An Examination of Model Track Forecast Errors for Hurricane Ike (2008) in the Gulf of Mexico

MICHAEL J. BRENNAN

NOAA/NWS/NCEP/National Hurricane Center, Miami, Florida

SHARANYA J. MAJUMDAR

Division of Meteorology and Physical Oceanography, Rosenstiel School of Marine and Atmospheric Science,
University of Miami, Miami, Florida

ABSTRACT

Sources of dynamical model track error for Hurricane Ike (2008) in the Gulf of Mexico are examined. Deterministic and ensemble model output are compared against National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) analyses to identify potential critical features associated with the motion of Ike and its eventual landfall along the upper Texas coast. Several potential critical features were identified, including the subtropical ridge north of Ike and several synoptic-scale short-wave troughs and ridges over central and western North America, and Tropical Storm Lowell in the eastern North Pacific. Using the NCEP Gridpoint Statistical Interpolation (GSI) data assimilation scheme, the operational GSI analysis from the 0000 UTC 9 September 2008 cycle was modified by perturbing each of these features individually, and then integrating the GFS model using the perturbed initial state. The track of Ike from each of the perturbed runs was compared to the operational GFS and it was found that the greatest improvements to the track forecast were associated with weakening the subtropical ridge north of Ike and strengthening a midlevel short-wave trough over California. A GFS run beginning with an analysis where both of these features were perturbed produced a greater track improvement than either did individually. The results suggest that multiple sources of error exist in the initial states of the operational models, and that the correction of these errors in conjunction with reliable ensemble forecasts would lead to improved forecasts of tropical cyclone tracks and their accompanying uncertainty.

1. Introduction

Track forecasts of tropical cyclones are critically important in the days prior to a potential landfall. During this time, emergency preparations are made, evacuations are ordered, and relief supplies are positioned, even well beyond the time period at which the National Hurricane Center (NHC) issues hurricane watches and warnings (48 and 36 h prior to the arrival of tropical storm force winds, respectively). Over the past 30 years, NHC's average track forecast errors in the Atlantic basin have decreased markedly (e.g., Rappaport et al. 2009). For example, the average Atlantic basin 48-h track error from 2000 to 2008 was around 100 n mi, as compared to an average of around 250 n mi in the 1970s, a decrease of about 60% (see their Fig. 3). Despite this increase in forecast accuracy, evacuation lead times in many coastal communities have continued to increase as coastal development continues at a rapid pace. As a result, evacuation and other decisions by emergency officials are often being made 4–5 days prior to a potential landfall (M. Green 2010, personal communication).

The NHC nominally issues track forecasts every 6 h for all active tropical cyclones in its area of responsibility. These forecasts are made at projection times of 12, 24, 36, 48, 72, 96, and 120 h. In the case of Hurricane Ike (2008), track forecast errors from numerical model guidance and the NHC's official forecast increased substantially from late on 8 September through early on 9 September 2008 when compared to errors from forecasts made during the previous couple of days. For example, the average error of the four NHC 72-h track forecasts issued on 9 September

Corresponding author address: Dr. Michael J. Brennan, 11691 SW 17th St., Miami, FL 33165.
E-mail: michael.j.brennan@noaa.gov

was 92.8 n mi, a 38% increase compared to the average error of the four NHC 96-h track forecasts issued on the previous day (66.8 n mi). This period of increased track error occurred in the 3–4 days prior to Ike's final landfall in the Galveston, Texas, area early on 13 September (Berg 2011). During this time, much of the track model guidance and the official NHC forecasts exhibited a southward bias, taking the center of Ike to a landfall location along the central or southern Texas coast, instead of the upper Texas coast near Houston–Galveston where landfall ultimately occurred (Fig. 1). Of note during this period is the superior forecast, particularly at 4 and 5 days, of the National Centers for Environmental Prediction (NCEP) Global Ensemble Forecast System (GEFS) mean compared to both the NCEP Global Forecast System (GFS) deterministic model and the multimodel variable consensus model (TVCN). The TVCN is an average of at least two of the interpolated versions of the following models: the GFS, the Met Office Global Model, the Navy Operational Global Atmospheric Prediction System (NOGAPS), the Geophysical Fluid Dynamics Laboratory (GFDL) Hurricane Model, the navy version of the GFDL model (GFDN), the Hurricane Weather and Research Forecasting Model (HWRF), and the European Centre for Medium-Range Weather Forecasts (ECMWF) Global Model (the interpolation process is described in section 2b). (Fig. 1; Table 1). On average, consensus techniques, which are ensembles of track forecasts from different model systems, are superior to the forecast from the mean of a single model ensemble system (Rappaport et al. 2009; Franklin 2011).

