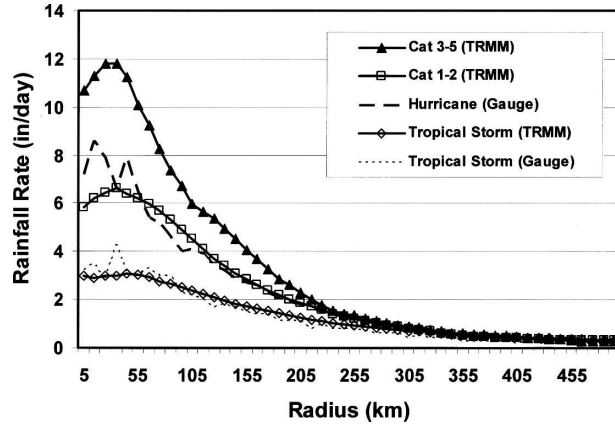


FIG. 6. Radial profiles of the rainfall rates estimated from 260 global tropical cyclones using TRMM and from 120 U.S. land-falling tropical cyclones using rain gauges. The gauge profiles are from the same data shown in Fig. 3 but rebinned to match the radial spacing of the TRMM rainfall rates.

agreement provides confidence that the TRMM profiles can be used for the climatological rain rate. The advantage of the TRMM data is that it provides rain-rate estimates for category 3–5 storms, which are not well represented in the gauge data, especially near the storm centers where gauges often fail.

The radial structure of the TRMM profiles suggests that the following equation could be used to represent the TRMM rain rates (TRR) as a function of radius and maximum wind ( $V$ ):

$$\text{TRR}(r, V) = T_0 + (T_m - T_0)(r/r_m) \quad r < r_m \quad (3a)$$

and

$$\text{TRR}(r, V) = T_m \exp[-(r - r_m)/r_e] \quad r \geq r_m, \quad (3b)$$

where  $T_0$  is the rain rate at  $r = 0$  and  $T_m$  is the maximum rain rate at  $r = r_m$ . In Eq. (3), TRR varies linearly with radius from  $r = 0$  to  $r = r_m$ , and then decays exponentially for  $r \geq r_m$  starting from  $T_m$ . Equation (3) has four parameters ( $T_0$ ,  $T_m$ ,  $r_m$ , and  $r_e$ ). From a least squares fit of the TRMM radial profiles in Fig. 6, it was found that the dependence on storm intensity could be accounted for by making these four parameters linear functions of the storm intensity as follows:

$$T_0 = a_1 + b_1 U, \quad (4a)$$

$$T_m = a_2 + b_2 U, \quad (4b)$$

$$r_m = a_3 + b_3 U, \quad (4c)$$

and

$$r_e = a_4 + b_4 U, \quad (4d)$$

TABLE 2. The constants from the fit of the TRMM rainfall rates as a function of radius and storm maximum wind for the R-CLIPER model. The bottom four rows are the bias-corrected constants used by the NHC in the operational version.

	Intercepts	Slopes
Preliminary	$a_1 = -2.11 \text{ in. day}^{-1}$	$b_1 = 3.63 \text{ in. day}^{-1}$
	$a_2 = -2.73 \text{ in. day}^{-1}$	$b_2 = 4.24 \text{ in. day}^{-1}$
	$a_3 = 69.1 \text{ km}$	$b_3 = -8.49 \text{ km}$
	$a_4 = 215 \text{ km}$	$b_4 = -35.8 \text{ km}$
NHC	$a_1 = -1.10 \text{ in. day}^{-1}$	$b_1 = 3.96 \text{ in. day}^{-1}$
	$a_2 = -1.60 \text{ in. day}^{-1}$	$b_2 = 4.80 \text{ in. day}^{-1}$
	$a_3 = 64.5 \text{ km}$	$b_3 = -13.0 \text{ km}$
	$a_4 = 150 \text{ km}$	$b_4 = -16.0 \text{ km}$

where  $U$  is the normalized maximum wind given by

$$U = 1 + (V_m - 35)/33 \quad (5)$$

and  $V_m$  is the maximum wind speed in knots. Table 2 lists the  $a_1$ – $a_4$  and  $b_1$ – $b_4$  that were obtained from the fit of Eqs. (4) and (5) to the TRMM rainfall profiles, assuming that the maximum winds of the three profiles in Fig. 6 are 45, 80, and 115 kt. The units of the coefficients were chosen to give the rain rate in units of inches per day. Equations (3)–(5) provide an excellent representation of the TRMM rain-rate profiles. The least squares fit explained 99% of the variance, with an RMSE of only 0.28 in. day<sup>−1</sup> (7.1 mm day<sup>−1</sup>).

The R-CLIPER model with the coefficients in the top part of Table 2 (TRMM R-CLIPER) was implemented on an experimental basis at the NHC, beginning with the 2001 hurricane season. Marks et al. (2002) performed an evaluation of the R-CLIPER results using rain gauge data, which suggested a low bias. They adjusted the coefficients to help correct the bias. The adjusted coefficients are shown at the bottom of Table 2. It is possible that the bias over land in the TRMM R-CLIPER is due to the fact that the inner-core rain rate observed from the gauges shown in Fig. 5 appears to decay more slowly than the maximum winds for the first several hours after landfall. Thus, the original coefficients in Table 2 might be more appropriate for storms over water, but the adjusted coefficients are better over land. Topographic effects on rainfall for land-falling storms can be significant (see Fig. 1 in Appalachian Mountains).

With the coefficients from Table 2, the rain rate as a function of radius can be determined for any wind speed, which can then be integrated along the storm track. The rainfall model with the adjusted TRMM rain rates will be referred to as operational R-CLIPER or simply as R-CLIPER. This version has been run operationally at the NHC since the 2004 hurricane season using their official track and intensity forecasts as input. Figure 2a shows an example of the R-CLIPER rainfall

forecast for Hurricane Fran (1996), where the track and intensity come from either the GFDL or the best-track observations (referred to as best-track R-CLIPER).

