P2H.9 VERIFICATION OF QUANTITATIVE PRECIPITATION FORECAST GUIDANCE FROM NWP MODELS AND THE HYDROMETEOROLOGICAL PREDICTION CENTER FOR 2005–2007 TROPICAL CYCLONES WITH CONTINENTAL U.S. RAINFALL IMPACTS

Michael J. Brennan, Jessica L. Clark, and Mark Klein
NOAA/NWS/NCEP Hydrometeorological Prediction Center, Camp Springs, MD

1. INTRODUCTION

The National Weather Service's Hydrometeorological Prediction Center (HPC) routinely produces quantitative precipitation forecast (QPF) guidance for days 1 through 5 over the Continental United States (CONUS). This guidance is used by National Weather Service (NWS) River Forecast Centers (RFCs) and Weather Forecast Offices (WFOs) in the preparation of river forecasts and gridded NWS precipitation products. HPC computes skill scores for all its QPFs in addition to corresponding numerical weather prediction (NWP) model QPFs; in particular, threat score and bias (e.g., Wilkes 2006) will be examined here. In 2004, HPC began performing regional verification for its 24-hour QPF products for days 1 through 3. The introduction of this process was the first step in an HPC initiative to assess QPF performance for specific meso-β scale systems. By breaking down the skill assessment of QPFs over the CONUS regionally, HPC can more accurately asses its skill and the skill of the NWP models at forecasting precipitation for specific phenomena, including rainfall associated with tropical cyclones (TCs) and their remnants. This study highlights TC QPF verification from both HPC and NWP forecasts from the National Weather Service's Global Forecast System (GFS), North American Mesoscale (NAM) model, and ECMWF model guidance for the 2005-2007 hurricane seasons. This period was chosen because in 2005, 24-h QPF from the European Centre for Medium-Range Weather Forecasts (ECMWF) global model was added to the guidance suite available at HPC.

Rainfall from TCs can occur well away from the center of circulation and may persist for days after the low-level circulation center has dissipated. This rainfall often generates flash flooding and river flooding. Considering that more annual fatalities in the United States are attributable to flash flooding than any other weather phenomena during the 30-year period from 1977-2006 (NWS 2008), and that freshwater flooding was responsible for more than half of the loss of life in the United States from TCs in the period from 1970-1999 (Rappaport 2000), HPC is particularly interested in improving the skill of TC QPFs in the interest of saving lives and property. The goal of this study is to explicate the methodology, verification, and evaluation of TC QPFs at HPC. Section 2 will describe the QPF forecast and verification methodology at HPC. Section 3 will present seasonal verification results for the 2005, 2006, and 2007 hurricane seasons, while section 4 contains individual case study QPF examples. Section 5 provides a summary and describes upcoming changes to the QPF verification methodology at HPC.

Corresponding author address:
Dr. Michael J. Brennan
5200 Auth Rd., Camp Springs, MD 20746
Email: Michael.J.Brennan@noaa.gov

2. METHODOLOGY

QPF from HPC and three NWP models were verified for TCs which impacted the CONUS from 2005-2007. The models verified included the National Weather Service's GFS and NAM models and the ECMWF model. Table 1 provides a list of those TCs included in this verification study.

Threat score and bias statistics were calculated for the 12-36 hour forecast period from the 0000 UTC model cycle (hereafter denoted as "day 1").

Threat score is defined as:

$$TS = C/(F + O + C) \quad (1),$$

where *F* is the area forecast, *O* is the area observed and *C* is the area "correct" (i.e., where the area forecast for a given threshold overlaps with the area where that threshold was observed). The units of *F*, *O*, and *C* for a specified precipitation threshold amount are in km^2.

Bias is defined as:

$$B = F / O \quad (2),$$

where B is the forecast coverage area (F) for a particular threshold divided by the observed area of that threshold (O), regardless of accuracy in location.

At HPC, the day 1 forecast consists of an accumulation of four spatially averaged QPFs at 6-hour intervals. Due to the variable spatial resolution among the numerical models, each of the forecasts, including those from HPC, were mapped to a standard grid (approximately 32-km grid resolution at 60°N) that has been used for verification at HPC for several years (Charba et al. 2003). This method ensures a fair comparison among each of the forecast products. During the three-year study period, the numerical models underwent several changes, including upgrades to the GFS (before the 2007 tropical season; NCEP 2008b) and ECMWF (before the 2006 tropical season), and a transition in the NAM from the Eta model to the Weather Research and Forecast (WRF) model NMM core prior to the 2006 season (NCEP 2008a). The verification data consisted of a manual, quality-controlled analysis of the Climate Prediction Center's (CPC) daily rain gauge observations, of which there are over 8000 stations reporting across the United States and Mexico.

In order to adequately differentiate precipitation associated with a given tropical system from other synoptic or mesoscale features, the CONUS was divided into 14 climatologically similar regions (Fig. 1). Verification was only performed for those regions in which precipitation was directly associated with the TC or its remnants. The region or regions verified were chosen on a daily basis to capture the movement of the system over time. Determination of the regions was accomplished through examining the observed track of the TC, radar imagery of the event, and quality-controlled quantitative precipitation estimate (QPE) data from the RFCs over the CONUS. While there was some subjectivity to the evaluation of the regions, if it was deemed that a majority of the observed precipitation was not directly associated with the TC, that region was excluded from the statistics for that verifying date.

The day 1 regional verification was performed using the Forecast Verification System (FVS; Brill 2008), which is widely used at both HPC and NCEP's Environmental Modeling Center (EMC) to evaluate model performance. Statistics were calculated for each TC, and summary statistics were also computed for each of the tropical seasons. Additionally, storm-total precipitation analyses for the TCs presented here were produced at HPC using a methodology described by Roth (2008).

3. 2005–2007 SEASONAL VERIFICATION

Skill assessments of HPC QPF and the QPF from the models for the 2005-2007 seasons are shown in Figs. 2-4. For all three seasons, HPC outperformed the NWP QPF model guidance at most thresholds on day 1. HPC QPF generally shows the most improvement over the models for TCs that produce the heaviest, most widespread precipitation. However, HPC tended to have a high bias for amounts over 3 inches during the 2005 (Fig. 2) and 2006 (Fig. 3) seasons, over-forecasting at these thresholds. Conversely, the 2007 season (Fig. 4) saw a low HPC bias for amounts over 3 inches. This may have been, at least in part, due to overcompensation by HPC forecasters who were aware of their previous high bias at these thresholds.

While the NWP models generally do not have seasonal threat scores as high as HPC for amounts less than 3 inches, they show the most skill compared to HPC at these lower thresholds. The NAM is the poorest overall performer, and its threat score drops off dramatically as the threshold amount increases. The NAM also has an extremely low bias for amounts over 2 inches, rarely forecasting amounts that heavy. Overall, the GFS and ECMWF showed similar skill at forecasting precipitation at thresholds 1 inch and under, with the GFS showing more skill at amounts over 1 inch during the 2007 season while the ECMWF was the superior performer at these thresholds during 2006. Interestingly, the GFS and ECMWF were nearly perfect in terms of bias for all thresholds during the 2005 season, but during the 2006 and 2007 season, the GFS displayed a high bias for amounts over 3 inches while the ECMWF demonstrated a significant low bias. Perhaps the changes made to the ECMWF prior to the 2006 season contributed to the decrease in its aerial coverage for precipitation at these higher thresholds.

4. CASE EXAMPLES

Three cases will be presented to highlight the challenges of forecasting TC rainfall, even in the day 1 period. The performance of the model and HPC QPFs will be discussed along with the particular forecast challenges associated with that TC.

a) Hurricane Rita (2005)

Hurricane Rita made landfall on 24 September 2005 near the Texas/Louisiana border (Knabb et al. 2006). After landfall, Rita weakened and moved north-northeast through eastern Texas, Arkansas, southeastern Missouri and southern Illinois. Rita produced heavy rainfall in South Florida as well as across eastern Texas and Louisiana (Fig. 5), with the maximum rainfall amount of 16 in. reported at Bunkie, Louisiana.

The greatest challenge encountered with forecasting rainfall associated with Rita was the forecast motion of the system after landfall. For example, the official track forecast from TPC issued at 2100 UTC 23 September 2005 showed the cyclone moving very slowly over eastern Texas during the entire 5-day forecast period after landfall (Fig. 6). This forecast of very slow motion resulted in predictions of very heavy rainfall amounts (possibly exceeding 25 in.) over eastern Texas and western Louisiana (Rita Public Advisory 23). These totals did not occur, as Rita accelerated north and northeastward after landfall, moving from southeast Texas into central Arkansas in 24 hours (not shown). This type of forecast challenge demonstrates the importance of accurate TC track forecasting after landfall. These track forecasts are often quite challenging due to the potential for weakening, dissipation, or extratropical transition of the TC itself and interaction of the decaying TC with fronts and extratropical cyclones.