The first goal of this study is to investigate the critical synoptic-scale features that determined the observed track of Ike in the Gulf of Mexico. Next, potential sources of error in the deterministic GFS model forecast and other deterministic models are identified. The performance of the GEFS ensemble system and particularly the variability of its members with their handling of the relevant synoptic features are then examined, with a brief examination of the ECMWF ensemble system also included. Since the perturbations made to the GEFS at the time of the event were restricted to the initial conditions, the variability in the ensemble can be attributed to initial condition uncertainty. Based on synoptic analysis and the evolution of model forecasts, potential areas are identified in which the track forecast of Ike is "sensitive" to perturbations in the GFS initial conditions. Finally, in an attempt to quantify the sensitivity, balanced perturbations to the GFS analysis are created via the assimilation of synthetic temperature observations in selected locations, and the GFS forecast model is run with the new perturbed analysis. This perturbation technique considers initial condition error only, and does not account for errors inherent in the model

physics, including those related to physical processes within the tropical cyclone (TC) itself and how they interact with the large scale, which have been shown to be important in model forecasts of TC track (e.g., Wu and Emanuel 1995a,b; Henderson et al. 1999).

In the remainder of this paper, section 2 describes the datasets, methodology, and tools used for this study. Section 3 presents a synoptic overview of the case, while section 4 contains the model and ensemble evaluations. Section 5 presents the results of the perturbation experiments, and section 6 contains the conclusions.

2. Data and methodology

a. Datasets

The synoptic overview will be performed using the final (FNL) analyses from NCEP's GFS model. These analyses were obtained on a 1° global grid and are available every 6 h. The initial analysis for the GFS model was produced using the GSI three-dimensional variational data assimilation (3DVAR) scheme (Kleist et al. 2009b). At the time of this event, the GSI and the GFS were run with horizontal spectral truncation at wavenumber 382 and 64 vertical levels (T382L64), with this resolution maintained through the first 180 h of the GFS forecast. This truncation results in an approximate horizontal grid resolution of 35 km. The forecast of the GFS will also be evaluated along with the ECMWF and Met Office (UKMET) global models. All model output was archived at NHC in real time on 1° global grids.

To examine the performance of ensemble forecast systems in this case, output from the GEFS will be analyzed. The 2008 version of the GEFS is composed of 20 members integrated 4 times per day at T126 L28 (~90 km) resolution, out to 16 days.[1] An ensemble transform method is employed to create the initial perturbations (Wei et al. 2008), with a tropical cyclone relocation procedure applied to all ensemble members that places all TCs at the analyzed position provided by NHC, the Central Pacific Hurricane Center, or the Joint Typhoon Warning Center (Liu et al. 2000). As with the deterministic model output, the GEFS fields were archived on 1° global grids. A brief examination of the output from the ECMWF ensemble system is also provided. At the time of this event, the ECMWF ensemble system had 51 perturbed members and was run with a resolution of T399L62 out