### 3. Experimental design

One significant difficulty in evaluating the accuracy of rainfall forecasts for landfalling tropical cyclones is that these systems tend to produce their heaviest rainfall over relatively small areas. Thus, if the spatial resolution of the observations is not fine enough to accurately depict these areas of heaviest rainfall, differences in the forecast field and rain gauge analysis may be due to a lack of observations rather than a bad forecast. To avoid this possibility, in this study the model rainfall values are spatially interpolated to the locations of the gauge data; that is, verification is performed only where there are observations.

As described in the introduction, the high-density daily rain gauge data are used for ground truth. All observations within 800 km of the storm track were included in the verification. Figure 3 showed that the average rain rates become very small for radii greater than about 500 km. However, the GFDL model tracks are not always perfect, so the radial extent was increased to 800 km for the verification. Also, in a few cases the rain rates were larger than the mean at radii outside of 500 km (e.g., when the storm interacted with a trough or frontal zone). The 800-km radius allows these types of cases to be included in the verification. The rain gauge data provide 24-h rainfall totals from 1200 UTC to 1200 UTC the next day. For temporal consistency between the gauges and model forecasts, the GFDL model forecast initialized at 1200 UTC within approximately 24 h of landfall was selected for each storm case. There were 25 U.S. landfalling cases from 1995 to 2002, each with a corresponding 1200 UTC GFDL model forecast as shown in Table 3. A few landfalling storms from this period were not evaluated because the GFDL operational files were not available.

The daily gauge values and the corresponding model rainfall totals were summed over time to give a storm total. The storm total period was generally 3 days, but was less if the system dissipated and/or became extratropical before 3 days according to the NHC best track. The daily gauge data were summed over the same time periods as the GFDL model forecast runs shown in Table 3, and are always a multiple of 24 h. Note that for a particular case, every gauge within 800 km of the storm track was summed for the same time period. This may allow some extraneous nontropical storm rainfall and exclude some tropical storm rainfall, although the distinction between storm-related and unrelated rain-

TABLE 3. Model cases evaluated, landfall time and date, model start time and date, and forecast length.

Year/storm	Landfall time and date	Model start date (1200 UTC)	Model forecast length (h)
1995/Allison	1400 UTC 5 Jun	4 Jun	72
1995/Erin	0615 UTC 2 Aug	1 Aug	72
1995/Opal	2200 UTC 4 Oct	4 Oct	48
1996/Bertha	0000 UTC 13 Jul	12 Jul	72
1996/Fran	0030 UTC 6 Sep	5 Sep	72
1996/Josephine	0330 UTC 8 Oct	7 Oct	72
1997/Danny	1800 UTC 19 Jul	19 Jul	72
1998/Bonnie	0400 UTC 27 Aug	26 Aug	2
1998/Charley	1000 UTC 22 Aug	21 Aug	8
1998/Earl	0600 UTC 3 Sep	2 Sep	2
1998/Frances	0600 UTC 11 Sep	10 Sep	2
1998/Georges	1130 UTC 28 Sep	27 Sep	2
1998/Hermine	0500 UTC 20 Sep	19 Sep	24
1999/Bret	0000 UTC 23 Aug	22 Aug	48
1999/Dennis	2100 UTC 4 Sep	4 Sep	2
1999/Floyd	0630 UTC 16 Sep	15 Sep	2
1999/Harvey	1700 UTC 21 Sep	21 Sep	4
1999/Irene	2000 UTC 15 Oct	15 Oct	72
2000/Gordon	0300 UTC 18 Sep	17 Sep	2
2001/Barry	0500 UTC 6 Aug	5 Aug	72
2001/Gabrielle	1200 UTC 14 Sep	13 Sep	72
2002/Fay	0830 UTC 7 Sep	6 Sep	2
2002/Hanna	1430 UTC 14 Sep	14 Sep	4
2002/Isidore	0630 UTC 26 Sep	25 Sep	48
2002/Lili	1400 UTC 3 Oct	2 Oct	48

fall is somewhat subjective. The number of stations available for evaluation of the 25 model forecasts is voluminous (32 784), amounting to ~1300 daily gauges per storm. In contrast, the average number of hourly gauge locations in DT was ~200 per storm. One can see the typical difference in area coverage between the hourly and daily gauges in Fig. 7. The total rainfall attributable to the 25 landfalling cases is shown in Fig. 8. Two maxima exceeding 45 in. (1143 mm) are evident along the North Carolina coast and Florida panhandle and are indicative of the storm activity in these two regions and the notable rainfall producers of Hurricanes Floyd, Fran, Georges, and Opal. Extensive rainfall amounts exceed 20 in. (508 mm) along the Gulf and Atlantic coasts extending more than 200 n mi (367 km) inland. These amounts are a significant portion of the 8-yr (1995–2002) total rainfall in these regions. Note the gradual decrease in amounts inland and the tendency for a secondary maximum along the spine of the Appalachian Mountains.

Several measures of forecast skill were used to evaluate rainfall forecast quality. The overall quality was measured by the mean absolute error and bias averaged over all the gauge sites. The accuracy of the forecasted spatial pattern was obtained through the correlation