Over the course of Rita's life cycle, day 1 QPF forecasts from HPC had threat scores of 0.5-0.6 for thresholds of 1 in. and below (Fig. 7). HPC threat scores were higher than those from the NAM, GFS, and ECMWF through the 4 in. threshold. HPC QPF had a slight high bias for amounts of 1 in. or less, and a low bias for amounts of 2–5 in., corresponding to the trend of the GFS and ECMWF through this range of thresholds. Of the models, the GFS had the highest threat score for rainfall amounts of 2 in. or more, considerably so for amounts of 5 in. or greater (and higher than HPC for the 2 in. threshold). The ECMWF had the highest model threat score for lighter amounts and generally exhibited the best bias of the model guidance through the various thresholds. The NAM had the lowest threat score through all thresholds, as well as a severe low bias for amounts of 1 in. or greater.

On 25-26 September, the remnants of Rita were moving northeastward into the mid-Mississippi Valley and interacting with and ultimately becoming absorbed by a cold front moving eastward across the upper Mississippi Valley (Fig. 8). For the 24-h period ending at 1200 UTC 26 September, HPC QPF indicated two precipitation maxima, one over central and northern Mississippi and a second farther north over southern and central Illinois (Fig. 9a). In both of these areas HPC forecast more than 2 in. of precipitation. Observed rainfall amounts in Mississippi exceeded 2 in. over a larger area than forecast by HPC, while farther north observed totals were in the 1.5–2 in. range. The GFS successfully forecast the existence of a second rainfall maximum farther north, but over-forecast the rainfall amounts in this area (Fig. 9b). The axis of heaviest QPF in the NAM was displaced to the west of that observed (Fig. 9c). Additionally, the NAM failed to bring enough precipitation northward into the observed maximum over Illinois. The ECMWF properly indicated a maximum of precipitation in Mississippi; however forecast amounts were too high, with maxima approaching 4 in. where ~ 2 in. was observed (Fig. 9d). Like the NAM, the ECMWF's forecast amounts were too low farther north in Illinois and Indiana, as the model only forecast a maximum of 1.5 in. where a large area of 1.5-2 in. was observed.

This case highlights the challenge of forecasting QPF after TCs move inland and begin to interact with extratropical features such as fronts and cyclones. Model representation of intensity and motion of the TC vortex and the transport of moisture associated with the TC remnant after landfall is critical to properly forecasting the distribution and amount of precipitation.

b) Tropical Storm Erin (2007)

Tropical Storm Erin made landfall on the central Texas coast on 16 August 2007 (Knabb 2008). Erin produced heavy rainfall across much of south-central and central Texas, extending into

Oklahoma on 17-19 August, including maxima of 10.20 in. in Sisterdale, Texas, and 12.81 in. near Eakly, Oklahoma (Fig. 10).

During Erin, HPC's 24-h day 1 QPF threat score was superior to that of the models for all thresholds except for 0.5 in., where the threat scores of the GFS and ECMWF were equal to or higher than HPC (Fig. 11). HPC's threat scores were much higher than the model QPF guidance at the 5 in. threshold. The ECMWF's threat score decreased dramatically above the 1 in. threshold, as that model suffered a severe low bias at high amounts. These biases were even lower than that of the NAM, which again had the lowest threat score and a very low bias for the heavier thresholds. The GFS showed a low bias for amounts of 4 in. or less, but a very high bias at the 5 and 6 in. thresholds, a trend opposite that displayed by the HPC forecast.

The sample forecast shown for Erin is valid for the 24-h period ending at 1200 UTC 27 August, when over 6 in. of rain was observed in south-central Texas along the track of Erin as it moved inland (Fig. 12). HPC's QPF was quite accurate for this period, showing a maximum of 6.24 in. slightly to the east of where the 6 in. maximum was observed (Fig. 12a). A secondary precipitation maximum occurred farther east near Houston, where over 2 in. of rainfall occurred. Between these maxima, a relative minimum of 0.5-1 in. of precipitation fell, and HPC's QPF showed a high bias in this region. Of the models, the QPF patterns from the GFS (Fig. 12b) and ECMWF (Fig. 12c) indicated that two precipitation maxima would occur, one near Houston and a second well inland. However, the QPF amounts from the GFS and ECMWF were much too low with the inland maximum with forecast amounts of only 1.5-3 in. The NAM QPF (Fig. 12c) for this period was very poor, producing heavy precipitation along the coast but little more than 0.1-0.25 in. where the 6+ in. maximum was observed, as the model failed to produce enough precipitation inland.

This case illustrates the challenges of forecasting QPF for tropical cyclones that are poorly initialized, resulting in poor short term forecasts by NWP models. In this instance, the ability of the GFS to accurately forecast the track of the low-level potential vorticity (PV) maximum associated with Erin as it moved inland (Fig. 13) allowed the GFS to properly forecast the axis of heavy precipitation, even though it was unable to produce the heaviest precipitation amounts. On the other hand, the NAM tracked the low-level PV maximum much too far to the south (Fig 14), resulting in westerly or even west-northwesterly 850-hPa flow into southern Texas, preventing the transport of deep moisture into that region and giving the model little chance to accurately depict the heavy precipitation that occurred well inland.

c) Hurricane Ernesto (2006)

Ernesto made two landfalls in the U.S., the first as a weak tropical storm in Florida on 30 August 2006, and a second landfall as a strong tropical storm in North Carolina on 1 September (Knabb and Mainelli 2006). Ernesto produced a precipitation maximum of 7 inches in southwest Florida and several 10-in. maxima in eastern North Carolina and southeast Virginia, including and a storm-total maximum of over 14 in. at Wrightsville Beach, North Carolina (Fig. 15).

As Ernesto moved northward into the Mid-Atlantic, it became involved with a front and transitioned into an extratropical cyclone late on 1 September (Fig. 16, Knabb and Mainelli 2006). This case exemplifies challenges of forecasting QPF when a TC undergoes the extratropical transition process, which can result in an asymmetric distribution of precipitation around the system, as well as enhancement of precipitation due to interaction with mid-latitude jets and fronts (e.g., Jones 2003, Atallah and Bosart 2003, Atallah et al. 2007).

During Ernesto, HPC's QPF had the highest threat score for amounts of 4 in. and greater, however the higher skill of HPC's QPF for heavier amounts came at the expense of over-forecasting the areal coverage, particularly at the 4 and 5 in. thresholds (Fig. 17). At thresholds of 3 in. and lower, the ECMWF had the most skillful forecasts, with the highest threat score and a near perfect bias. The GFS trailed HPC in terms of threat score at all thresholds, but generally had a better bias. The NAM had the lowest threat score at all threshold values and a very low bias for heavier amounts.

The sample QPF for Ernesto is for the 24-h day 1 period ending at 1200 UTC 1 September 2006, when the system was moving across eastern North Carolina and eastern Virginia. During this period, a large area of 6 in. or more of precipitation occurred along the central coast of

North Carolina (Fig. 18), with much of eastern North Carolina and southeastern Virginia receiving at least 2 in. of rainfall. HPC's QPF accurately predicted that maximum amounts would be greater than 6 in. in eastern North Carolina, but over-predicted the coverage of the aerial extent of amounts of 4 in. or more, bringing the heavy precipitation too far to the west relative to observations. The GFS (Fig. 18b) was more accurate in depicting the sharp western gradient of the heavy precipitation; however the GFS produced too much precipitation farther to the north in eastern Virginia, with totals over 10 in., while only producing 3-4 in. in eastern North Carolina where 6 in. was observed. The NAM (Fig. 18c) concentrated its heaviest precipitation along the coast, depicting 4-5 in. in southeast North Carolina, but failed to extend the heavy amounts far enough to the north. The ECMWF QPF (Fig. 18d) was too light, only producing a maximum of 3.79 in. The ECMWF also showed heavier precipitation extending farther west into the Piedmont, failing to depict the tight precipitation gradient that was observed.

5. SUMMARY

Verification statistics for HPC QPF, as well as QPF from the GFS, NAM, and ECMWF models for TCs with CONUS rainfall impacts from 2005-2007 have been presented. Seasonal summary results as well as examples from individual TCs show that HPC provides considerable added value over raw model QPF guidance, particularly for the heaviest precipitation amounts. This added value is the product of forecaster experience, accurate track forecast guidance from the National Hurricane Center, and locally conducted research, including the development of an extensive TC precipitation climatology.

It is difficult to judge the performance of the NWP models from year to year, given the variation in the number of TCs that produced precipitation impacts in the CONUS and the significant changes that occurred to the NWP models presented here during the study period. However, some general conclusions can be made.