---

[1] While not included in 2008, stochastic model perturbations are now included in the operational ensemble, together with a resolution upgrade to T190 L28.
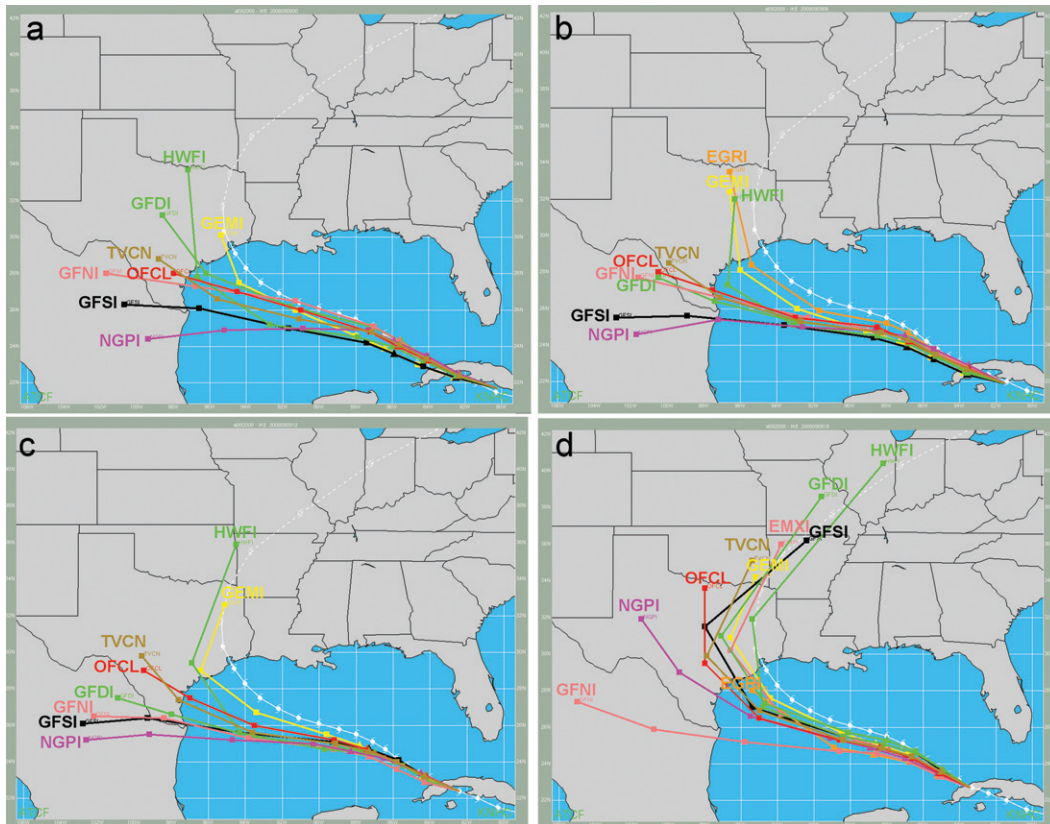
FIG. 1. Final NHC best track of Ike (white) with various track guidance models and the NHC official forecast (OFCL, red) from (a) 0000 UTC 9 Sep, (b) 0600 UTC 9 Sep, (c) 1200 UTC 9 Sep, and (d) 1800 UTC 9 Sep 2008. Models shown include GFSI (previous cycle GFS, interpolated; black), EGRI (previous cycle UKMET with subjective tracker, interpolated; orange), EMXI (previous cycle ECMWF, interpolated; pink), NGPI (previous cycle NOGAPS, interpolated; purple), GFDI (previous cycle GFDL, interpolated; green), HWFI (previous cycle HWRF, interpolated; green), GFNI (previous cycle GFDN, interpolated; pink), GEMI (previous cycle GEFS ensemble mean, interpolated; yellow), and the multimodel consensus TVCN (brown), an average of at least two of GFSI, EGRI, NGPI, GFDI, HWFI, GFNI, and EMXI.

through 10 days, with degraded resolution thereafter (ECMWF 2011).

b. Methodology

Errors of operational TC track model guidance and the NHC official forecast are computed using data from NHC's Automated Tropical Cyclone Forecast (ATCF) software system (Sampson and Schrader 2000). The ATCF maintains a database for each TC that includes a variety of track model guidance from individual, ensemble, and consensus models as well as the final "best track" of the TC. The best track consists of a subjectively smoothed analysis of TC position, intensity, and structure (i.e., wind radii) every 6 h through the life cycle of the TC (Rappaport et al. 2009). Track forecast error is defined as the great-circle distance between the forecast position and the best-track-analyzed position of the TC.

When evaluating model TC track error for a particular forecast cycle (e.g., the 1200 UTC NHC forecast cycle that culminates in the issuance of the forecast at 1500 UTC), the deterministic and ensemble track guidance being discussed will be from the previous model cycle. This allows for a more fair comparison to the official NHC official forecast since, for example, the 1200 UTC run of the GFS model is not available in time for use in the 1200 UTC NHC forecast cycle. However, the ATCF software smooths and adjusts the model track from the previous GFS run (0600 UTC) so that the 6-h forecast position of the TC (valid at 1200 UTC) matches the analyzed position of the TC at that time. This same adjustment is then made to the forecast positions at all forecast valid times. A similar process is performed for TC track guidance from other dynamical models; although for some models, such as the ECMWF and UKMET, the previous available run is sometimes 12 h old, since those models

TABLE 1. Average 72-, 96-, and 120-h forecast track errors [nautical miles (n mi), where 1 n mi = 1852 m] from the interpolated GFS ensemble mean (GEMI), the interpolated GFS (GFSI), and the multimodel TVCN consensus from the four operational forecast cycles on 9 Sep 2008.

Model	Forecast valid time (h)		
	72	96	120
GEMI	74.9	82.4	220.5
GFSI	131.6	239.9	644.7
TVCN	97.3	180.0	435.6

are only run out through the 120-h forecast periods at 0000 and 1200 UTC. For a more complete description of this process, see Goerss et al. (2004), Sampson et al. (2006), and Franklin (2011).

To quantify the sensitivity of GFS forecasts to selected local modifications to the model analyses, synthetic "observations" are assimilated into the GSI system. The GSI framework permits the assimilation of a single synthetic observation of horizontal wind components, temperature, or specific humidity at any prescribed pressure level. In this paper, synthetic observations of temperature are specified at a given location and level, with an artificially low observation error of 0.1 K and departures from the operational analysis of 4–5 K. Through the assimilation of temperature, the other meteorological fields are correspondingly adjusted in the horizontal and vertical via a tangent linear normal mode constraint on the analysis increment (Kleist et al. 2009a). Examples of analysis increments from single-observation experiments are presented in Kleist et al. (2009b). For perturbations involving more than one observation, the single-observation experiments are performed serially, with the first-guess field being the analysis from the assimilation of the previous observation. All assimilations in this paper are performed at 0000 UTC 9 September 2008.