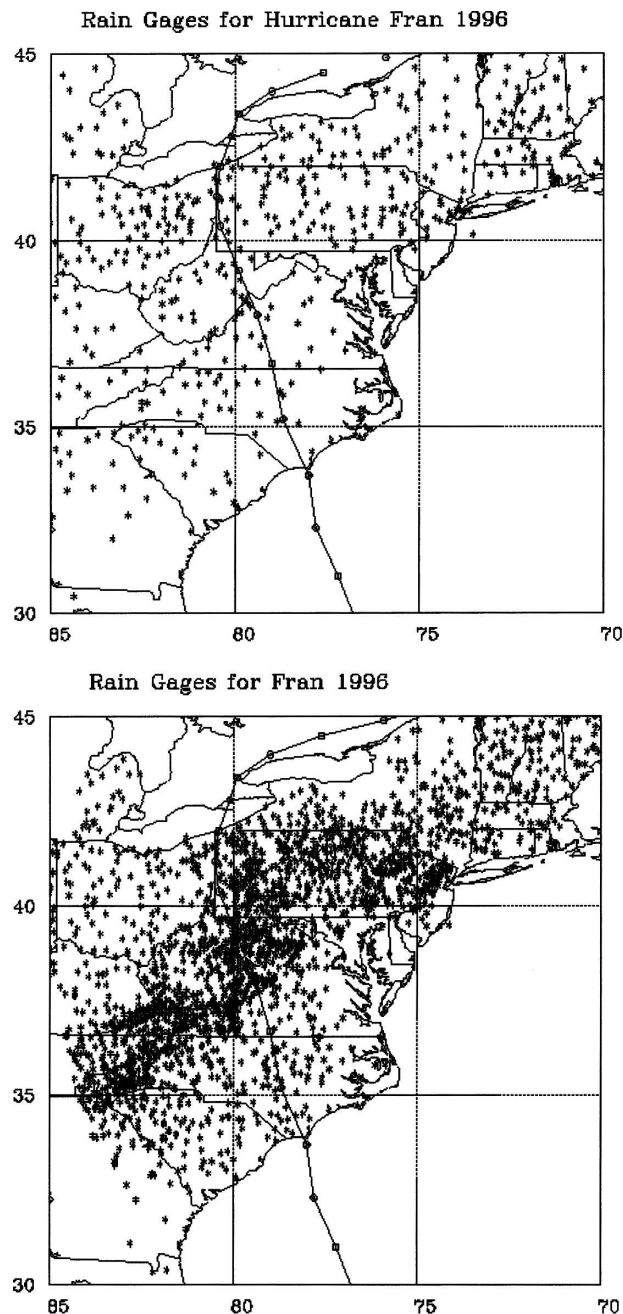


FIG. 7. Locations of (top) hourly and (bottom) daily rain gauges for Hurricane Fran (1996).

coefficient,  $r$ , comparing the observed and predicted rainfall over all gauge sites. These parameters could be obtained both on a case-by-case basis or for the entire set of storms. The preliminary study of DT used these same skill parameters to evaluate rainfall skill. In the current study, additional precipitation verification scores were also calculated including equitable threat and bias scores (Ebert et al. 2003). These measures

compare forecasted and observed rainfall areas (or the number of validation points) equaling or exceeding a specified threshold. In this study, the methods were applied using rainfall thresholds of 0.1, 0.25, 0.50, 1.0, 1.5, 2.0, 3.0, 5.0, and 9.0 in. (2.5, 6.4, 13, 25, 38, 51, 76, 127, and 229 mm). These two measures routinely measure rainfall in a particular fixed geographical domain. In the present study, the domain is variable from case to case and is based on the observed storm track. This is done to restrict the evaluation as much as possible to tropical storm-related rainfall.

The bias score (BS) is defined as

$$BS = F/O, \quad (6)$$

where  $F$  and  $O$  are the numbers of forecasts and observations, respectively, satisfying the threshold criteria. Note that this measure does not involve any specific geographical relation of forecasts to observations and, thus, is independent of track error provided that the forecast track and its associated tropical storm rainfall are within the evaluation domain. Another measure of rainfall skill is the equitable threat score (ETS), which measures the number of forecast fields that match the observed threshold amount for a particular geographic location, which can be written as

$$ETS = (H - H_{\text{random}})/(F + O - H - H_{\text{random}}), \quad (7)$$

where  $H$  is the number of locations where a forecast successfully meets the same criteria as the observation. This “equitable” measure accounts for the random chance that both forecast and observed meet the criteria  $H_{\text{random}} = FO/(F + O)$ .

The ETS, unlike the BS, is therefore quite dependent on the geographical distribution of the predicted rainfall relative to the observed rainfall. For tropical cyclone rainfall, this measure may be a severe test in evaluating the amount and location of rainfall because of the sensitivity to errors in the forecasted track. Additional rainfall verification techniques have been suggested to account for shifts in the rainfall amounts (Ebert et al. 2003). Such adjustments may be especially applicable for mesoscale systems, including tropical storm rainfall patterns. However, this difference is minimized to some extent by evaluating cases about to make landfall where the track error will not be large for forecast periods up to 3 days. As shown in the development of R-CLIPER in section 2a, rainfall is highly controlled by the storm forcing, which in turn is controlled by the storm track. In this study the effects of track and intensity error are investigated by contrasting R-CLIPER results using best-track observations with