All of the models evaluated in this study struggle to accurately depict the heaviest precipitation amounts, as threat scores for the models decrease substantially for amounts exceeding 3 in. The GFS and ECMWF showed some skill at predicting these heavier amounts in 2005, but their performance dropped off considerably in 2006 and 2007, with threat scores of 0.1 or less for thresholds at and above 4 in. In particular, the ECMWF has displayed a severe low bias for heavy precipitation amounts in 2006 and 2007, after showing very good bias at these thresholds in 2005.

The NAM provided the least accurate QPF guidance of the three models examined here. The NAM's threat scores rapidly decrease for amounts greater than 1 in., and the model, despite its higher resolution, consistently struggles to produce precipitation amounts that exceed 3 in., as the bias for the NAM at these thresholds is close to zero.

The case examples presented here demonstrate some of the challenges associated with forecasting TC QPF, including:

(i) The need for accurate track forecasts of the TC and its remnant vortex from both NHC and model guidance.
(ii) NWP models' difficulty in properly analyzing the TC vortex and the impact of these problems on short term forecasts of track and QPF.
(iii) Properly forecasting the interaction of the TC with fronts and jets, extratropical cyclones, and topography.
(iv) Anticipating the redistribution of precipitation associated with a TC as it undergoes extratropical transition.

Several model changes will occur prior to the 2008 hurricane season. The NAM underwent and upgrade in April 2008 and the GFS is scheduled to undergo an upgrade in May 2008. Also, HPC now has access to 6-h QPF from the ECMWF model that can be utilized in the operational QPF process at HPC; this was not the case during the entire period of this study.

Finally, beginning in 2008, the verification methodology for TC QPF at HPC will be modified. Instead of using predefined verification regions (Fig. 1), the area over which the verification for that TC will be performed will be defined by an amorphous shape that can vary in size and location from day to day. This will allow greater flexibility in performing the verification, and more accurately define the region where precipitation is directly associated with the TC for the computation of verification statistics.

6. ACKNOWLEDGEMENTS

Thanks to David Roth of HPC for producing and providing the TC rainfall accumulation graphics used here. The complete set of accumulation graphics can be accessed online at: http://www.hpc.ncep.noaa.gov/tropical/rain/tcrainfall.html

7. REFERENCES

Atallah, E.H., and L.F. Bosart, 2003: The Extratropical Transition and Precipitation Distribution of Hurricane Floyd (1999). *Mon. Wea. Rev.,* **131,** 1063–1081.

———, ———, and A.R. Aiyyer, 2007: Precipitation Distribution Associated with Landfalling Tropical Cyclones over the Eastern United States. *Mon. Wea. Rev.*, **135**, 2185–2206.

Brill, K./Hydrometeorological Prediction Center, cited 2008: Welcome to FVS -- the EMC/HPC Forecast Verification System. [Available online at: http://www.emc.ncep.noaa.gov/mmb/papers/brill/FVShelpfile07.txt]

Charba, J, D.W. Reynolds, B.E. McDonald, G.M. Carter, 2003: Comparative Verification of Recent Quantitative Precipitation Forecasts in the National Weather Service: A Simple Approach for Scoring Forecast Accuracy. *Wea. Forecasting*, **18**, 161-182.

Knabb, R.D., 2008: Tropical Cyclone Report: Tropical Storm Erin. [Available online at: http://www.nhc.noaa.gov/pdf/TCR-AL052007_Erin.pdf].

———, and M. Mainelli, 2006: Tropical Cyclone Report: Hurricane Ernesto. [Available online at: http://www.nhc.noaa.gov/pdf/TCR-AL052006_Ernesto.pdf].

———, D.P. Brown, and J.R. Rhome, 2006: Tropical Cyclone Report: Hurricane Rita. [Available online at: http://www.nhc.noaa.gov/pdf/TCR-AL182005_Rita.pdf].

NCEP 2008a, Briefing on WRF in NAM. [Available online at: http://www.emc.ncep.noaa.gov/WRFinNAM/].

NCEP 2008b, Model Changes Since 1991. [Available online at: http://www.emc.ncep.noaa.gov/gmb/STATS/html/model_changes.html].

NWS 2008, NWS Weather Fatality, Injury and Damage Statistics. [Available online at: http://www.weather.gov/os/hazstats.shtml].

Rappaport, E.N., 2000: Loss of Life in the United States Associated with Recent Atlantic Tropical Cyclones. *Bull. Amer. Meteor. Soc.,* **81,** 2065–2073.

Roth, D., cited 2008: Tropical Cyclone Rainfall Climatology Methodology. [Available online at: http://www.hpc.ncep.noaa.gov/tropical/rain/TCrainmethodology.html].

Wilkes, D.S., 2006: *Statistical Methods in the Atmospheric Sciences, Second Edition.* Academic Press, 627 pp.

Table 1. Tropical cyclones from 2005-2007 for which day 1 24-hour QPF verification statistics were calculated.

2005	2006	2007
Arlene	Alberto	Andrea
Cindy	Ernesto	Barry
Dennis		Erin
Katrina		Gabrielle
Ophelia		Henriette (East Pacific)
Rita		Humberto
Tammy		T.D. 10
Wilma		Noel

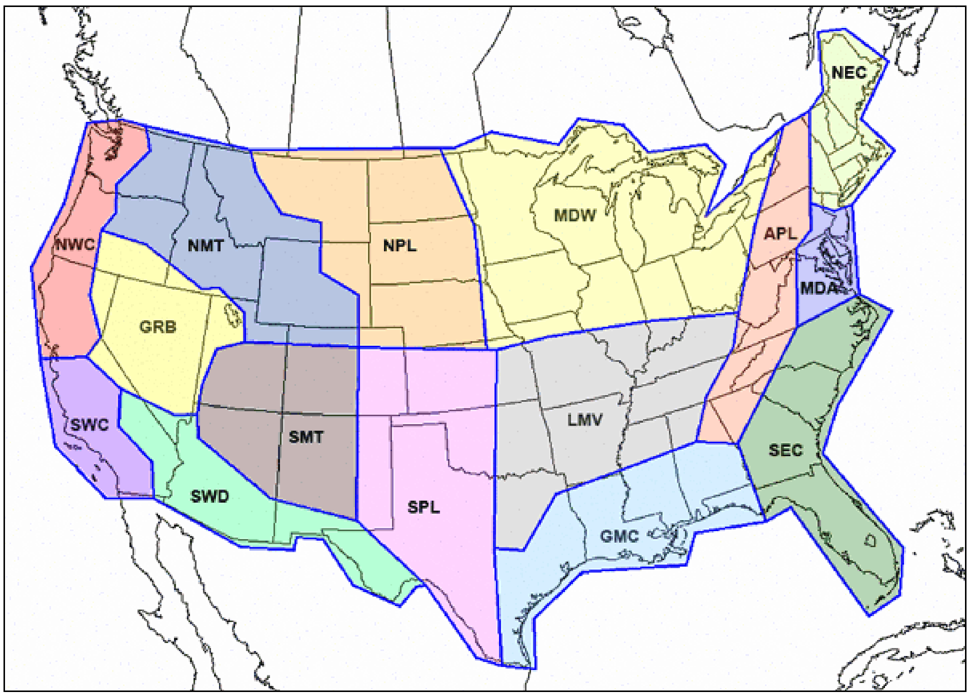

Figure 1. QPF verification regions used in this study.

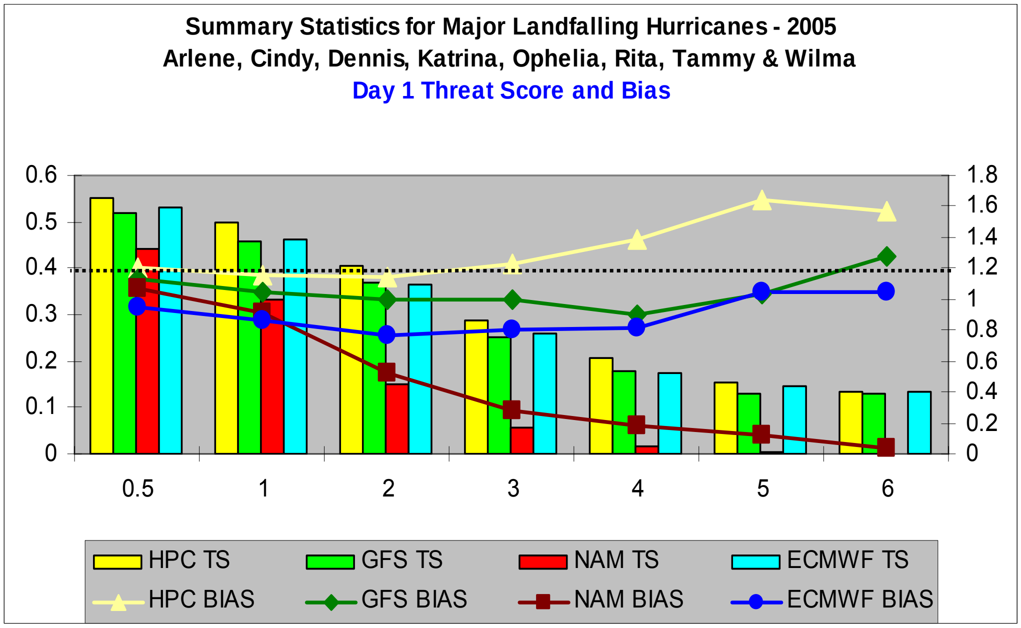

Figure 2. 2005 Seasonal threat score (bar graph, left y-axis) and bias (line graph, right y-axis) for HPC, GFS, NAM, and ECMWF day 1 QPF for thresholds (in.) depicted on x-axis.