3. Synoptic overview

At 0000 UTC 8 September 2008, Hurricane Ike was located over eastern Cuba moving westward at 11 kt (1 m s^{-1} = 1.94 kt). North of Ike, a mid- and upper-level ridge axis extended from the western Atlantic across northern Florida and into the Gulf of Mexico, with a 200-hPa anticyclone centered over the extreme western Gulf (Fig. 2a). The midlatitude flow over North America showed a long-wave ridge over western Canada and the eastern North Pacific transitioning to a zonal pattern over the eastern United States with a 100-kt upper-level jet centered over the Great Lakes.

By 0000 UTC 9 September, Ike had moved to a position just offshore of central Cuba, and was moving westward at around 12 kt. Several features were apparent

in the synoptic environment of Ike, and are labeled in Fig. 2b. The ridge (A) north of Ike had weakened and the midlevel ridge axis had shifted northward while the 200-hPa anticyclone had moved westward into northern Mexico. Farther north, a short-wave trough (B) was evident over the northern Mississippi River valley, with weak ridging (C) over the Rockies. Farther west a pair of short-wave troughs was evident, one over British Columbia (D) and a second digging south along the coast of California (E). Upstream over the eastern Pacific, a long-wave ridge had amplified south of Alaska and was centered just west of 140°W (F). At 200 hPa the flow was dominated by a 100-kt jet extending from the mid–Mississippi River valley to the upper Great Lakes, with only weak westerly flow present along the Gulf coast and in the northern Gulf. Also noted is the presence of Tropical Storm Lowell (G) southwest of the Baja California peninsula.

On 10 September at 0000 UTC, Ike had moved offshore of the western tip of Cuba and turned toward the west-northwest with a slower forward speed of 8 kt (Fig. 2c). The subtropical ridge axis north of Ike had shifted slightly to the south and again extended from north Florida across the northern Gulf of Mexico. The upper-level anticyclone associated with Ike dominated the flow pattern over the eastern Gulf and downstream over the western Atlantic. Over western North America, the long-wave mid- and upper-level trough amplified as the short-wave troughs over the Pacific Northwest and central California also intensified. Downstream, the mid- and upper-level flow over the central United States remained generally zonal as the short-wave trough previously over the Great Lakes moved into the northeastern United States.

Ike moved into the eastern Gulf of Mexico by 0000 UTC 11 September and turned toward the northwest with a forward speed of 7 kt (Fig. 2d). The flow over western North America continued to amplify with a positively tilted mid- to upper-level trough extending from Saskatchewan, Canada, southwestward to west of Baja California. Downstream ridging developed over the central United States and a 90-kt southwesterly 200-hPa jet intensified over the southwestern United States between the upper trough and the 200-hPa anticyclone over northeastern Mexico.

