NOTES AND CORRESPONDENCE

An Example of the Value of Strong Climatological Signals in Tropical Cyclone Track Forecasting: Hurricane Ivan (2004)

BRADFORD S. BARRETT, LANCE M. LESLIE, AND BRIAN H. FIEDLER

School of Meteorology, University of Oklahoma, Norman, Oklahoma

(Manuscript received 1 March 2005, in final form 28 September 2005)

ABSTRACT

Since 1970, tropical cyclone (TC) track forecasts have improved steadily in the Atlantic basin. This improvement has been linked primarily to advances in numerical weather prediction (NWP) models. Concurrently, with few exceptions, the development and operational use of statistical track prediction schemes have experienced a relative decline. Statistical schemes provided the most accurate TC track forecasts until approximately the late 1980s. In this note, it is shown that increased reliance on the global NWP models does not always guarantee the best forecast. Here, Hurricane Ivan is used from the 2004 Atlantic TC season as a classical example, and reminder, of how strong climatological signals still can add substantial value to TC track forecasts, in the form of improved accuracy and increased timeliness at minimal computational cost.

In an 8-day period in early September 2004, Hurricane Ivan was repeatedly, and incorrectly, forecast by 12 operational NWP models to move with a significant northward (poleward) component. It was found that the mean 24-h trajectory forecasts of a consensus of five commonly used NWP track prediction aids had a statistically significant right-of-track bias. Furthermore, the official track forecasts, which relied heavily on erroneous numerical guidance over this period, were also found to have significant poleward trajectory errors. At the same time, a climatology-based prediction technique, drawn entirely from the historical record of motion characteristics of TCs in geographical locations similar to Ivan, correctly and consistently indicated a more westward motion component, had a small directional spread, and was supported by a large number of archived cases. This climatological signal was in conflict with the deterministic NWP model output, and it is suggested that the large errors in the official track forecast for TC Ivan could have been reduced considerably by taking into greater account such a strong climatological signal. The potential impact of such an error reduction is a saving of lives and billions of dollars in both actual damage and unnecessary evacuations costs, for just this one hurricane. We also suggest that this simple strategy of examining the strength of the climatological signal be considered for all TCs to identify cases where the NWP and official forecasts differ significantly from strong, persistent climatological signals.

1. Introduction

Tropical cyclone (TC) track forecasts in the Atlantic basin have steadily improved over the last 30 yr. Franklin et al. (2003), updating the work of McAdie and Lawrence (2000), found that position errors in the National Hurricane Center's (NHC) official track forecasts for the Atlantic basin decreased at an average annual rate of 1.3%, 1.9%, and 2.0% at 24, 48, and 72 h, respectively, from 1970 to 2001. However, in contrast to the basinwide track forecast improvements, forecasts of landfall location for TCs approaching the U.S. coastline have not improved significantly since 1976 (Powell and Aberson 2001). Powell and Aberson (2001) attribute this lack of significant improvement in part to a "conservative, least-regret" forecast philosophy.

Over this same period, global and regional numerical weather prediction (NWP) models have become much more skillful (Shuman 1989; Kalnay et al. 1990; Caplan et al. 1997; Bender and Ginis 2000; Aberson 2001), and several NWP models designed specifically for TC prediction have been introduced, including the National Oceanic and Atmospheric Administration (NOAA)

Corresponding author address: Bradford S. Barrett, School of Meteorology, University of Oklahoma, 120 David L. Boren Blvd., Norman, OK 73072.
E-mail: bradb@ou.edu

Geophysical Fluid Dynamics Laboratory (GFDL) model (Kurihara et al. 1993, 1995, 1998; Bender et al. 1993) and the Florida State University "multimodel superensemble" (Krishnamurti et al. 1999; 2000a,b; 2001). Recently developed statistical techniques, such as the self-adapting analog ensemble prediction method (Sievers et al. 2000; Fraedrich et al. 2003) and the simple consensus forecasts and ensemble averages determined from multiple NWP models (Goerss 2000; Goerss et al. 2004) also have produced TC track forecasts superior to the individual model components. Because track forecasts from global and regional NWP models have improved so much in the last decade (Weber 2003), they now are used operationally by forecasters at the Tropical Prediction Center (TPC) in Miami, Florida, and the Joint Typhoon Warning Center (JTWC) in Pearl Harbor, Hawaii (Goerss 2000). Much of the steady reduction in TC official forecast errors therefore appears to have resulted from an increased reliance upon improved NWP model forecasts (Sheets 1990; McAdie and Lawrence 2000).

One of the consequences of the rapid advance of dynamical models has been a relative decline in the improvement, development, and operational use of statistical–climatological prediction schemes (Bessafi et al. 2002). However, it will be argued here that these climatology-based models still have and should continue to have a role to play in TC track prediction. For example, they can be used to 1) provide a convenient reference from which to assess the performance of NWP model predictions (Neumann and Pelissier 1981a; Aberson 2001); 2) evaluate forecast difficulty of particular storms and TC basins (Franklin et al. 2003; Goerss et al. 2004); 3) conveniently generate bogus TC tracks (Bessafi et al. 2002); 4) provide very early TC track, speed, and heading forecasts in all localities of a basin (Neumann and Pelissier 1981b); and 5) provide an accurate forecast when departures from climatology and persistence are minor (Neumann and Hope 1972). Such statistical methods are therefore capable of providing rapid and valuable guidance with wide-ranging functionality.

This study has two main objectives. First is to illustrate the continued operational usefulness, and therefore necessity, of statistical methods that rely largely or entirely upon the archived TC climate record. We show that large along-track trajectory (speed and direction) errors can be reduced when there is a strong climatological signal that has a small spread and is based on a large number of archived cases. Hurricane Ivan, which occurred during the 2004 Atlantic season, is an excellent recent example. Ivan was a classical long-lived, long-track major hurricane that was responsible for 25 deaths and over $14 billion in U.S. losses (Stewart 2004). It afforded the NWP models many opportunities to predict its track. However, trajectory forecasts from statistical methods based upon the climate record were significantly more accurate, over an 8-day period, than the tracks predicted by the NWP models. Ivan is an example that demonstrates powerfully that the NWP and official forecasts can have large trajectory errors when their predictions are significantly different from the tracks suggested by the climatological scheme used in this study.