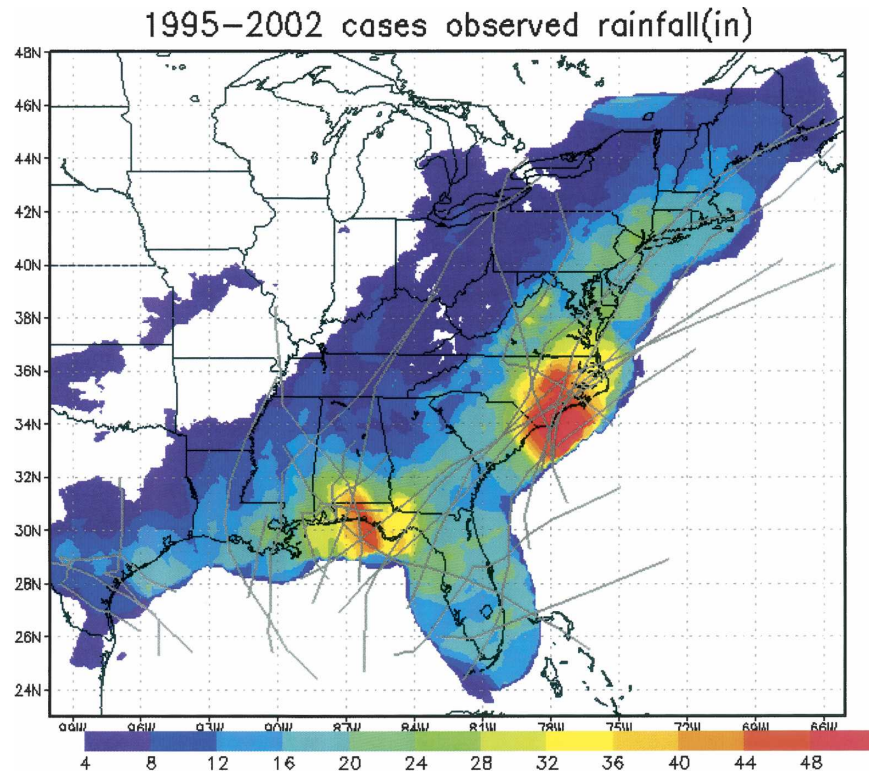


FIG. 8. Total gauge-observed rainfall for all 25 cases evaluated for the 1995–2002 period listed in Table 3. Storm tracks are indicated by gray lines.

R-CLIPER results using the GFDL forecasted track and intensity. In addition, the correlation between GFDL model and gauge rainfall as a function of track error and storm intensity is also calculated. The topic of verifying a 3D model predicted rainfall pattern through techniques that make geographic shifts in rainfall based on track error is left for future study.

#### 4. Results

Table 4 indicates an overall correlation ( $r$ ) of 0.54 between rain gauge measurements and GFDL model forecasts of storm total rainfall for these 25 cases. The mean absolute error and bias for the 25 cases were 0.94 in. (24 mm) and +0.33 in. (8.4 mm), respectively. In the 16-case, low-resolution model study of DT, the correlation coefficient, mean absolute error and bias were 0.48, 1.2 in. (31 mm), and −0.3 in. (−7.6 mm), respectively. The results are similar in the two studies, although the model forecasts are slightly more accurate for the higher-resolution observational and model datasets used in the current study.

In the DT study, the maximum storm total rainfall for each case from any gauge location in the analysis domain was compared to the corresponding parameter

from the model forecast (one value per case) regardless of geographical location. Also in DT, the gauge average rainfall was also compared to the corresponding model rainfall for each of the 16 cases. These parameters can be considered measures of the ability of the GFDL model forecasts to distinguish a “wet” (above average rainfall producing) landfalling tropical storm from a “dry” (below average rainfall producing) one. The mean absolute error, bias, and correlation for both the maximum and average storm rainfall were calculated for the 25-storm sample. All three of the maximum indices were degraded in the present study relative to that in DT, with the correlation falling from 0.79 to 0.50, the mean absolute error increasing from 3.0 to 5.2 in. (76 to 132 mm), and the bias increasing from +0.6 to +1.7 in. (+15 to +43 mm). For the gauge average statistics, the correlation in the present study relative to that in DT degraded from 0.66 to 0.51 although the bias improved from +0.4 to +0.33 in. This degradation may be due in part to the spatial smoothing of the GFDL model output to  $1^\circ$  resolution in DT, which would tend to reduce the maximum amounts—and thus their variation from storm to storm—relative to representing them at  $1/3^\circ$  model resolution. Gallus (2002) found that skill measurements of warm season rainfall often de-

TABLE 4. Overall rainfall statistics for the operational GFDL and several versions of R-CLIPER along the GFDL model forecast track. The R-CLIPER model utilized along the best track is also shown. Statistics are based on the 25 cases of Table 3 over 32 784 matching gauge observations and model predictions of storm total amounts. Also shown is the ratio of the model-predicted to the observed rainfall volume for the GFDL and operational and best-track R-CLIPER models on a  $1/3^\circ$  analysis grid within 800 km of the storm track.

	GFDL model	R-CLIPER			
		Gauge based	Operational	TRMM based	Best track
Correlation coefficient ( $r$ )	0.54	0.34	0.35	0.34	0.49
Mean absolute error (in./mm)	0.94/24	0.94/24	0.92/23	0.91/23	0.81/21
Mean bias (in./mm)	+0.33/+8.4	-0.28/-7.1	-0.34/-8.6	-0.39/-9.9	-0.42/-11
Model volume/observed volume	1.21		0.64		0.61

graded when the verification grid resolution was made finer. Nevertheless, the GFDL model appears capable of predicting the relative amount of rainfall that will fall from one storm to another, explaining  $\sim 25\%$  of the variance of the maximum and domain average rain.

As a benchmark for the evaluation of rainfall forecast skill, the R-CLIPER model was run using the GFDL model forecast track and intensity. As shown in Table 4, the correlation with the gauges was only 0.35, thus indicating some degree of skill for the GFDL model in predicting the precipitation distribution. Both the GFDL and R-CLIPER exhibited a rather large mean absolute error of approximately 0.9 in. (23 mm); the GFDL model was unable to improve upon the much simpler, straightforward R-CLIPER approach according to this measure. Also shown in Table 4 are the results of the gauge R-CLIPER and the preliminary TRMM-based R-CLIPER discussed in section 2b. Note that the operational R-CLIPER model results in slightly better statistics than the other R-CLIPER models run up the GFDL forecast track. The operational R-CLIPER is similar to the gauge R-CLIPER, but has the capability to predict higher rainfall amounts due to a stronger dependence of rain rate on maximum wind speed [cf. Eq. (3) with Eq. (1), which does not explicitly account for maximum wind speed].

To assess the impact of track and intensity error, R-CLIPER was rerun with the best track rather than the GFDL forecast track for all 25 cases. Table 4 shows that the mean absolute error reduced to 0.81 in. (20 mm) and the correlation increased to 0.49 for the best-track R-CLIPER. The elimination of the track and intensity errors thus led to a reduction of 11% for the mean absolute error and increased the variance explained ( $r^2$ ) by 12% points. Note that the GFDL model correlation is slightly better than that of the best-track R-CLIPER, but its mean absolute error is worse. Note that in the best-track R-CLIPER, both the observed track and intensity were used. The best-track R-CLIPER can be considered to represent the upper limit

of predictability for R-CLIPER. In real time, the operational R-CLIPER is run using the official NHC forecast track and intensity.