By 0000 UTC 12 September, Ike had turned back toward the west-northwest at around 10 kt and was located in the north-central Gulf (Fig. 2e). The flow over the United States continued to amplify, with an upper-level ridge axis along the Mississippi River valley and a mid-level anticyclone situated over the southeastern United States northeast of Ike, along with southeasterly midlevel flow over much of the northeastern Gulf (not shown). The 200-hPa anticyclone shifted north into south-central Texas with southwesterly upper-level flow between it and the trough over northern Baja California.
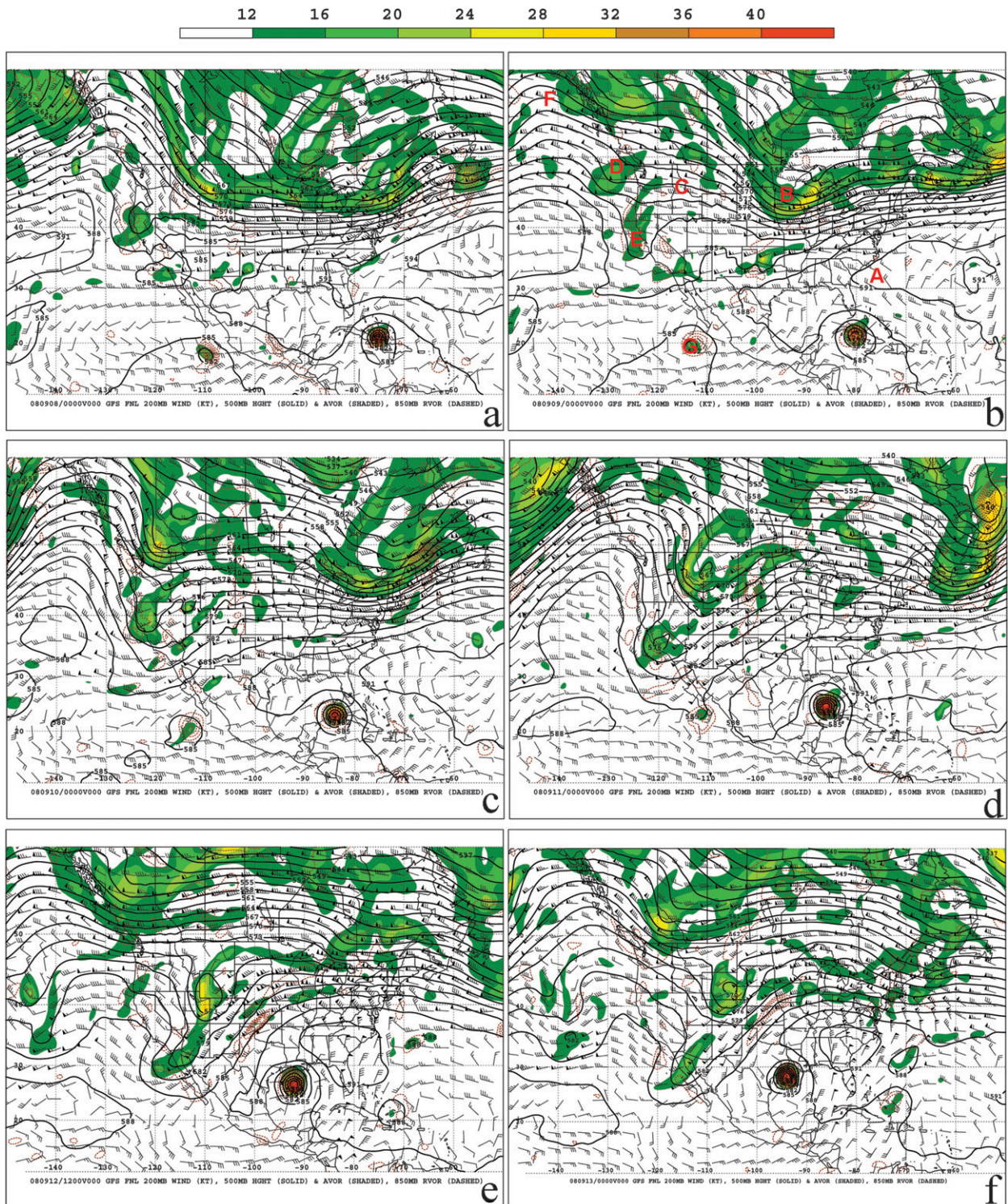
FIG. 2. GFS final (FNL) analysis showing 200-hPa wind [barbs, kt ($=0.5144$ m s^{-1}], 500-hPa geopotential height (solid contours every 3 dam) and absolute vorticity (shaded every 4×10^{-5} s^{-1} beginning at 12×10^{-5} s^{-1}), and 850-hPa relative vorticity (dashed contours, every 4×10^{-5} s^{-1} starting at 4×10^{-5} s^{-1}) valid at 0000 UTC (a) 8 Sep, (b) 9 Sep, (c) 10 Sep, (d) 11 Sep, (e) 12 Sep, and (f) 13 Sep 2008. Features discussed in the text related to the track of Ike are labeled in (b).

TABLE 2. GFS track forecast errors for Ike (n mi) for model runs initialized from 0000 UTC 8 Sep through 1200 UTC 10 Sep 2008. N/A indicates that Ike was not a tropical cyclone at the forecast's verifying time.

Model initial time	Forecast cycle time	Forecast valid time (h)				
		24	48	72	96	120
0000 UTC 8 Sep	0600 UTC 8 Sep	18	33	42	104	235
0600 UTC 8 Sep	1200 UTC 8 Sep	23	13	27	108	234
1200 UTC 8 Sep	1800 UTC 8 Sep	13	39	70	111	245
1800 UTC 8 Sep	0000 UTC 9 Sep	21	62	113	192	527
0000 UTC 9 Sep	0600 UTC 9 Sep	46	97	156	311	762
0600 UTC 9 Sep	1200 UTC 9 Sep	33	75	137	319	N/A
1200 UTC 9 Sep	1800 UTC 9 Sep	41	83	120	139	N/A
1800 UTC 9 Sep	0000 UTC 10 Sep	28	76	153	276	N/A
0000 UTC 10 Sep	0600 UTC 10 Sep	21	55	151	489	N/A
0600 UTC 10 Sep	1200 UTC 10 Sep	36	81	226	N/A	N/A
1200 UTC 10 Sep	1800 UTC 10 Sep	12	47	203	N/A	N/A

Finally, by 0000 UTC 13 September Ike accelerated toward the northwest with a forward speed of 11 kt as it approached the upper Texas coast (Fig. 2f). At this time Ike was situated between a deep-layer trough over the western United States and a midlevel ridge over the southeast, setting the stage for Ike's recurvature into the central United States ahead of the trough.