Second, this study reminds TC forecasters and other users of climate data of the continued utility of climatological data, especially when it provides an early means of alerting TC forecasters to NWP predictions that have potentially large track errors. This study and its conclusions are based on a wide range of input data. Over 500 forecasts from 14 different operational NWP models and statistical prediction methods were examined, and over 400 historical Atlantic TC records were used to compute the climatological signal most relevant to Ivan.

Section 2 introduces the TC motion climatology concept employed in this study and describes its computational aspects. Section 3 applies the climatology to Hurricane Ivan, performs a statistical comparison between numerical and statistical prediction methods, and discusses the role of the synoptic-scale steering flow. Section 4 presents our conclusions and suggestions motivated by the results from Ivan.

2. TC motion climatology

Associated with every geographical location in the North Atlantic basin is a TC "motion climatology" derived from the historical movement characteristics of all TCs that passed near it. A technique was developed for this study to calculate and display graphically the TC motion climatology. In brief, the Atlantic TC dataset is used to compute motion tendencies (speeds and directions) of past TCs at or near a specified geographical point (Barrett et al. 2004). The focus here is on the 24-h motion climatology because this time period is critical for operational warning decisions (Sheets 1990). Other historical analog techniques, such as the hurricane analog (HURRAN) method of Hope and Neumann (1970) and the self-adapting ensembles of Sievers et al. (2000) and Fraedrich et al. (2003), generate forecasts by adapting entire tracks of any storm in the historical database. In contrast, our climatological technique focuses on individual motion characteristics of storms located within a specified geographical radius of influence (as defined below). Furthermore, unlike the widely used operational climatology and persistence model, known as CLIPER (Neumann 1972; Neumann et al. 1981; Leslie et al. 1990), our system generates and displays probabilistic estimates of future 24-h TC trajectories rather than the single-point forecasts produced by CLIPER. A description of the dataset used to calculate the motion climatology is given in section 2a. The computational aspects of the motion climatology are discussed in section 2b.

a. Best-track dataset

The so-called best-track Atlantic hurricane dataset, described by Jarvinen et al. (1984) and updated annually by the NOAA Tropical Prediction Center, was used to compute the motion climatology statistics used in this study. The dataset uses all available surface, satellite, and aircraft reconnaissance observations—including those not accessible in real time—to revise and refine the official poststorm estimates of TC position and intensity (Neumann and Pelissier 1981b). This dataset is a record of all TC activity in the Atlantic basin dating back to 1851. For this study, only the most recent (1970–2003) records are used in an attempt to maximize the stationarity of the dataset and minimize any discontinuities due to secular improvements in observing technology or changes in operational classification schemes (Landsea 1993; Landsea et al. 1996; Buckley et al. 2003; Barrett and Leslie 2005). The most recent 34 yr of the dataset provide three pieces of information critical to any TC climatology study: geographical location (latitude and longitude); temporal location (month, day, and year); and intensity (maximum sustained 1-min surface winds and minimum sea level pressure). These data are available four times daily (0000, 0600, 1200, and 1800 UTC) over the life of each TC.

b. The TC motion climatology prediction scheme

The motion climatology for a specified geographical point is calculated by first searching the best-track dataset to find all TC records located within a prescribed distance, or "radius of influence" of that point (Barrett et al. 2004). The speed and direction vector components are computed directly from the great-circle distance (GCD) traveled by the TC in the 24-h period, that is,

$$\mathrm{GCD} = 111(\cos^{-1})[\sin(\varphi_1)\sin(\varphi_2) + \cos(\varphi_1)\cos(\varphi_2)\cos(\lambda_2 - \lambda_1)], \qquad (1)$$

where (φ_1, λ_1) and (φ_2, λ_2) are the initial and final latitudes and longitudes of the center of the TC.

Because each motion vector contains both a speed and a directional component, it is possible to divide the vectors into convenient radial "bins" of direction (in degrees) and speed (in knots). For this study, we divided the vector space into 180 bins: 36 radial categories, each 10° in azimuth, and five translational speed categories, each 5 kt in range. Each historical TC record can then be sorted into its corresponding radial sector and speed bins. These bin totals are converted into relative frequencies and displayed graphically in a format analogous to a probabilistic "wind rose." These relative frequencies, which range from 0.00 to 1.00, represent the historical mean 24-h trajectories for that specific geographical point. The calculations and graphics are computationally negligible, requiring about 2 s on a desktop PC. With just one program command, the climatological mean TC speed and heading information is available for any point in the Atlantic basin. Furthermore, because the technique is initialized using just the TC initial position, the prediction is available at the beginning of the forecast period, as there is no need to wait several hours for a numerical analysis.

3. TC Ivan, September 2004

North Atlantic TC activity reached record levels in 2004. Fifteen TCs formed (counting Subtropical Storm Nicole), including nine hurricanes and a record seven TCs in August (Gray et al. 2004). Eight TCs, including five hurricanes, made landfall on the U.S. mainland. Five TCs, including a record four hurricanes, came ashore in Florida alone. Hurricane Ivan was the ninth TC and the sixth hurricane of the 2004 Atlantic season. It also happened to be the first hurricane for which we closely monitored and predicted the trajectory, in real time, using the climatological technique developed by the first two authors.