The mean bias was also computed for these 25 cases. The GFDL mean bias in Table 4 is +0.33 in. (8.3 mm), which can be contrasted with the operational R-CLIPER and best-track R-CLIPER negative values of -0.34 in. (-8.6 mm) and -0.42 in. (-10.7 mm), respectively. An underprediction (not shown) of rainfall amounts greater than 0.5 in. (13 mm) contributes to the general negative bias in rainfall amounts in the R-CLIPER predictions. Also computed in this study was the rainfall volume for the GFDL and operational and best-track R-CLIPER models on a  $1/3^\circ$  analysis grid within 800 km of the storm track. The ratio of the model predicted to the observed volume is shown in Table 4. The GFDL predicts 21% too much rain while the R-CLIPER models predicts  $\sim 60\%$  of the observed rain volume. This is consistent with the mean biases at the gauge sites.

Figure 9 illustrates the ETS as a measure of the relative skill for a distribution of rainfall amounts, including greater than 9 in. (229 mm) to evaluate possible copious storm rainfall at tropical storm landfall. The GFDL model forecast threat score increases sharply with amount up to a peak of 0.34 for a threshold of 0.75 in. (19 mm), followed by a gradual decrease. Both the operational and best-track R-CLIPERs also have the same tendency but have lower scores for all rainfall amounts, again indicating that the GFDL has rainfall forecast skill relative to the climatology-persistence approach. Note that the magnitude of the GFDL skill relative to the R-CLIPERs indicated in Fig. 9 appears to be larger than the global mean statistics in Table 4. This is due in part to the distribution of rainfall being heavily skewed toward low amounts in which 60% of the gauge sites recorded less than 0.25 in. The GFDL ETS skill is actually lower than that of R-CLIPER for these rainfall amounts. The R-CLIPER model is designed to minimize the global mean error. On the other

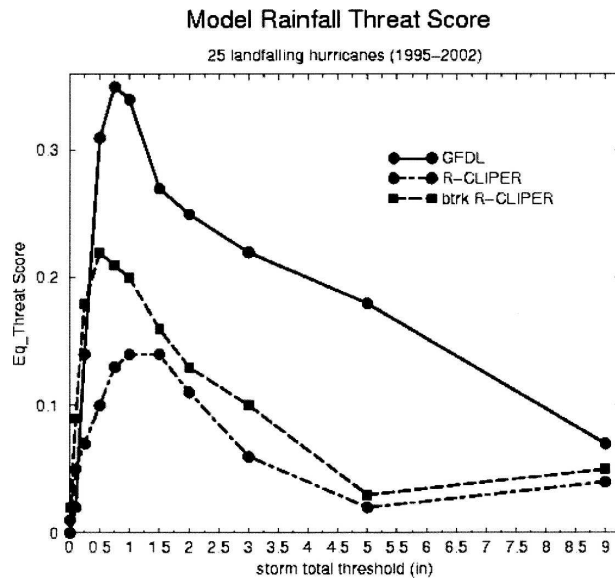


FIG. 9. GFDL threat scores (solid), R-CLIPER scores evaluated along the GFDL track (dashed-dotted), and the NHC best track (long dashed) for all 25 landfalling cases listed in Table 1. The number of observed cases in 0.1-, 0.25-, 0.50-, 1.0-, 1.5-, 2.0-, 3.0-, 5.0-, and 9.0-in. thresholds were 22 064, 19 257, 16 860, 14 000, 11 852, 10 049, 7455, 5377, 2987, 1067, and 106, respectively.

hand, especially at outer radii, considerable azimuthal variations can be anticipated due to the interaction with fronts and other synoptic features that cannot be determined as simple function of radius, intensity, and time after landfall. Thus, for moderate to heavy rainfall, the skill of the GFDL forecasts relative to R-CLIPER is much greater than that implied by the global statistics in Table 4.

The corresponding bias scores are shown in Fig. 10. The R-CLIPERs have a high bias for all thresholds below 0.5 in. (13 mm), and a low bias for higher thresholds. The underestimation of the large rain amounts for R-CLIPER is not surprising because the climatological rain rates represent averages over many storms, with different radial profiles of rainfall that average out to a relatively low value. This is a significant weakness of climatological techniques. On the other hand, the GFDL model has a high bias at all thresholds in Fig. 10. For low thresholds below 0.25 in., GFDL predicts rain at 99% of the gauge sites yielding bias scores exceeding 1.5. This bias is partially responsible for the low threat scores below 0.5 in. (13 mm) seen in Fig. 9. Interestingly, the R-CLIPERs have similarly low threat scores for low rainfall amounts. In contrast to the overprediction of low amounts in the GFDL model, the R-CLIPERs are less biased for these amounts but fail to predict the correct location of much of the lighter rainfall (as indicated by the low ETS).

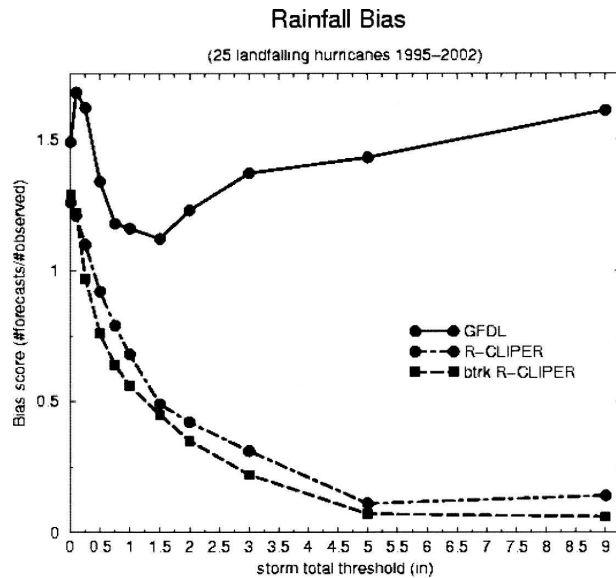


FIG. 10. GFDL bias scores (solid), R-CLIPER bias scores evaluated along the GFDL track (dashed-dotted), and the NHC best track (long dashed) for all 25 U.S. landfalling cases listed in Table 1. Note that the best-track R-CLIPER is run with both observed storm track and intensity.