Overall, in the 5-day period leading up to Ike's landfall, the mid- and upper-level flow over the United States evolved from a relatively zonal pattern to a highly amplified pattern, with downstream ridge development over the central and eastern United States ahead of the amplifying western North American trough and likely some contribution to the amplification of the flow due to the upper-level outflow of Ike itself. How the operational models handled the details of the flow evolution over North America offers several potential sources of forecast uncertainty, including the timing and phasing of individual short-wave troughs and subtle differences in the evolution of the ridge north of Ike. In the next section, the possible sources of error in the operational GFS forecast from the 0000 UTC cycle on 9 September will be examined, since this was the GFS run that took Ike farthest south away from the observed track.

4. Operational model evaluation

a. NCEP GFS

The GFS model had a large increase in 4- and 5-day track forecast errors in the 0000 UTC 9 September run (Table 2). Therefore, the large-scale pattern forecast from this run of the GFS will be evaluated in comparison to the analyses described in the previous section to see if differences in the synoptic-scale steering features critical to the track of Ike can be identified.

The 24-h forecast valid at 0000 UTC 10 September (Fig. 3a) does not show large departures from the GSI analysis valid at that time (Fig. 2c) with any of the large-scale features. Ike is forecast to be a little farther west than analyzed; however, the midlevel ridge north of Ike in the GFS forecast is displaced farther east than analyzed, particularly when examining the 591-dam 500-hPa height contour. Farther upstream, there are subtle differences in the depth and location of the short-wave trough over northern California, with the analysis showing a sharper trough and higher vorticity extending farther south than the forecast.

By 0000 UTC 11 September (Fig. 3b), the 48-h GFS forecast track of Ike has begun to depart more significantly from the observed track, and is now located about 70 n mi (130 km) west-southwest of the observed position (Fig. 2d). The forecast position of the 500-hPa ridge axis north of Ike over the southeastern United States is displaced a little to the south compared with its analyzed location across central Georgia and Alabama. Over the western United States, the GFS forecast of the northern short-wave trough over the Rockies is too amplified and too slow, with a much stronger short-wave forecast over southern Idaho compared to the analyzed vorticity maximum near the Montana–Idaho border. Conversely, the GFS forecast is too weak with the southern stream short-wave trough over southern California. At upper levels, the forecast 200-hPa 90-kt jet over the southwestern and central United States is too strong compared with the analysis, though the upper-level flow immediately poleward of Ike is similar.

At the 72-h forecast time, valid at 0000 UTC 12 September (Fig. 3c), the GFS forecast continues to take Ike farther westward and southward relative to the observed track (Fig. 2e). Large differences are now apparent in the configuration of the subtropical ridge north of the hurricane. The forecast orientation of the midlevel ridge axis is east to west over the central and eastern Gulf coast (Figs. 3c and 4b), as compared to a more northwest to southeast orientation in the analysis (Figs. 2e and 4a). This results in a much larger southerly component in the midtropospheric wind field east of Ike than is forecast (Fig. 4). Over the western United States, the GFS forecast continues to be too strong and too slow with the northern stream short wave now located over northern Utah and is too weak with the southern stream vorticity maximum over southern California and northern Baja California.