Ivan formed in the eastern tropical North Atlantic basin on 3 September and traveled west-northwest through the southern Windward Islands into the Caribbean Sea and Gulf of Mexico, eventually making landfall near Gulf Shores, Alabama, on 16 September (Fig. 1; Stewart 2004). From 0000 UTC 5 September through 1200 UTC 13 September, Ivan moved from the south-central tropical North Atlantic Ocean, through the southern Windward Islands, to the western tip of Jamaica, reaching category 5 (Simpson 1974) at its peak intensity. This 8-day period is of most interest to us for four reasons: 1) Ivan remained a well-organized, long-track hurricane (intense hurricane) for 30 (25) of the 35 forecast periods; 2) over 500 forecasts were generated by 14 different operational prediction methods; 3) Ivan's westward motion was repeatedly, and consis-

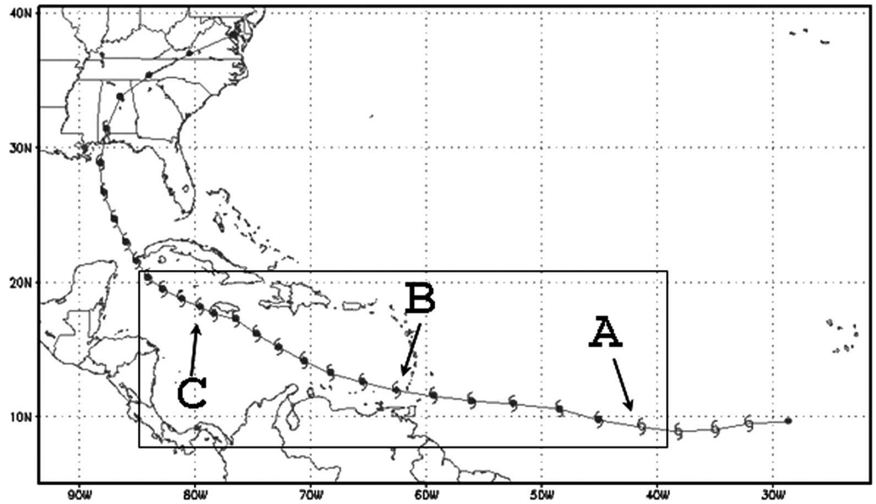

FIG. 1. Track of Hurricane Ivan (2004). Boxed region highlights the 8-day period from 0000 UTC 5 Sep to 1200 UTC 13 Sep when a strong climatological signal repeatedly conflicted with NWP forecasts. Letters A, B, and C indicate the locations of the motion climatologies depicted in Fig. 2.

tently, underforecast by almost every operational NWP model; and 4) the climatological signal in this part of the tropical Atlantic basin clearly indicated a preference for a continuing westward motion. Section 3a presents the climatological signal associated with Ivan, obtained by applying the motion climatology scheme described in section 2. Section 3b contains a summary of NWP models examined for this study and discusses the trajectory errors associated with the prediction methods. Section 3c details an analysis of the deep-layer mean synoptic steering flow, computed from the archived analyses using code developed by the third author, and referred to here as FLOW.

a. *Motion climatology interpretation*

Hurricane Ivan was chosen from the 2004 TC season as a case study that clearly illustrated the potential errors in NWP model track forecasts when they differ repeatedly from the strong climatological signal calculated using the climatological scheme developed for this study. In Fig. 2, the length of each radial sector corresponds to the probabilistic 24-h trajectory preferences of all TCs located within a radius of influence of the chosen location. The longer radial sectors denote preferred TC trajectories. In Figs. 2a and 2b, it is easily seen that approximately 80% of TCs comprising the climatology (the numbers of cases are 102 and 114, respectively) have directional headings in a small range between 270° and 295°. For comparison, the observed motion vector for Hurricane Ivan is superimposed onto the historical motion climatology in Figs. 2a–c. We note the remarkable agreement between Ivan and climatology in Figs. 2a and 2b, where the climatological signal is strong and has a narrow spread.

This type of climatological product has several key features. First, as discussed already, it quickly and simply displays the climate information relevant to each TC track. Second, it gives an indication of the variability of the synoptic steering flow. A strong, unimodal preference for westerly directional headings with average speeds of 11 to 20 kt is apparent in Figs. 2a and 2b. However, a more evenly distributed synoptic signal with westerly through northeasterly directional headings is present in Fig. 2c. Third, unlike many statistical methods such as CLP5 and A98E, this product does not specify a point forecast, but instead displays the spread of past TC trajectories.

As a consequence, the climatological scheme adds considerable value to a real-time forecasting setting. Because "tropical cyclone tracks tend to be repetitive and are associated with likewise repetitive synoptic patterns" (Bessafi et al. 2002), these climatological relative frequencies convey highly valuable probabilistic information to forecasters, especially so in the deep Tropics where synoptic steering patterns tend to be more repetitive than in the subtropics and in basins where TC tracks are not as erratic (Pike and Neumann 1987).

b. *NWP models*

To assess Ivan's predictability, the performances of 14 operational prediction methods initialized between 0000 UTC 4 September through 1200 UTC 12 Septem-