It was anticipated that track errors would have a detrimental impact on rainfall forecasts. This can be seen in Fig. 9 in the improvement of R-CLIPER evaluated along the observed best track instead of along the GFDL model track. For these cases it appears that the detrimental track effect has a significant impact only for values less than  $\sim 1.5$  in. (38 mm). However, this is probably because of the failure of R-CLIPER to forecast sufficient amounts of heavy rainfall, as indicated by the very low bias scores in Fig. 10. As mentioned, the best-track R-CLIPER is run with both the observed storm track and intensity. Since the bias score is independent of track differences, it appears from Fig. 10 that any influence of GFDL intensity errors does not lead to any further degradation of bias in the R-CLIPER model. Apparently, overprediction of intensity by the GFDL model upon landfall leads to slightly higher rainfall and less bias in the operational R-CLIPER than the best-track R-CLIPER (Table 4). These results appear to support the conclusion that, assuming a relatively good track forecast made near landfall, other factors such as topographical and synoptic forcing must be considered in order to produce forecasts of large rainfall amounts. The GFDL model has these effects included conceptually through the initial conditions and the numerical modeling approach of 3D dynamics and physics parameterizations, allowing it to produce forecasts of heavy rainfall (somewhat too much of it according to Fig. 10), though the low ETS



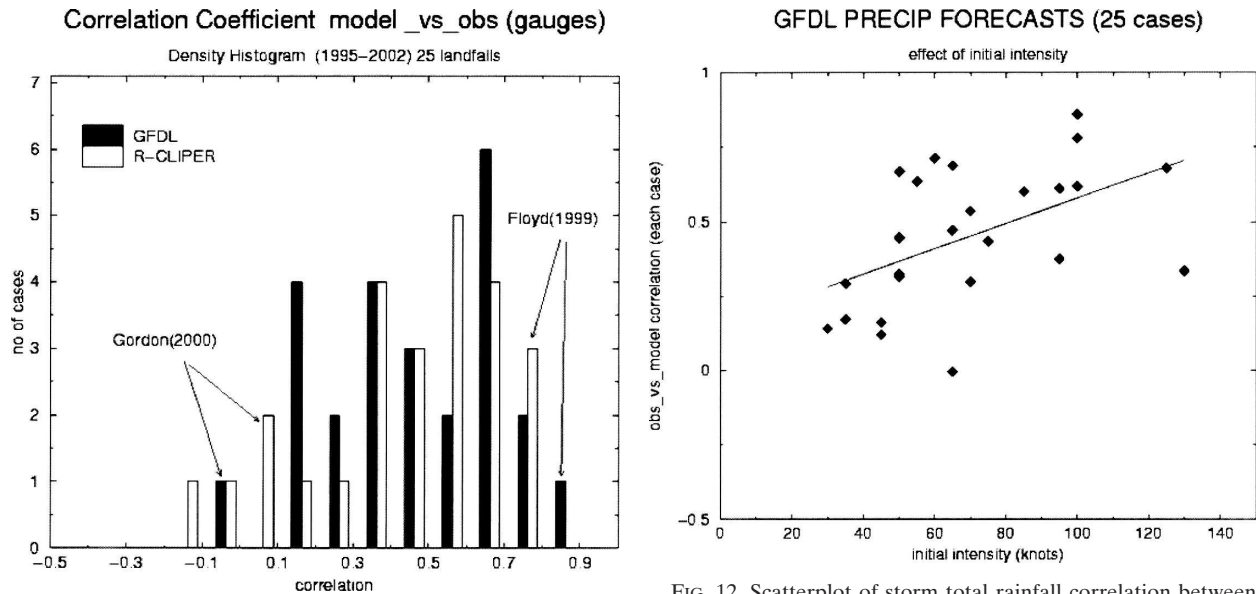


FIG. 11. Histograms of the correlation coefficient between gauge and predicted amounts for each case listed in Table 1. Black denotes GFDL model and white R-CLIPER, both using the GFDL forecast track.

values indicate that the forecasts of location are not very accurate. This can also be seen by the GFDL model rainfall forecast for Fran (Fig. 2) in which it was relatively successful in capturing the heavy rains near the North Carolina–Virginia border as well as the rain in the Shenandoah region of Virginia. The present version of R-CLIPER does not account for these topographical and synoptic effects.

One can also analyze the distribution of GFDL model–gauge correlation (Fig. 11), which indicates a wide spread of performance ranging from a small negative correlations (Gordon in 2000) to correlations of near 0.90 for some cases (e.g., Floyd in 1999). The R-CLIPER model also displays a wide case-to-case variability. For these 25 cases, 11 cases of the GFDL model have correlations above 0.5 and R-CLIPER has 12. By this rough pattern correlation method it appears that both GFDL and R-CLIPER can capture the basic pattern of maximum rainfall along the track for the majority of cases. On the other hand, R-CLIPER had two cases of negative correlation compared to one case for the GFDL model, with the GFDL model having an overall higher correlation over the entire dataset as previously stated. Furthermore, bias score and equitable threat scores indicate that the GFDL model is more capable of forecasting the distribution and location for most amounts. However, the GFDL forecasts of high amounts may be overdone as seen from the high bias of the GFDL model. And as the low ETS indicates, the

FIG. 12. Scatterplot of storm total rainfall correlation between the GFDL model and gauges for 25 cases as a function of initial storm intensity. The line denotes a linear regression relationship between the two.

locations of extreme amounts are rarely correctly forecasted.