These differences in the location and amplitude of the western U.S. trough continue to grow by the 96-h forecast valid at 0000 UTC 13 September (cf. Figs. 3d and 2f). The GFS forecast continues to be too weak with the southern short-wave trough, resulting in a weaker height gradient over northern Mexico, New Mexico, and western Texas. Consequently, the magnitude of the 200-hPa southwesterly
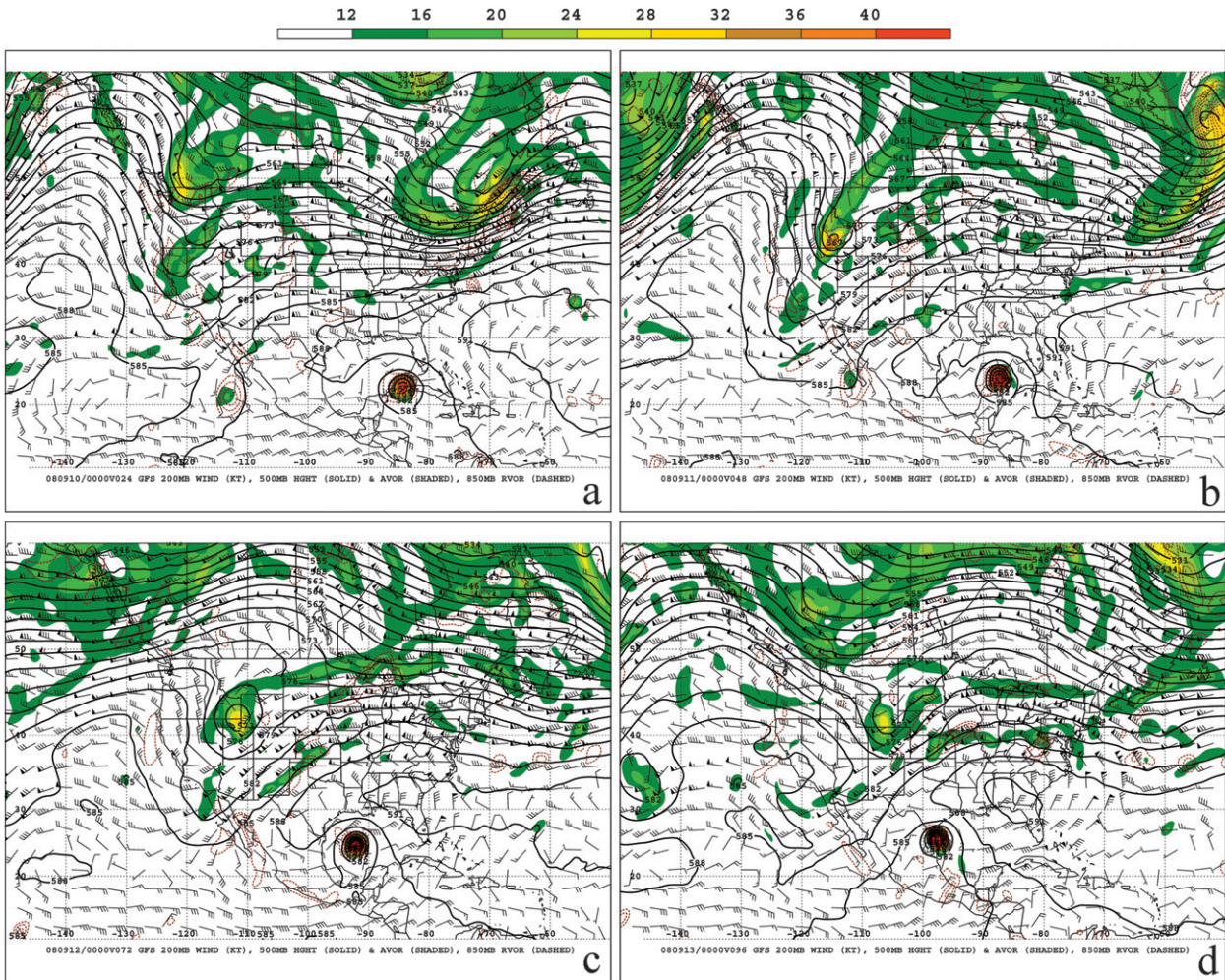
FIG. 3. As in Fig. 2, but for GFS forecast from the 0000 UTC 9 Sep 2008 cycle valid at (a) 0000 UTC 10 Sep, (b) 0000 UTC 11 Sep, (c) 0000 UTC 12 Sep, and (d) 0000 UTC 13 Sep 2008.

flow northwest of Ike is too weak in the GFS forecast. Immediately north of Ike, the forecast midlevel ridge is displaced southwest of the analyzed position, lying closer to the Gulf coast, compared to the location over the Tennessee Valley in the analysis.

The errors in the forecast of the midlevel ridge location and the southern extent of the western U.S. trough appear to be responsible for the operational GFS moving Ike due westward into Mexico and avoiding recurvature ahead of the advancing western U.S. trough into the central United States after landfall.

b. Other deterministic models

The ECMWF deterministic model forecast track of Ike from the 0000 UTC 9 September model run was also well south of the cyclone's actual landfall location (Figs. 5a and 5b), although it was slightly farther north than the track of the GFS (Figs. 3c and 3d). Differences

with the GFS FNL analysis remain rather small through 48 h (not shown); however, by 0000 UTC 12 September, differences in the midlevel ridge strength and orientation are more apparent (cf. Figs. 6a and 4a). Even though the ECMWF position of Ike is only a little to the south of the analyzed position (cf. Figs. 5a and 2e), the ECMWF shows a weaker 500-hPa ridge north of Ike and the orientation of the ridge axis is more east to west in the ECMWF forecast (Fig. 6a) compared to the GFS analysis (Fig. 4a). This weaker ridge is also noticeable in the ECMWF forecast valid at 0000 UTC 13 September (cf. Figs. 5b and 2f). Also, by 12–13 September, the ECMWF is noticeably weaker with the short-wave trough over the southwestern United States and northwestern Mexico (Figs. 5a and 5b).