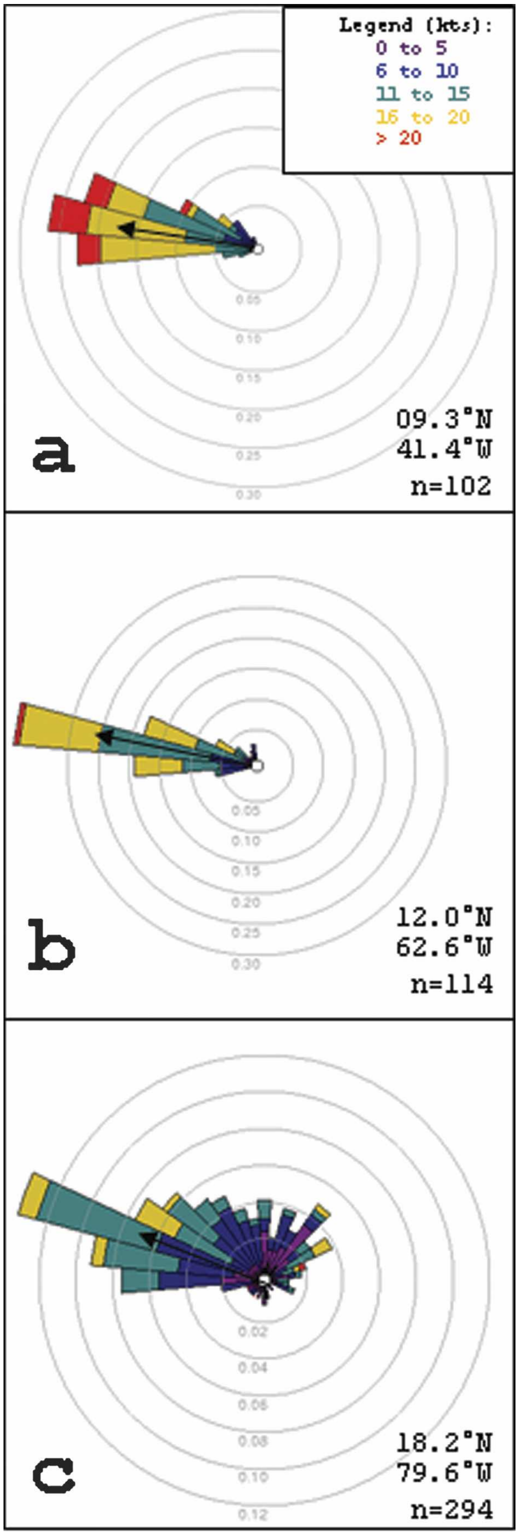

FIG. 2. (a)–(c) Historical motion climatology for three locations along the track of Hurricane Ivan. Length of each sector corresponds to relative frequency of a TC moving with that trajectory; concentric circles are labeled with relative frequency increasing away from the center. Colors represent mean 24-h speeds; a legend is provided in the upper-right corner of (a). In the lower-right corner of each panel the geographical center of each climatology and the number of historical TCs comprising each climatology (e.g., $n = 102$) are given. The radius of influence used to determine each climatology was 500 km. Dark arrows represent Ivan's actual motion vector.

ber were evaluated: two statistical–climatological schemes, six NWP models, four limited-area barotropic models, and three ensembles of NWP models (see Table 1 for a summary of each prediction method). In addition, the TPC official (OFCL) forecast was included. All the TC track forecasts were provided by the Hurricane Research Division (HRD) in Miami, Florida, and the Naval Research Laboratory (NRL) in Monterey, California.

The 24-h forecast positions for each of the 12 nonstatistical prediction methods were found to be consistently to the right (poleward) of Ivan's actual track (as indicated by positive trajectory errors in Table 2). The largest trajectory errors were associated with the dynamical models, namely, the GFDL, U.K. Met Office (UKMET), National Centers for Environmental Prediction (NCEP) Global Forecast System (AVNO), and U.S. Navy Operational Global Atmospheric Prediction System (NOGAPS), and their ensemble forms (CONU, GUNS, and GUNA). During the 8-day period in early September, 35 forecasts were generated by each of the GFDL, AVNO, and NOGAPS models, while 16 forecasts were made by the UKMET model (see Figs. 3a and 3b for a representative sample of NWP track forecasts). The 24-h mean trajectory errors from these models ranged from +4.0° to +6.3° (positive values indicate poleward track biases). Each trajectory error was tested for statistical significance using a two-tailed Student's *t* test. Three null hypotheses, comparing the NWP forecasts to zero (no error), CLP5, and OFCL, were considered. A summary of the statistical *p* values is presented in Table 3. The null hypothesis required *p* values less than 0.005 to be rejected at the 99% confidence level. It was found that all of the dynamical model forecasts (GFDL, UKMET, AVNO, and NOGAPS) and their ensembles (CONU, GUNS, and GUNA) had a statistically significant right-of-track bias (at the 99% confidence level). Furthermore, the OFCL forecast was also found to have a statistically significant right-of-track bias (also at the 99% confidence level).

In marked contrast, the two statistical–climatological methods examined in this study, A98E and CLP5, had much smaller mean trajectory errors than the dynamical models (Table 3), and the errors were *not* significant at the 99% confidence level. The error verifications

TABLE 1. Tropical cyclone prediction methods used in this study.

Name	Acronym(s)	Reference(s)	Organization(s)	Comments
Statistical and dynamical hurricane track model	A98E	Neumann and McAdie (1991)	NOAA TPC	
NCEP Global Forecast System	AVNO	Kanamitsu (1989); Lord (1993)	NCEP	Aviation run
Beta and advection models	BAMS BAMM BAMD	Marks (1992); Holland (1983)	NOAA TPC	S—Shallow layer M—Medium layer D—Deep layer
Climatology and persistence model	CLP5	Neumann (1972)	NOAA TPC	
Canadian Meteorological Centre model	CMC	Côté et al. (1998)	CMC	
Geophysical Fluid Dynamics Laboratory Hurricane Forecast System	GFDL	Kurihara et al. (1993, 1995, 1998); Bender et al. (1993)	NOAA GFDL	
Consensus forecast models	CONU GUNS GUNA	Goerss (2000)	NCEP TPC	CONU: A consensus of at least two of GFDL, GFDN (U.S. Navy run of GFDL), GFS, NOGAPS, and UKMET. GUNS: A consensus of GFDL, UKMET, and NOGAPS. GUNA: A consensus of GFDL, UKMET, NOGAPS, and GFS.
Limited-area sine transform barotropic model	LBAR	Chen et al. (1997); Horsfall et al. (1997)	NOAA TPC	
U.S. Navy Operational Global Atmospheric Prediction System	NOGAPS	Hogan and Rosmond (1991); Goerss and Jeffries (1994)	Fleet Numerical Meteorological and Oceanographic Center (FNMOC)	
U.K. Met Office Model	UKMET	Cullen (1993); Heming et al. (1995)	UKMET	
TPC official forecast	OFCL		NOAA TPC	

revealed that these methods captured Ivan's preference for continued westward motion far better than the global and regional models. Moreover, the OFCL forecasts were found to be statistically different from both A98E and CLP5 (at a 99% confidence level), but not from AVNO, GFDL, NOGAPS, or UKMET models, or the three consensus models CONU, GUNA, and GUNS. This finding agrees well with Stewart (2004), who suggested that the OFCL forecasts relied heavily on the dynamical model forecasts rather than the climatological models.