The model performance was further investigated by examining the sensitivity of the forecasts to various parameters. Figure 12 shows that there is a weak positive correlation of model performance with initial storm intensity. Strong storms appear to have better rainfall forecasts than weak storms, which would be expected given the generally greater degree of large-scale organization for stronger storms. About 25% of the variance appears to be explained by the initial intensity of the storm for these 25 cases. In a similar manner as shown by observations in Fig. 6, it was found that model maximum and average rainfall increased with the observed intensity of the storm. On the other hand, it is not apparent that the model rainfall prediction skill is related to model intensity skill at landfall. There was little correlation between the model rainfall bias or the over-/underprediction of maximum or the mean rainfall amounts and the over-/underprediction of surface winds. However, since there is quite a large case-to-case variability that may mask any rainfall skill and intensity skill relationship, further research may be warranted.

The effect of track error on GFDL rainfall errors was also investigated. As shown in Fig. 9 the best-track R-CLIPER results improved relative to the R-CLIPER with the forecast track, especially for rainfall thresholds less than 1.5 in. For large rainfall amounts, minor improvements were seen, which may largely be attributed to the overall failure of R-CLIPER to predict sufficient

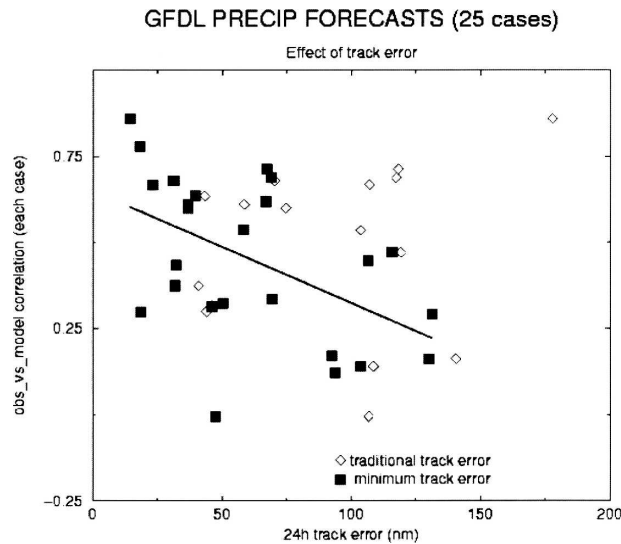


FIG. 13. Scatterplot of storm total rainfall correlation between the GFDL model and gauges for 25 cases as a function of traditional 24-h track error (open diamonds) and “minimum” 24-h track error (black squares). Minimum track error is defined as the error of closest approach of any 6-h forecast position (6–72 h) with that of the observed 24-h observed position. The line denotes a best-fit regression line for the minimum track error.

amounts of heavy rainfall outside the inner core. For the GFDL model in Fig. 13 (open diamonds), there appears to be a small correlation between 24-h mean forecast track error and the GFDL rainfall error. The correlations with track errors at other forecast times are similar. This result may be because of the relatively small track error for these cases just before landfall. One can speculate that for the storm total, the along-track error (i.e., the errors in forward speed) may not be as critical as the cross-track error (i.e., the errors in direction of motion). If one compares the observed 24-h position with the closest approach of the forecast track, regardless of forecast time, the correlation becomes much stronger (approximately  $-0.5$ ), yielding an expected negative correlation between adjusted minimum 24-h track error and forecast–observed rainfall correlation (Fig. 13, black squares with the best-fit line also depicted). In some cases, for example, Floyd (1999), the track error is reduced from  $>175$  n mi to less than 25 n mi with the along-track adjustment. The GFDL model forecast had Floyd move too slowly, but still basically along the observed track. Therefore the correlation pattern of gauge versus model rainfall ( $r = 0.86$ ) was quite high.

## 5. Summary and discussion

A thorough quantitative evaluation of rainfall from 25 U.S. landfalling tropical storms has been performed

for the 1995–2002 seasons for the operational GFDL hurricane model. This is one of the first attempts to quantify the accuracy of model forecasts of tropical storm specific rainfall for a wide variety of U.S. cases. The analysis utilized high-resolution daily rain gauges and  $1/3^\circ$  model resolution and emphasized storm total rainfall near the storm track. The GFDL model was compared to a baseline rainfall CLIPER (R-CLIPER) model to assess relative skill. The details of the development of the R-CLIPER were also described. Both R-CLIPER and GFDL had comparable mean absolute errors of  $\sim 0.9$  in. (23 mm) at 32 784 gauge sites for the combined 25 cases. The GFDL model exhibited a higher pattern of correlation than did R-CLIPER, but still only explained  $\sim 30\%$  of the spatial variance. The GFDL model also had higher equitable threat scores than R-CLIPER, partially because of the known large low bias of R-CLIPER for amounts larger than 0.5 in. The GFDL model suffers from a high bias for practically all threshold rainfall levels. A large case-to-case variability was found that was dependent to some extent on both storm intensity and track error. It appears that this study was successful in evaluating rainfall for these 25 U.S. landfalling tropical storms despite the questions about track sensitivity and model–gauge representativeness and compatibility. It is speculated that by emphasizing storm total rainfall for model forecasts initialized a short time ( $\leq 24$  h) before landfall, some of these problems may have been alleviated. Furthermore, landfalling tropical storms are undoubtedly subjected to topographical and extratropical forcings, which may make skillful rainfall forecasting more complicated yet perhaps more attainable and less sensitive to track errors.

This study should be viewed as just the beginning of a more thorough analysis of tropical storm specific rainfall model forecasts. There are many questions remaining. From the modeling standpoint, the operational GFDL model was changed little for the period of validation from 1995 to 2002. On the other hand, a significant upgrade occurred in 2003. The effects of the upgrade on rainfall forecasts for many of these same cases are the subject of a follow-up study. In addition, this upcoming study on tropical storm rainfall will compare the GFDL model with both the NWS global and regional forecast models. From the observational viewpoint, the present study utilized a straightforward, yet not optimal, approach of analysis at gauge sites only. A multisensor approach of integrating gauge, radar, and satellite data into a consistent gridded database is probably a more sound approach. This may overcome to some extent the known negative bias of rain gauges at high wind speed (Groisman and Legates 1994) and de-

termine whether the high bias of GFDL rainfall is exaggerated by using rain gauges for verification. The upcoming study will utilize the multisensor gridded database as a validation dataset and account for track deviations in its validation methodology.