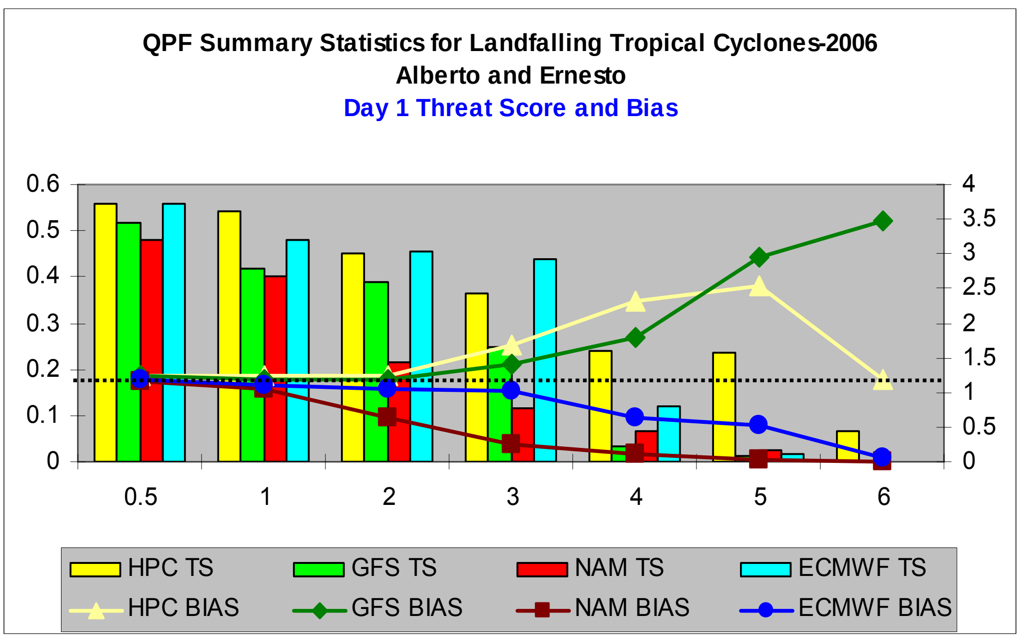

Figure 3. As in Fig. 2, except for 2006.

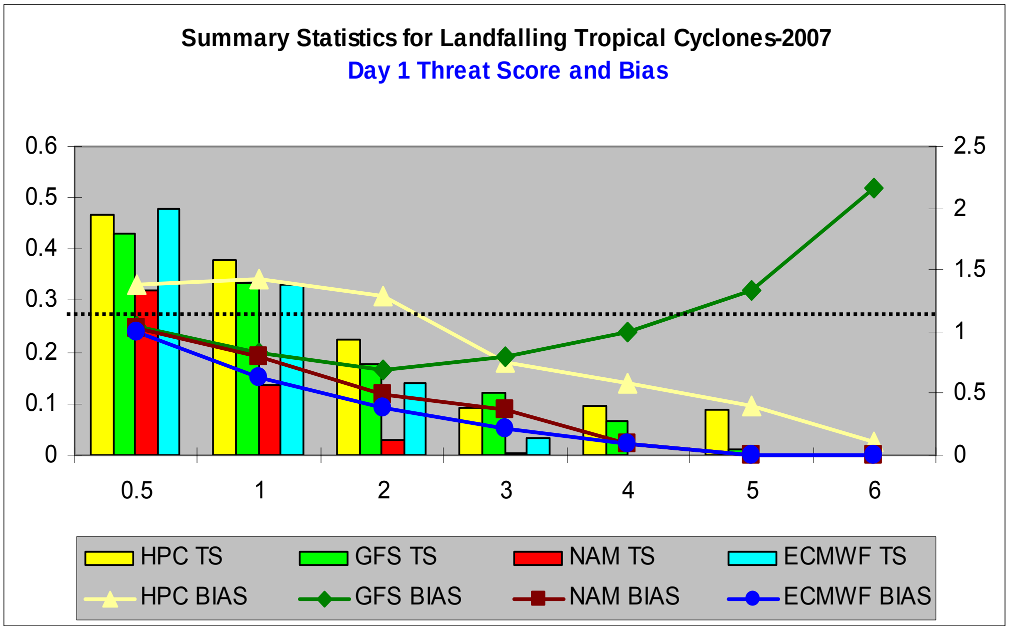

Figure 4. As in Fig. 2, except for 2007.

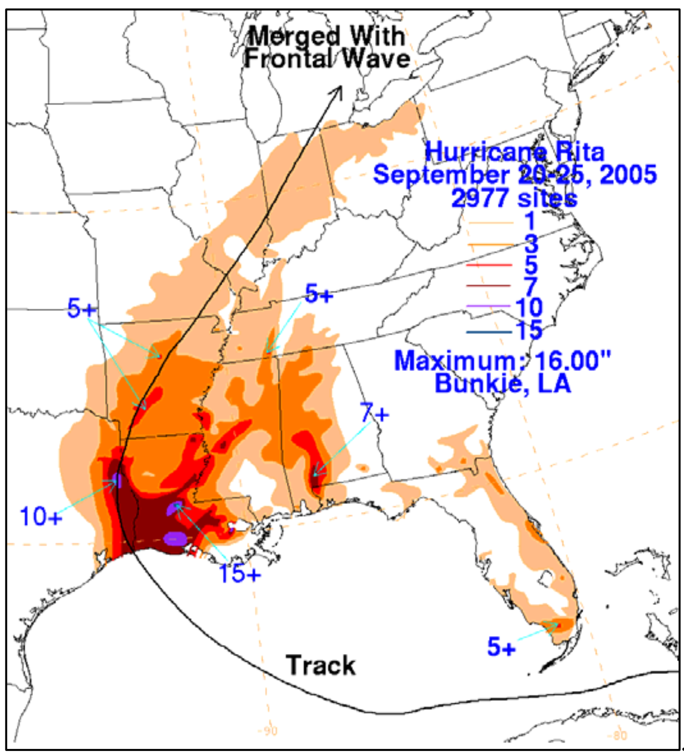

Figure 5. Analysis of storm-total rainfall (.in) associated with Hurricane Rita (2005).

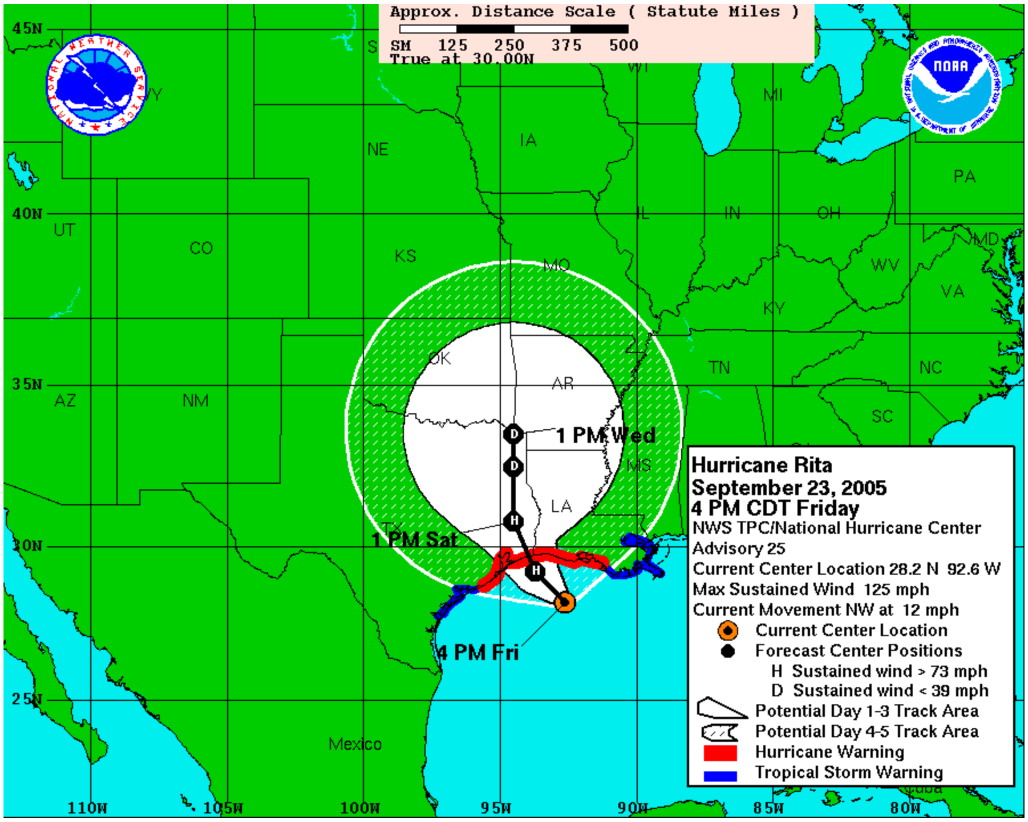

Figure 6. NHC 5-day track forecast for Hurricane Rita issued at 2100 UTC 23 September 2005.