c. *Steering flow for Ivan*

Stewart (2004) concluded that the right-of-track bias in the NWP models can be attributed largely to the models' premature erosion of the strong subtropical ridge in the mid-Atlantic. To examine the synoptic currents in the vicinity of Ivan, we calculated a deep-layer mean steering flow from NCEP reanalysis-2 dataset us-

TABLE 2. Hurricane Ivan trajectory errors for the thirty-five 24-h forecasts generated between 0000 UTC 4 Sep and 1200 UTC 12 Sep. Mean left-of-track heading errors are noted in bold. The acronyms are defined in Table 1.

Acronym	Heading error (°)	Position error (km)
A98E	**−0.4**	74
AVNO	6.2	115
BAMD	3.7	107
BAMM	1.5	83
BAMS	3.2	139
CLP5	**−0.8**	87
CMC	3.0	108
CONU	5.9	82
GFDL	4.0	86
GUNA	6.3	91
GUNS	5.8	77
LBAR	3.2	97
NOGAPS	4.3	75
UKMET	4.6	78
NHC OFCL	5.8	85

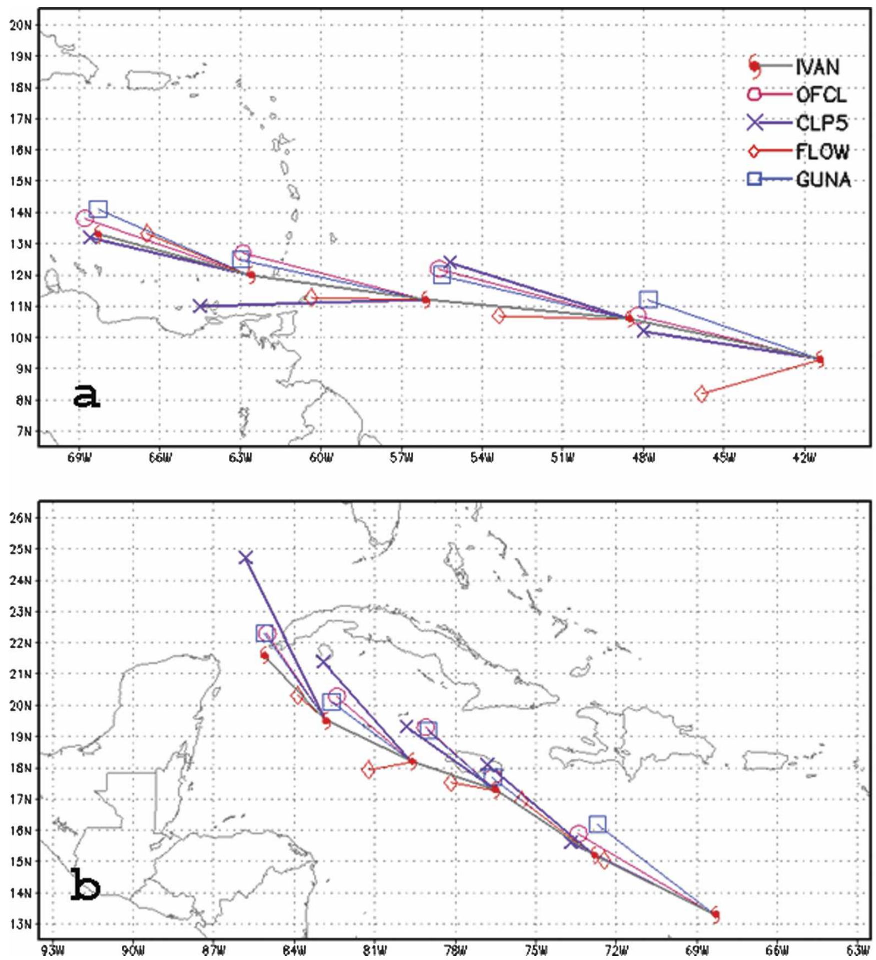

FIG. 3. Ivan track and model spread, and 24-h track forecasts (initialized at 0000 UTC) for OFCL (open circle), CLP5 (cross), FLOW (open diamond), and GUNA (open square) models. (a) The period from 5 to 9 September. (b) The period from 10 to 14 September. In (a), steering flow is primarily from E to W; in (b), steering flow becomes more S to N.

ing a 7 × 7 box averaged over 850–200 mb. The deep-layer mean was used, as it has been suggested that the deep-layer mean is the most appropriate choice for the strongest storms (see, e.g., Velden and Leslie 1991). This computation applied the trapezoidal rule to the 6-hourly values at 200, 250, 300, 400, 500, 600, 700, and 850 hPa. The resulting track forecast, which we refer to hereafter as "FLOW," is indicated in Figs. 3a and 3b by an open diamond. When Ivan was south of Hispaniola, the steering flow–based trajectory was more accurate than most NWP and consensus models. Unlike these models, FLOW did not exhibit a statistically significant right-of-track bias. However, its mean square trajectory errors were comparable to the NWP models, and it can be seen in Figs. 3a and 3b that the FLOW trajectory forecast was often left-of-track. We believe this equatorward pattern of errors in the FLOW trajectory forecast reveals the strength of the synoptic-scale ridge centered north of Ivan. Thus, the official forecast for Ivan, which relied heavily upon the NWP models instead of the statistical and climatological models, had significant right-of-track errors.