Another interesting result of this study is that the simple R-CLIPER model provides mean rainfall errors similar to those of the GFDL model. The R-CLIPER model could be improved by making the rainfall rate a function of the environmental flow, which has been shown to induce asymmetries in the rainfall rate. The hourly rain gauge data could also be used to develop an orographic correction. In addition, the initial rainfall rates used in R-CLIPER could be adjusted to match rainfall rates estimated from satellites (e.g., Scofield and Kuligowski 2003). The use of satellite rainfall rates in an extrapolation method for short-term tropical cyclone precipitation forecasts in a method called the tropical rainfall potential (TRaP) has shown promise (Kidder et al. 2005). The combination of the TRaP approach with a generalized R-CLIPER that included synoptic and topographic effects might provide a skillful rainfall forecast relative to the baseline R-CLIPER. Another interesting and quite relevant topic is the use of quantitative model rainfall forecasts to predict riverflow and streamflow. This topic is left for further research.

**Acknowledgments.** The authors thank Frank Marks and Manuel Lonfat for providing the TRMM rainfall profiles, Roger Phillips for his assistance with the rain gauge data processing, and Evgeny Yarosh of the Climate Prediction Center for providing the daily rain gauge data. They would also like to thank Tim Marchok and Rob Rogers for taking steps to expand this project to other models and verification techniques. Partial funding for this project was through the Joint Hurricane Testbed. Computer time from both the Environmental Modeling Center and GFDL/NOAA were utilized. The views, opinions, and findings in this report are those of the authors and should not be construed as an official NOAA and or U.S. government position, policy, or decision.

#### REFERENCES

- Aberson, S. D., 1998: Five-day tropical cyclone track forecasts in the North Atlantic Basin. *Wea. Forecasting*, **13**, 1005–1015.
- Corbosiero, K. L., and J. Molinari, 2003: The relationship between storm motion, vertical wind shear, and convective asymmetries in tropical cyclones. *J. Atmos. Sci.*, **60**, 366–376.
- DeMaria, M., and R. E. Tuleya, 2001: Evaluation of quantitative precipitation forecasts from the GFDL hurricane model. Preprints, *Precipitation Extremes: Prediction, Impacts, and Responses*, Albuquerque, NM, Amer. Meteor. Soc., 340–343.
- , and J. M. Gross, 2003: Evolution of prediction models. *Hurricane! Coping with Disaster*, R. Simpson, Ed., Amer. Geophys. Union, 103–126.
- Ebert, E. E., U. Damrath, W. Wergen, and M. E. Baldwin, 2003: The WGEN assessment of short-term quantitative precipitation forecasts (QPFs) from operational numerical weather prediction models. *Bull. Amer. Meteor. Soc.*, **84**, 481–492.
- Gallus, W. A., 2002: Impact of verification grid-box size on warm-season QPF skill measures. *Wea. Forecasting*, **17**, 1296–1302.
- Groisman, P. Ya., and D. R. Legates, 1994: The accuracy of United States precipitation data. *Bull. Amer. Meteor. Soc.*, **75**, 215–227.
- Kaplan, J., and M. DeMaria, 1995: A simple empirical model for predicting the decay of tropical cyclone winds after landfall. *J. Appl. Meteor.*, **34**, 2499–2512.
- Kidder, S. Q., S. J. Kusselson, J. A. Knaff, R. R. Ferraro, R. J. Kuligowski, and M. Turk, 2005: The tropical rainfall potential (TRaP) technique. Part I. *Wea. Forecasting*, **20**, 456–464.
- Knaff, J. A., M. DeMaria, B. Sampson, and J. M. Gross, 2003: Statistical, 5-day tropical cyclone intensity forecasts derived from climatology and persistence. *Wea. Forecasting*, **18**, 80–92.
- Kurihara, Y., 1973: A scheme of moist convective adjustment. *Mon. Wea. Rev.*, **101**, 547–553.
- , R. E. Tuleya, and M. A. Bender, 1998: The GFDL hurricane prediction system and its performance in the 1995 hurricane season. *Mon. Wea. Rev.*, **126**, 1306–1322.
- Lonfat, M., F. D. Marks, and S. Chen, 2004: Precipitation distribution in tropical cyclones using the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager: A global perspective. *Mon. Wea. Rev.*, **132**, 1645–1660.
- Marks, F. D., G. Kappler, and M. DeMaria, 2002: Development of a tropical cyclone rainfall climatology and persistence (R-CLIPER) model. Preprints, *25th Conf. on Hurricanes and Tropical Meteorology*, San Diego, CA, Amer. Meteor. Soc., 327–328.
- Mellor, G. L., and T. Yamada, 1974: A hierarchy of turbulence closure models for planetary boundary layers. *J. Atmos. Sci.*, **31**, 1791–1806.
- Riehl, H., 1954: *Tropical Meteorology*. McGraw-Hill, 392 pp.
- Rogers, R. F., S. S. Chen, J. E. Tenerelli, and H. E. Willoughby, 2003: A numerical study of the impact of vertical shear on the distribution of rainfall in Hurricane Bonnie (1998). *Mon. Wea. Rev.*, **131**, 1577–1599.
- Scofield, R. A., and R. J. Kuligowski, 2003: Status and outlook of operational satellite precipitation algorithms for extreme-precipitation events. *Wea. Forecasting*, **18**, 1037–1051.
- Tuleya, R. E., 1994: Tropical storm development and decay: Sensitivity to surface boundary conditions. *Mon. Wea. Rev.*, **122**, 291–304.