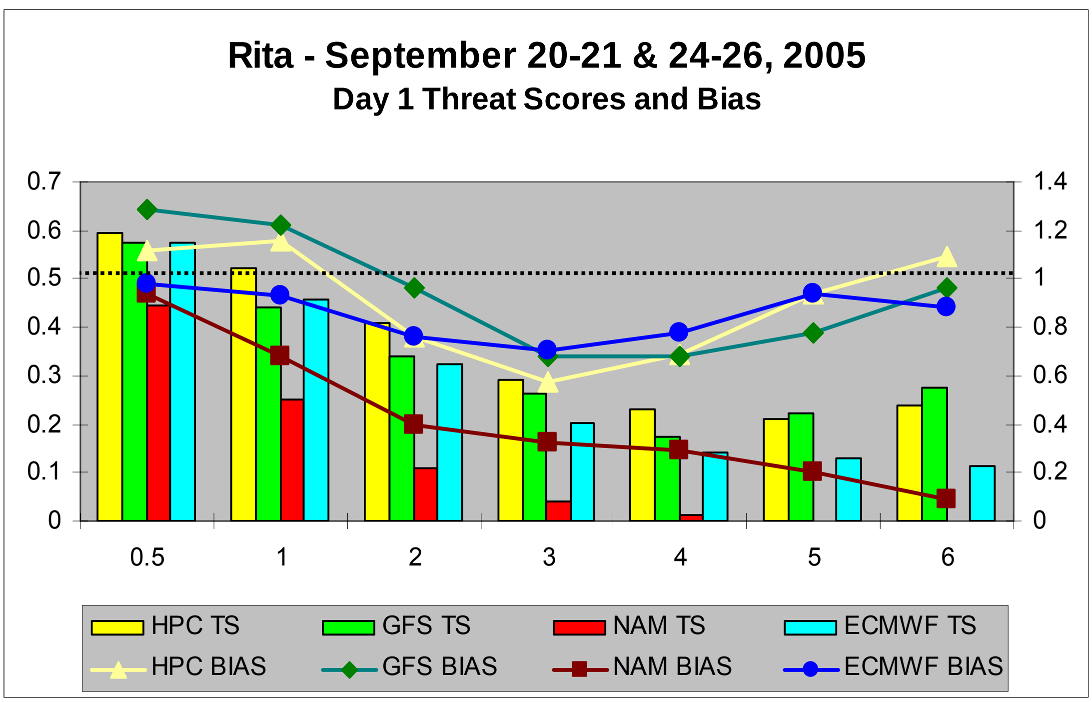

Figure 7. As in Fig. 2, except threat score and bias computed for Hurricane Rita (2005).

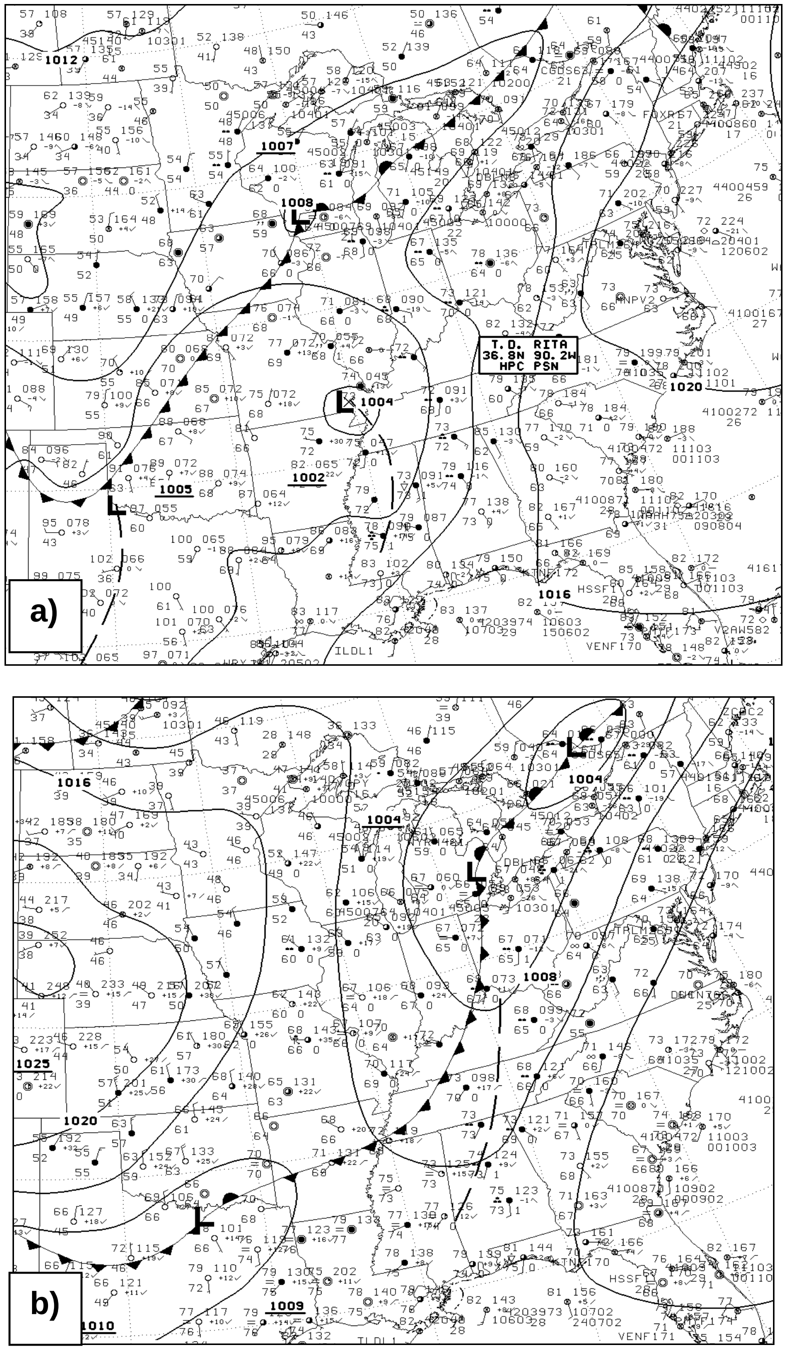

Figure 8. HPC surface analysis valid at (a) 0000 UTC and (b) 1200 UTC 26 September 2005.