In addition to the above treatment of Ivan, we examined the steering flow for TC Lili, from the 2002 Atlantic season, in the same manner. Lili formed and tracked over a similar path to Ivan and also reached hurricane intensity. However, the NWP models did not exhibit the same right-of-track bias as for Ivan. With Lili, the GUNA dynamical ensemble and the official forecast both had smaller 24-h position errors (49 and 54 km, respectively) than the climatology model CLP5 (87 km; Lawrence 2002; Pasch et al. 2004). These findings, which are in contrast with those for Ivan, provide further support for our advocating a return to weighting

TABLE 3. The p values (for $\alpha = 0.01$, two-tailed t test) comparing model trajectory errors to zero, CLP5, or OFCL. Here, p values less than 0.005 imply that the model errors are significantly different from zero, CLP5, or OFCL at the 99% confidence level, and p values in boldface represent cases where there is no statistically significant difference between the models' heading errors and no (zero) error, CLP5, or OFCL. The acronyms are defined in Table 1.

Acronym	No (zero) error	CLP5	OFCL
A98E	**0.718**	**0.726**	0.000
AVNO	0.000	0.000	**0.630**
BAMD	0.000	0.000	0.022
BAMM	**0.072**	**0.007**	0.000
BAMS	0.014	0.002	0.037
CLP5	**0.441**	1.000	0.000
CMC	0.000	0.000	0.001
CONU	0.000	0.000	**0.814**
GFDL	0.000	0.000	0.020
GUNA	0.000	0.000	**0.326**
GUNS	0.000	0.000	**0.966**
LBAR	0.001	0.000	**0.005**
UKMET	0.000	0.000	**0.191**
NOGAPS	0.000	0.000	**0.074**
NHC OFCL	0.000	0.000	1.000

more heavily the predictions available from statistical and climatological methods, particularly when the climatological scheme consistently has a strong track prediction signal that differs from the NWP models and the number of cases making up the climatology is large.

4. Conclusions

A consequence of the rapid advance of dynamical models has been a move away from the operational use of TC prediction schemes based on climatology and statistical methods. In this study, we showed that neglecting these methods is a strategy that is easily remedied. We devised a simple climatological scheme that provides graphical displays of climatological TC motion data in a quick and timely manner. When the climatological signal from the scheme is strong and has a small spread, deviations from the climatologically derived synoptic direction predictions, while still possible, are expected to be minimal. Conversely, when the signal is weak and has a large spread, the climatological scheme is not expected to be of much value, other than to suggest that additional care should be taken to examine the various components that comprise the resultant steering of the TC.

This case study examined here was TC Ivan, which reached hurricane intensity during the 2004 Atlantic season and caused significant loss of life and property in the southeast United States. Our focus here is on an earlier period when, as a consequence of the sustained poleward track errors from the NWP models and the official forecast, evacuation orders were issued for the Florida Keys at 1200 UTC 9 September. However, Ivan passed more than 450 km to the west of Key West, in the open Gulf of Mexico waters. The evacuation was initiated because 12 different NWP models consistently, and incorrectly, predicted a poleward motion component that was not observed as the TC traversed the tropical North Atlantic as far as western Cuba. This poleward bias was shown to be statistically significant for all of the dynamical models and for the official forecast at the 99% confidence level. These forecast errors contrast with the contradictory strong climatological signal that correctly indicated a more westward motion. The forecast errors in the NWP models have since been attributed to the premature erosion of the mid-Atlantic subtropical ridge by the NWP models. The official forecast exhibited the same right-of-track bias due to its very heavy reliance on the (inaccurate in this case) NWP model predictions. Not all the operational forecast systems had poleward biases, however. The climatology and persistence model, the statistical–dynamical model, and the deep-layer mean steering flow forecast did not exhibit significant poleward biases and were found to have no directional bias at the 99% confidence level.

Hurricane Ivan is an example that shows how NWP and official forecasts can have large position errors when they are significantly different from the tracks produced by the climatological scheme used in this study. The best-track historical record contains many other recent TCs in which the statistical–climatological methods outperform the NWP models for at least some part of the forecast period. Table 4 summarizes the error statistics for these TCs from the 2004 Atlantic season. After our original submission of the manuscript, we examined Hurricane Emily of the 2005 Atlantic season, which we found to be remarkably similar to Ivan in several aspects. In mid-July, Emily tracked across the eastern North Atlantic and into the southeastern Caribbean Sea. While Emily was east of the Windward Islands, from 10 to 13 July, the suite of operational NWP track guidance models consistently predicted a northward motion component that did not develop. As with Ivan, the climatological signal for Emily indicated a strong preference for westward motion component, and it is noteworthy that again the TC followed the strong climatological signal and did not develop the northward motion component forecasted by the NWP guidance models.

In summary, TC Ivan has demonstrated that a greater role should be accorded to the statistical–

TABLE 4. Selected examples from the 2004 tropical Atlantic season where mean 24-h statistical–climatological model position errors are less than average 24-h NWP model position errors over the life of the TC. Acronyms are defined in Table 1.

Storm name	Statistical–climatological model	Statistical–climatological model position error (km)	NWP model	NWP model position error (km)
Bonnie	A98E	107	AVNO	170
			GUNA	181
			LBAR	135
Danielle	A98E	165	AVNO	212
	CLP5	192	UKMET	191
Ivan	A98E	80	AVNO	104
	CLP5	94	GFDL	96
			NOGAPS	96
			GUNA	83
Jeanne	A98E	107	GFDN	115
			NOGAPS	119
			UKMET	117

climatological methods when a strong climatological signal conflicts with the NWP or other deterministic predictions. The simple tool used here provides a means of identifying TCs that are potentially difficult to forecast by the NWP models. If the computed climatological signal is persistent, has a small spread, and is supported by a large number of archived cases, then our study demonstrates that the operational statistical–climatological schemes are potentially at least as accurate as the dynamical methods.