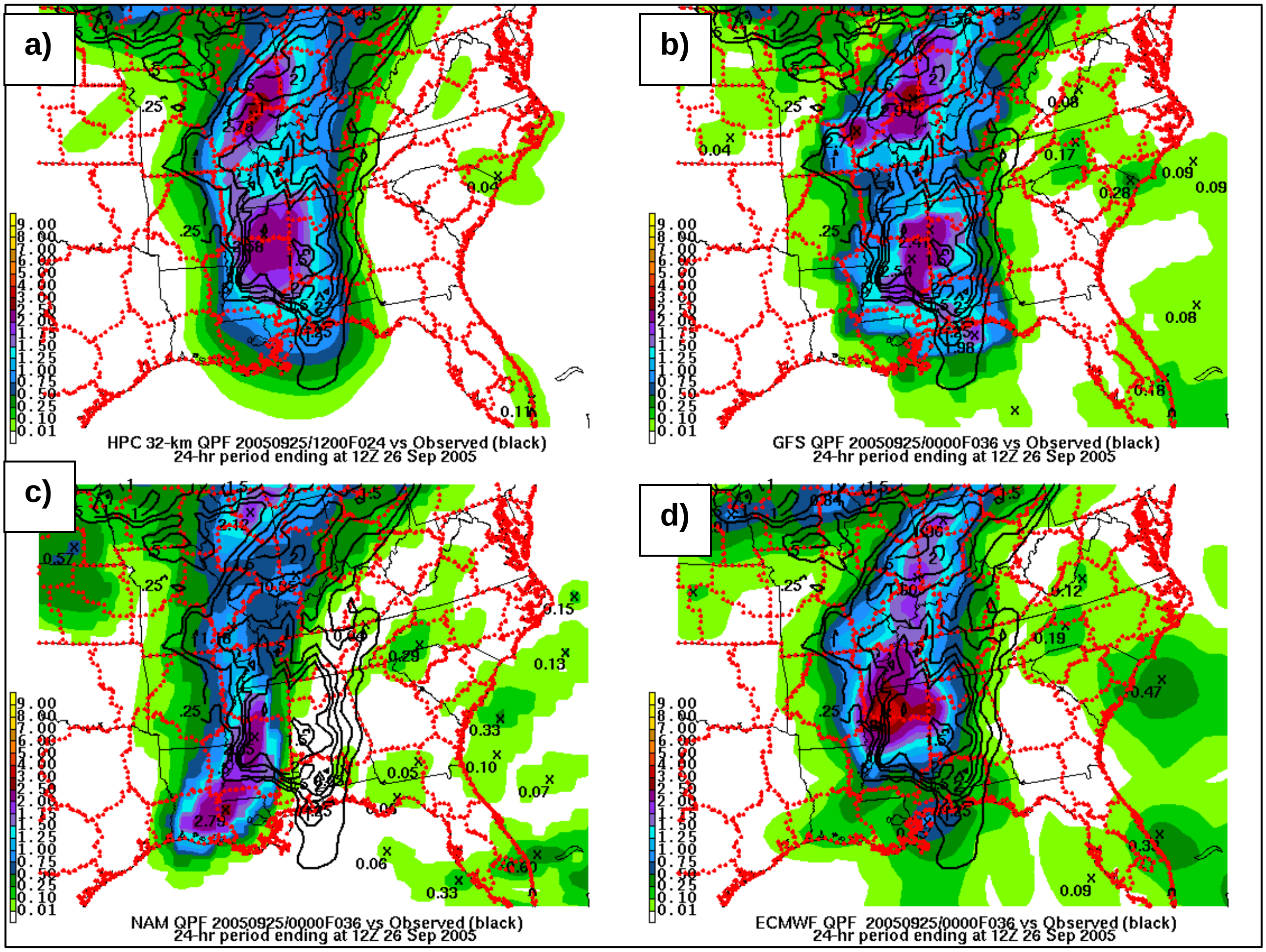

Figure 9. QPF from (a) HPC, (b) GFS, (c) NAM, and (d) ECMWF valid for the 24-h period ending at 1200 UTC 26 September 2005 (shading, in.) and observed precipitation for the same 24-h period (contours, in.).

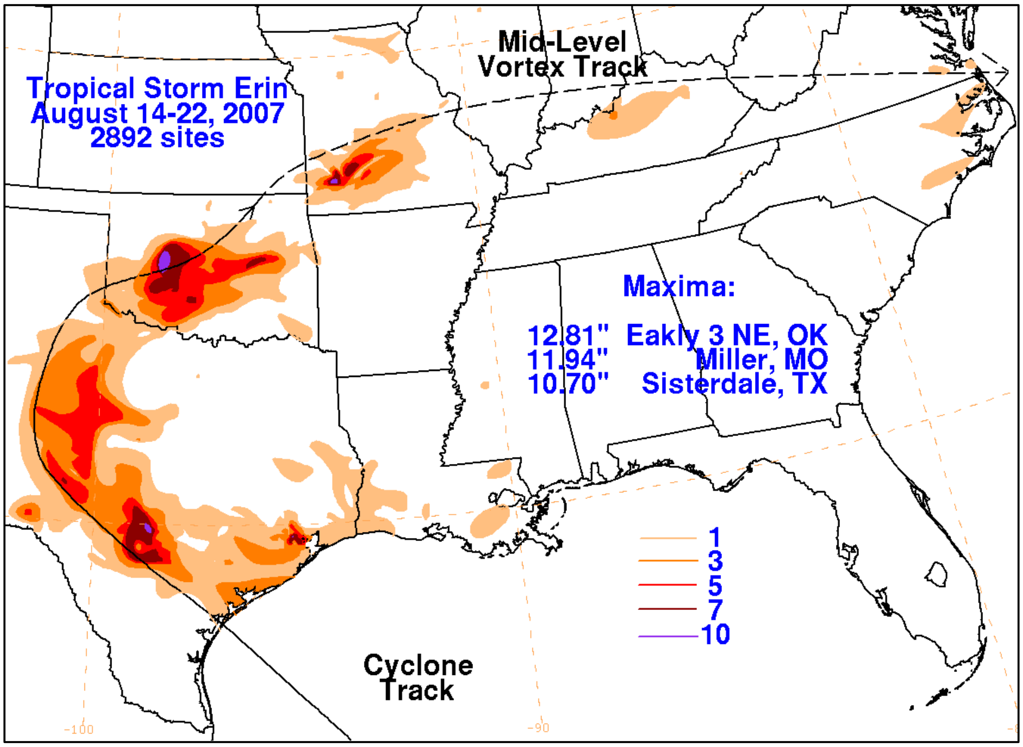

Figure 10. As in Fig. 5, except for Tropical Storm Erin (2007).

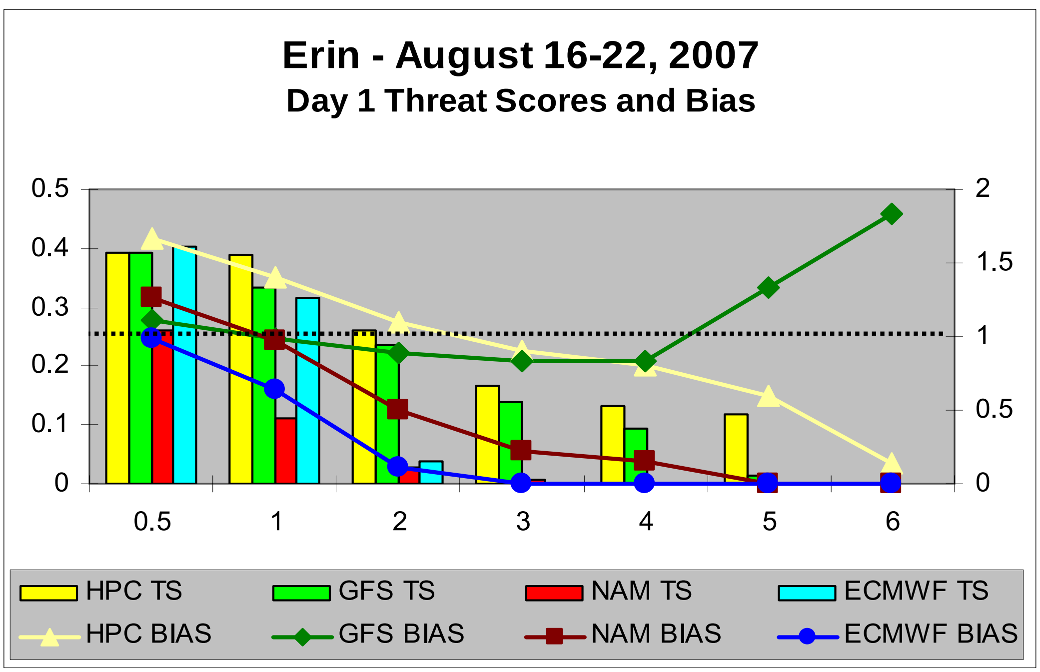

Figure 11. As in Fig. 7, except for Tropical Storm Erin (2007).