Acknowledgments. This work was supported by the U.S. Office of Naval Research under Research Grant N00014-00-1-0288. We thank Chris Landsea for providing the best-track dataset, and we thank Sim Aberson and Chi-Sann Liou for providing access to the NWP track forecasts. We also extend our gratitude to Dr. Peter J. Lamb for his many constructive comments and suggestions, which greatly improved the structure and content of the manuscript. We finally thank the anonymous reviewers for their valuable comments on the manuscript.

REFERENCES

Aberson, S. D., 2001: The ensemble of tropical cyclone track forecasting models in the North Atlantic basin (1976–2000). *Bull. Amer. Meteor. Soc.,* **82,** 1895–1904.

Barrett, B. S., and L. M. Leslie, 2005: An examination of the quality of the Atlantic tropical cyclone database. Preprints, *16th Conf. on Climate Variability and Change,* San Diego, CA, Amer. Meteor. Soc., CD-ROM, P1.30.

——, ——, and C.-S. Liou, 2004: Relationship between climatology and model track, bearing, and speed errors. Preprints, *20th Conf. on Weather Analysis and Forecasting,* Seattle, WA, Amer. Meteor. Soc., CD-ROM, 13.6.

Bender, M. A., and I. Ginis, 2000: Real-case simulations of hurricane–ocean interaction using a high-resolution coupled model: Effects on hurricane intensity. *Mon. Wea. Rev.,* **128,** 917–946.

——, R. J. Ross, R. E. Tuleya, and Y. Kurihara, 1993: Improvements in tropical cyclone track forecasts using the GFDL initialization system. *Mon. Wea. Rev.,* **121,** 2046–2061.

Bessafi, M., A. Lasserre-Bigorry, C. J. Neumann, F. Pignolet-Tardan, D. Payet, and M. Lee-Ching-Ken, 2002: Statistical prediction of tropical cyclone motion: An analog–CLIPER approach. *Wea. Forecasting,* **17,** 821–831.

Buckley, B. W., L. M. Leslie, and M. S. Speer, 2003: The impact of observational technology on climate database quality: Tropical cyclones in the Tasman Sea. *J. Climate,* **16,** 2640–2645.

Caplan, P., J. Derber, W. Gemmill, S.-Y. Hong, H.-L. Pan, and D. Parrish, 1997: Changes to the 1995 NCEP operational medium-range forecast model analysis–forecast system. *Wea. Forecasting,* **12,** 581–594.

Chen, Q.-S., L.-E. Bai, and D. H. Bromwich, 1997: A harmonic-Fourier spectral limited-area model with an external wind lateral boundary condition. *Mon. Wea. Rev.,* **125,** 143–167.

Côté, J., S. Gravel, A. Méthot, A. Patoine, M. Roch, and A. N. Staniforth, 1998: The operational CMC–MRB Global Environmental Multiscale (GEM) model. Part I: Design consideration and formulation. *Mon. Wea. Rev.,* **126,** 1373–1395.

Cullen, M. J. P., 1993: The unified forecast/climate model. *Meteor. Mag.,* **122,** 81–122.

Fraedrich, K., C. C. Raible, and F. Sielmann, 2003: Analog ensemble forecasts of tropical cyclone tracks in the Australian region. *Wea. Forecasting,* **18,** 3–11.

Franklin, J. L., C. J. McAdie, and M. B. Lawrence, 2003: Trends in track forecasting for tropical cyclones threatening the United States, 1970–2001. *Bull. Amer. Meteor. Soc.,* **84,** 1197–1203.

Goerss, J. S., 2000: Tropical cyclone track forecasts using an ensemble of dynamical models. *Mon. Wea. Rev.,* **128,** 1187–1193.

——, and R. Jeffries, 1994: Assimilation of synthetic tropical cyclone observations into the Navy Operational Global Atmospheric Prediction System. *Wea. Forecasting,* **9,** 557–576.

——, C. R. Sampson, and J. M. Gross, 2004: A history of western North Pacific tropical cyclone track forecast skill. *Wea. Forecasting,* **19,** 633–638.

Gray, W. M., P. J. Klotzbach, and W. Thorston, 2004: Summary of 2004 Atlantic tropical cyclone activity and verification of author's seasonal and monthly forecasts. Dept. of Atmospheric Science Rep., Colorado State University, Fort Collins, CO, 38 pp.

Heming, J. T., J. C. L. Chan, and A. M. Radford, 1995: A new scheme for the initialization of tropical cyclones in the UK Meteorological Office global model. *Meteor. Appl.,* **2,** 171–184.

Hogan, T. F., and T. E. Rosmond, 1991: The description of the Navy Operational Global Atmospheric Prediction System's spectral forecast model. *Mon. Wea. Rev.,* **119,** 1786–1815.

Holland, G. J., 1983: Tropical cyclone motion: Environmental interaction plus a beta effect. *J. Atmos. Sci.,* **40,** 328–342.

Hope, J. R., and C. J. Neumann, 1970: An operational technique for relating the movement of existing tropical cyclones to past tracks. *Mon. Wea. Rev.,* **98,** 925–933.

Horsfall, F., M. DeMaria, and J. M. Gross, 1997: Optimal use of large-scale boundary and initial fields for limited-area hurricane forecast models. Preprints, *22d Conf. on Hurricanes and Tropical Meteorology,* Fort Collins, CO, Amer. Meteor. Soc., 571–572.

Jarvinen, B. R., C. J. Neumann, and M. A. S. Davis, 1984: A tropical cyclone data tape for the North Atlantic basin, 1886–1983: Contents, limitations, and uses. NOAA Tech. Memo. NWS NHC-22, 28 pp.

Kalnay, E., M. Kanamitsu, and W. E. Baker, 1990: Global numerical weather prediction at the National Meteorological Center. *Bull. Amer. Meteor. Soc.,* **71,** 1410–1428.

Kanamitsu, M., 1989: Description of the NMC global data assimilation and forecast system. *Wea. Forecasting,* **4,** 335–342.

Krishnamurti, T. N., C. M. Kishtawal, T. LaRow, D. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, and S. Surendran, 1999: Improved skills for weather and seasonal climate forecasts from multimodel superensemble. *Science,* **285,** 1548–1550.

——, ——, ——, ——, ——, ——, ——, and ——, 2000a: Improving tropical precipitation forecasts from a multianalysis superensemble. *J. Climate,* **13,** 4217–4227.

——, ——, D. W. Shin, and C. E. Williford, 2000b: Multimodel superensemble forecasts for weather and seasonal climate. *J. Climate,* **13,** 4196–4216.

——, and Coauthors, 2001: Real-time multianalysis–multimodel superensemble forecasts of precipitation using TRMM and SSM/I products. *Mon. Wea. Rev.,* **129,** 2861–2883.

Kurihara, Y., M. A. Bender, and R. T. Ross, 1993: An initialization scheme of hurricane models by vortex specification. *Mon. Wea. Rev.,* **121,** 2030–2045.

——, ——, R. E. Tuleya, and R. J. Ross, 1995: Improvements in the GFDL hurricane prediction system. *Mon. Wea. Rev.,* **123,** 2791–2801.

——, 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.

Landsea, C. W., 1993: A climatology of intense (or major) Atlantic hurricanes. *Mon. Wea. Rev.,* **121,** 1703–1713.

——, N. Nicholls, W. M. Gray, and L. A. Avila, 1996: Downward trends in the frequency of intense Atlantic hurricanes during the past five decades. *Geophys. Res. Lett.,* **23,** 1697–1700.

Lawrence, M. B., 2002: Tropical cyclone report: Hurricane Lili 21 September-04 October. National Hurricane Center, Miami, FL, 14 pp. [Available online at http://www.nhc.noaa.gov/2002lili.shtml.]

Leslie, L. M., G. J. Holland, M. Glover, and F. J. Woodcock, 1990: A climatological-persistence (CLIPER) scheme for predicting Australian region tropical cyclone tracks. *Aust. Meteor. Mag.,* **38,** 87–92.

Lord, S. J., 1993: Recent developments in tropical cyclone track forecasting with the NMC global analysis and forecast system. Preprints, *20th Conf. on Hurricanes and Tropical Meteorology,* San Antonio, TX, Amer. Meteor. Soc., 290–291.

Marks, D. G., 1992: The beta and advection model for hurricane track forecasting. NOAA Tech. Memo. NWS NMC-70, 89 pp.

McAdie, C. J., and M. B. Lawrence, 2000: Improvements to tropical cyclone track forecasting in the Atlantic basin, 1970–1998. *Bull. Amer. Meteor. Soc.,* **81,** 989–999.

Neumann, C. J., 1972: An alternate to the HURRAN tropical cyclone forecast system. NOAA Tech. Memo. NWS SR-62, 22 pp.

——, and J. R. Hope, 1972: Performance analysis of the HURRAN tropical cyclone forecast system. *Mon. Wea. Rev.,* **100,** 245–255.

——, and J. M. Pelissier, 1981a: An analysis of Atlantic tropical cyclone forecast errors, 1970–1979. *Mon. Wea. Rev.,* **109,** 1248–1266.

——, and ——, 1981b: Models for the prediction of tropical cyclone motion over the North Atlantic: An operational evaluation. *Mon. Wea. Rev.,* **109,** 522–538.

——, and C. J. McAdie, 1991: A revised National Hurricane Center NHC83 model (NHC90). NOAA Tech. Memo. NWS NHC-44, 35 pp.

——, B. R. Jarvinen, and A. G. Pike, 1981: Tropical cyclones of the North Atlantic Ocean, 1871–1981. Historical Climatology Series 6-2, National Climatic Data Center, Asheville, NC, 174 pp.

Pasch, R. J., M. B. Lawrence, L. A. Avila, J. L. Beven, J. L. Franklin, and S. R. Stewart, 2004: Atlantic hurricane season of 2002. *Mon. Wea. Rev.,* **132,** 1829–1859.

Pike, A. C., and C. J. Neumann, 1987: The variation of track forecast difficulty among tropical cyclone basins. *Wea. Forecasting,* **2,** 237–241.

Powell, M. D., and S. D. Aberson, 2001: Accuracy of United States tropical cyclone landfall forecasts in the Atlantic basin (1976–2000). *Bull. Amer. Meteor. Soc.,* **82,** 2749–2768.

Sheets, R. H., 1990: The National Hurricane Center: Past, present and future. *Wea. Forecasting,* **5,** 185–232.

Shuman, F. G., 1989: History of numerical weather prediction at the National Meteorological Center. *Wea. Forecasting,* **4,** 286–296.

Sievers, O., K. Fraedrich, and C. C. Raible, 2000: Self-adapting analog ensemble predictions of tropical cyclone tracks. *Wea. Forecasting,* **15,** 623–629.

Simpson, R. H., 1974: The hurricane disaster potential scale. *Weatherwise,* **27,** 169, 186.

Stewart, S. R., 2004: Tropical cyclone report: Hurricane Ivan 2–26 September 2004. National Hurricane Center, Miami, FL, 44 pp. [Available online at http://www.nhc.noaa.gov/2004ivan.shtml.]

Velden, C. S., and L. M. Leslie, 1991: The basic relationship between tropical cyclone intensity and the depth of the environmental steering layer in the Australian region. *Wea. Forecasting,* **6,** 244–253.

Weber, H. C., 2003: Hurricane track prediction using a statistical ensemble of numerical models. *Mon. Wea. Rev.,* **131,** 749–770.