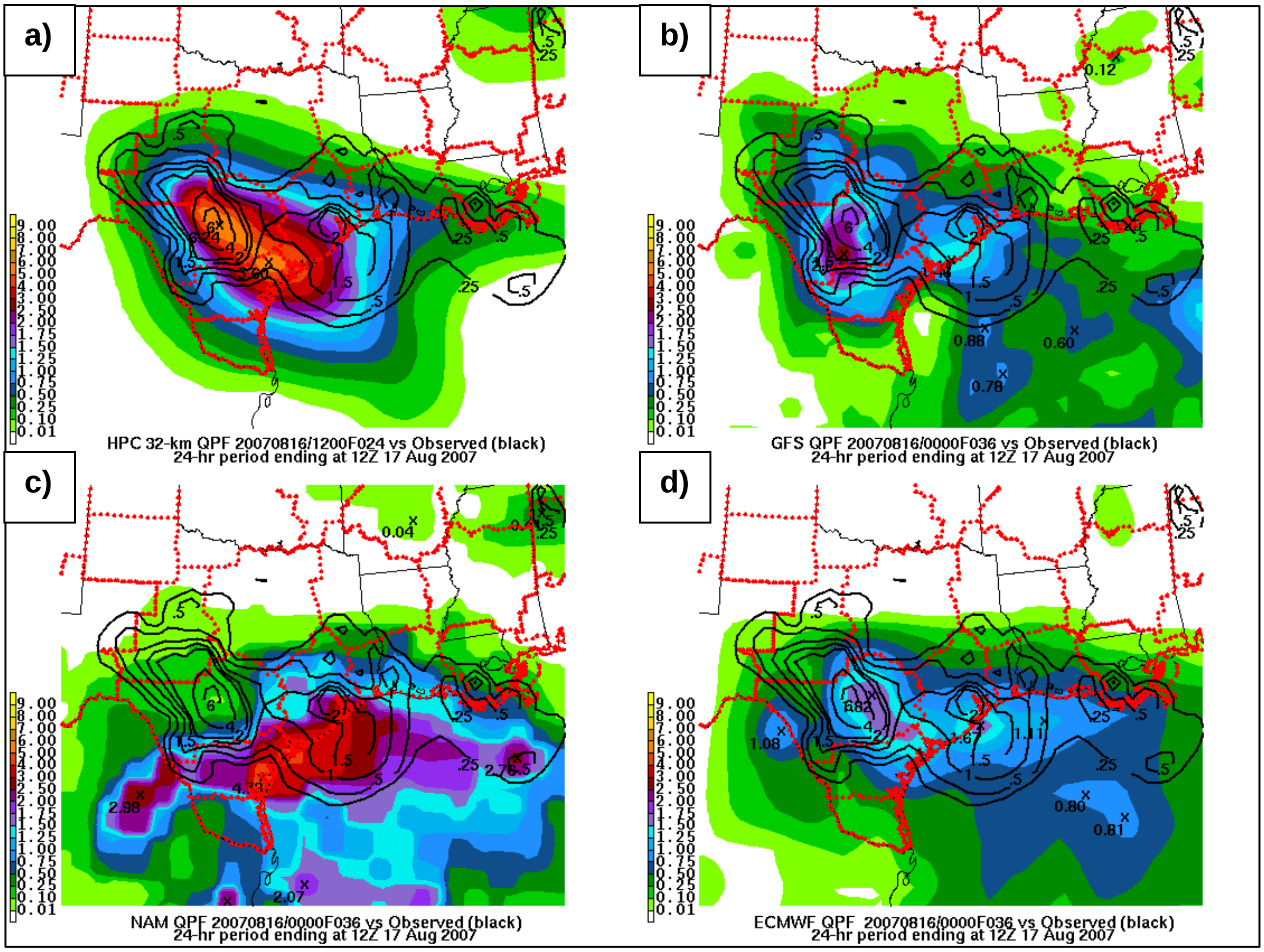

Figure 12. As in Fig. 9, except for Tropical Storm Erin valid for the 24-h period ending 1200 UTC 17 August 2007.

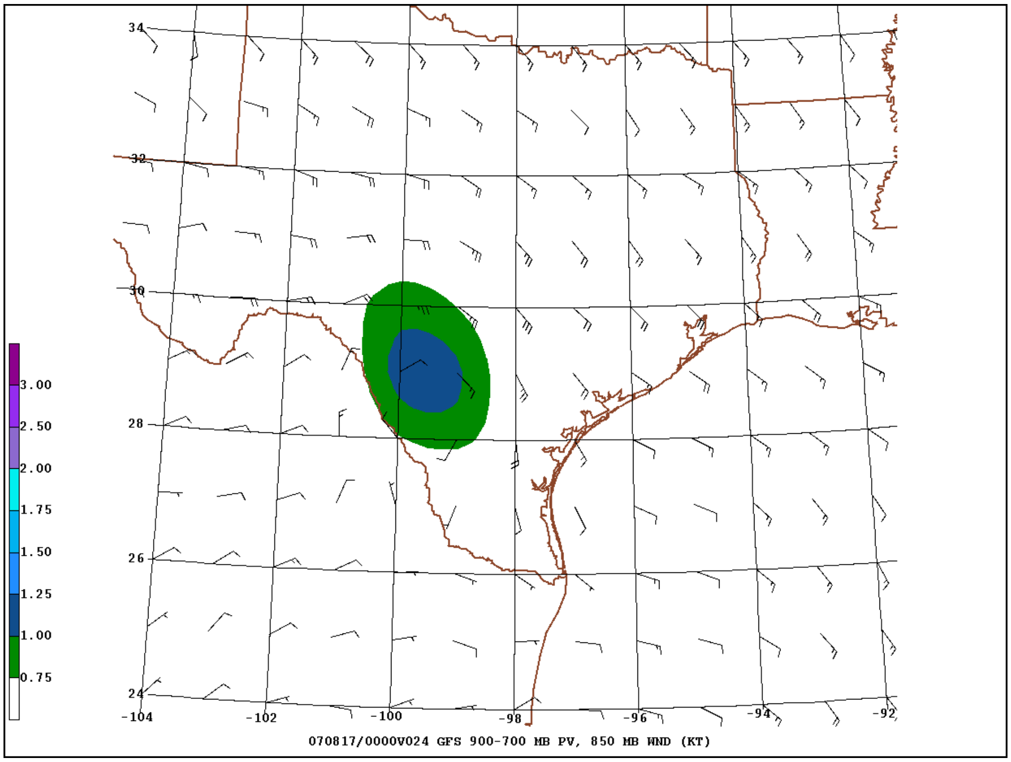

Figure 13. 24-h forecast of 900-700 hPa PV (shaded, PVU) and 850-hPa winds (barbs, kt) from the GFS model valid at 0000 UTC 17 August 2007.

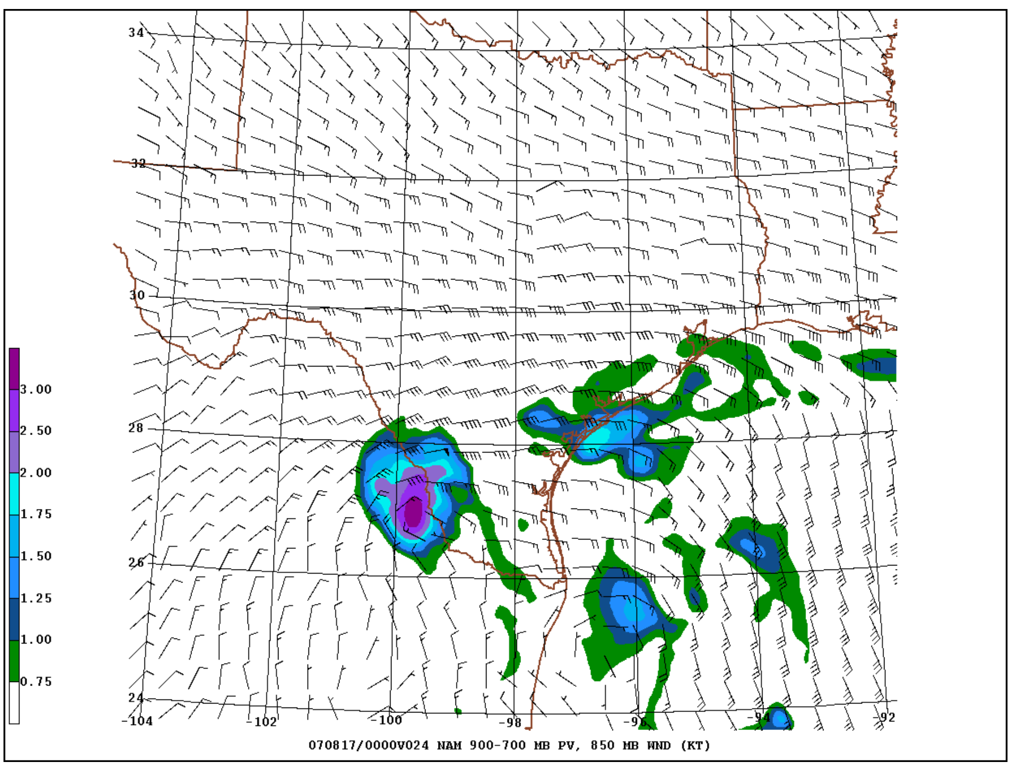

Figure 14. As in Fig. 13, except from the NAM model.

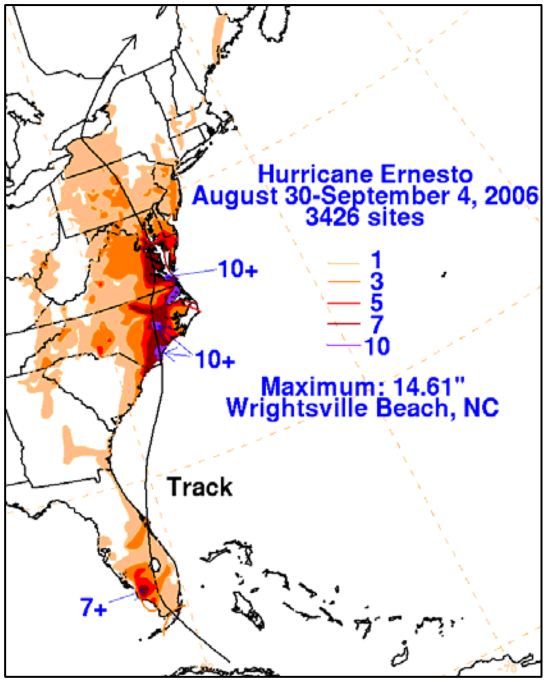

Figure 15. As in Fig. 5, except for Hurricane Ernesto (2006).

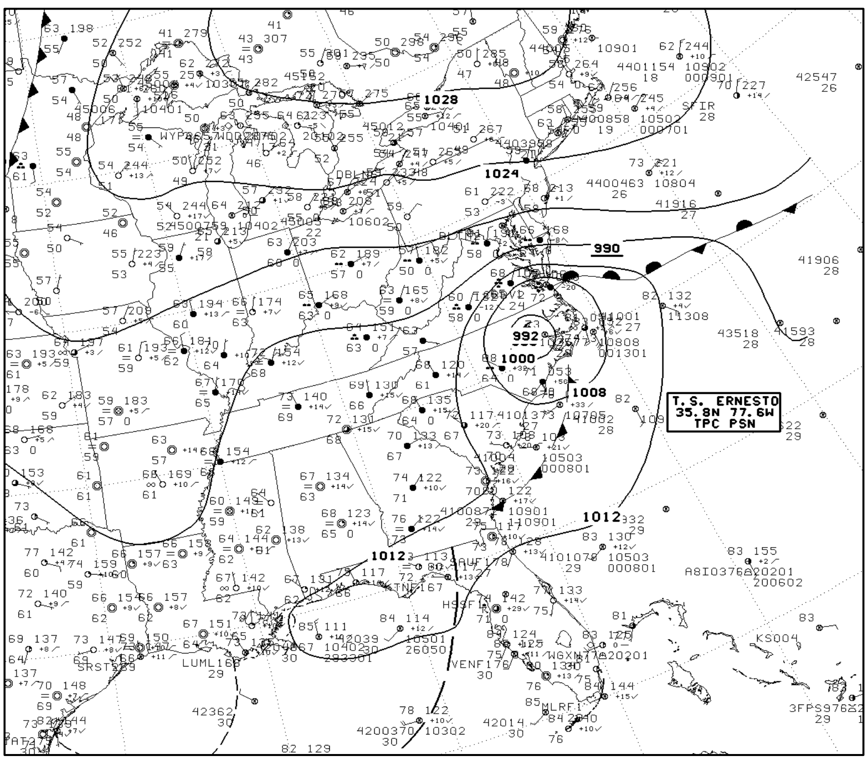

Figure 16. As in Fig. 8, except valid at 1200 UTC 1 September 2006.

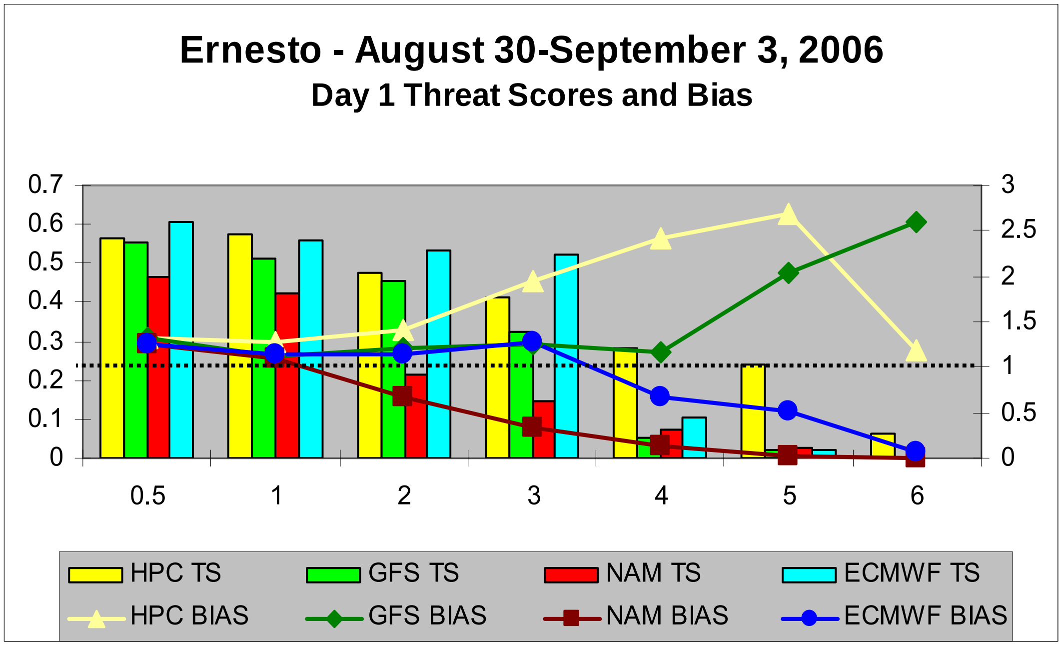

Figure 17. As in Fig. 7, except for Hurricane Ernesto (2006).

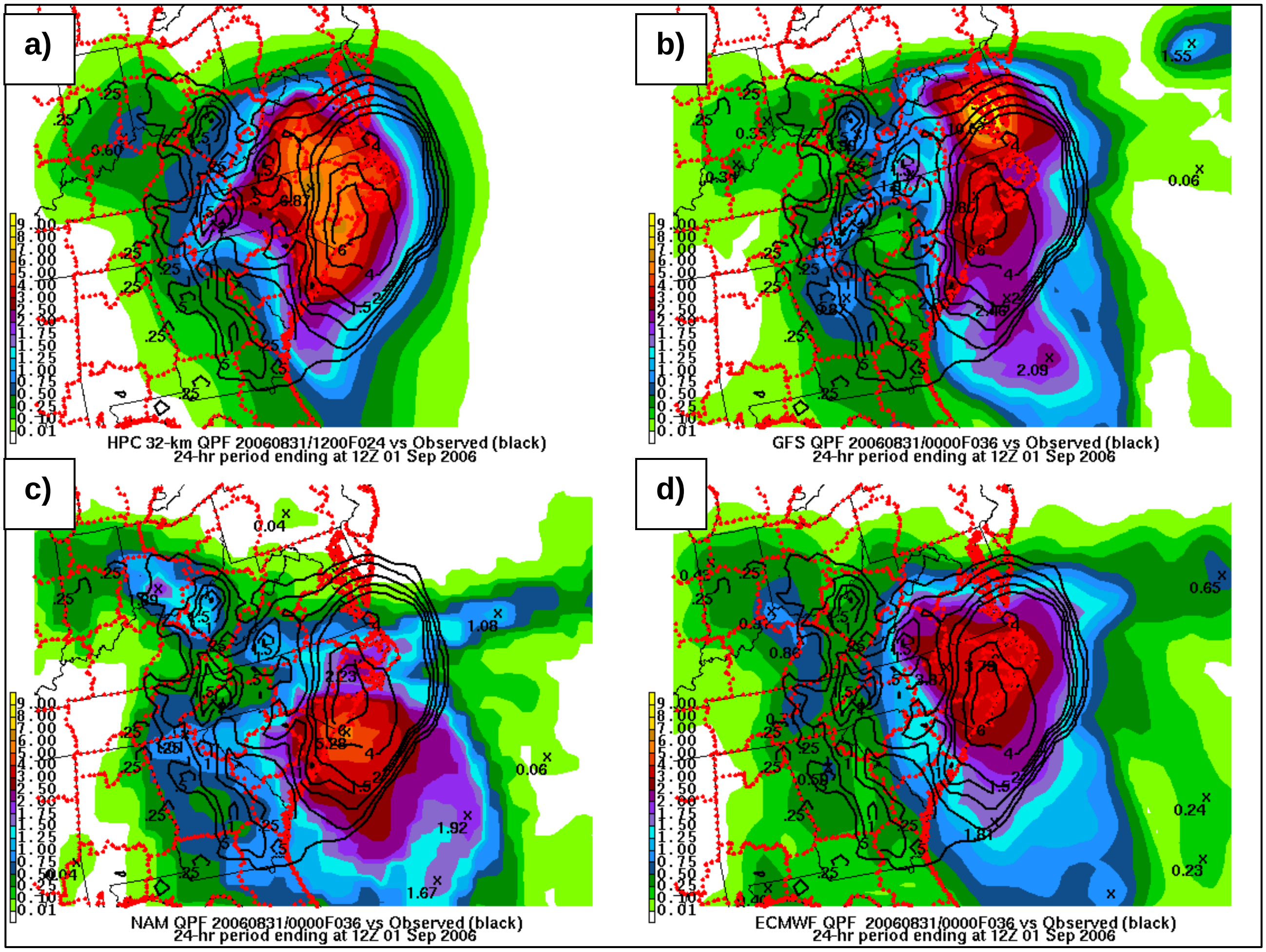

Figure 18. As in Fig. 9, except for Tropical Storm Ernesto valid for the 24-h period ending 1200 UTC 1 September 2007.