The forecast evolution of the large-scale pattern in the UKMET model run (Figs. 5c and 5d) is quite similar to the ECMWF results, with a weaker midlevel ridge apparent
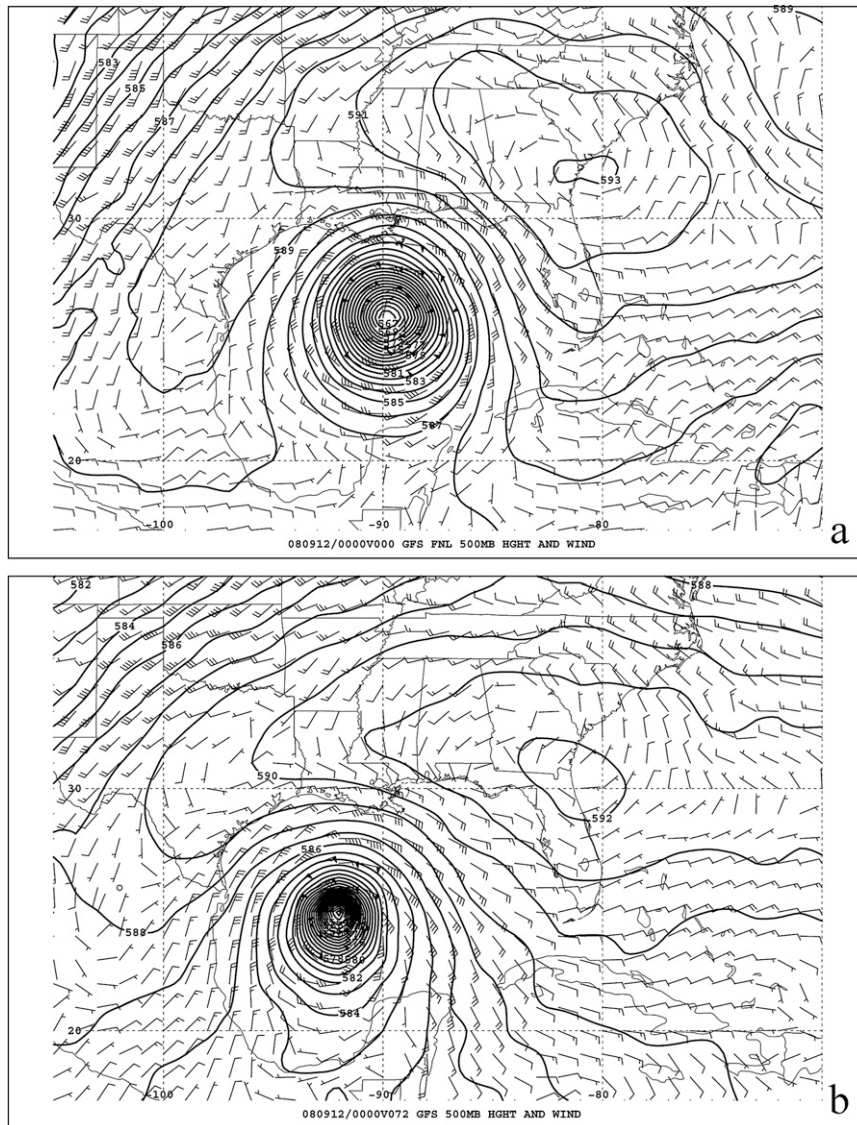
FIG. 4. Comparison of 500-hPa geopotential height (contours every 1 dam) and winds (barbs, kt) from (a) GFS analysis and (b) 72-h GFS forecast valid at 0000 UTC 12 Sep 2008.

by 12 September compared to the GFS analysis (cf. Figs. 6b and 4a). This trend continues on 13 September, and at this time the UKMET is also too weak with the southwestern U.S. short-wave trough (Fig. 5d). Interestingly, at this time the UKMET shows Ike's vortex shearing apart, with the low-level vortex turning northward toward the upper Texas coast (Fig. 5d) while the midlevel vortex weakens and continues westward toward northern Mexico. The reasons for this are unclear and beyond the scope of this study, however.

Both the ECMWF and UKMET forecasts initialized at 0000 UTC 9 September had similar issues with the forecast strength and orientation of synoptic-scale features as

seen in the GFS forecast, although the track of Ike in both of these models was somewhat better than the GFS. This finding suggests that the forecast challenge with Ike was not simply due to errors inherent to the GFS model or the GSI analysis scheme, but may have been due to errors in the initial analysis at 0000 UTC 9 September that were apparent in multiple modeling systems.

c. NCEP GEFS

Initial condition uncertainty can be further examined using GEFS initialized at 0000 UTC 9 September. The forecast tracks of the 20 perturbed GEFS members show one cluster of solutions near the observed track of
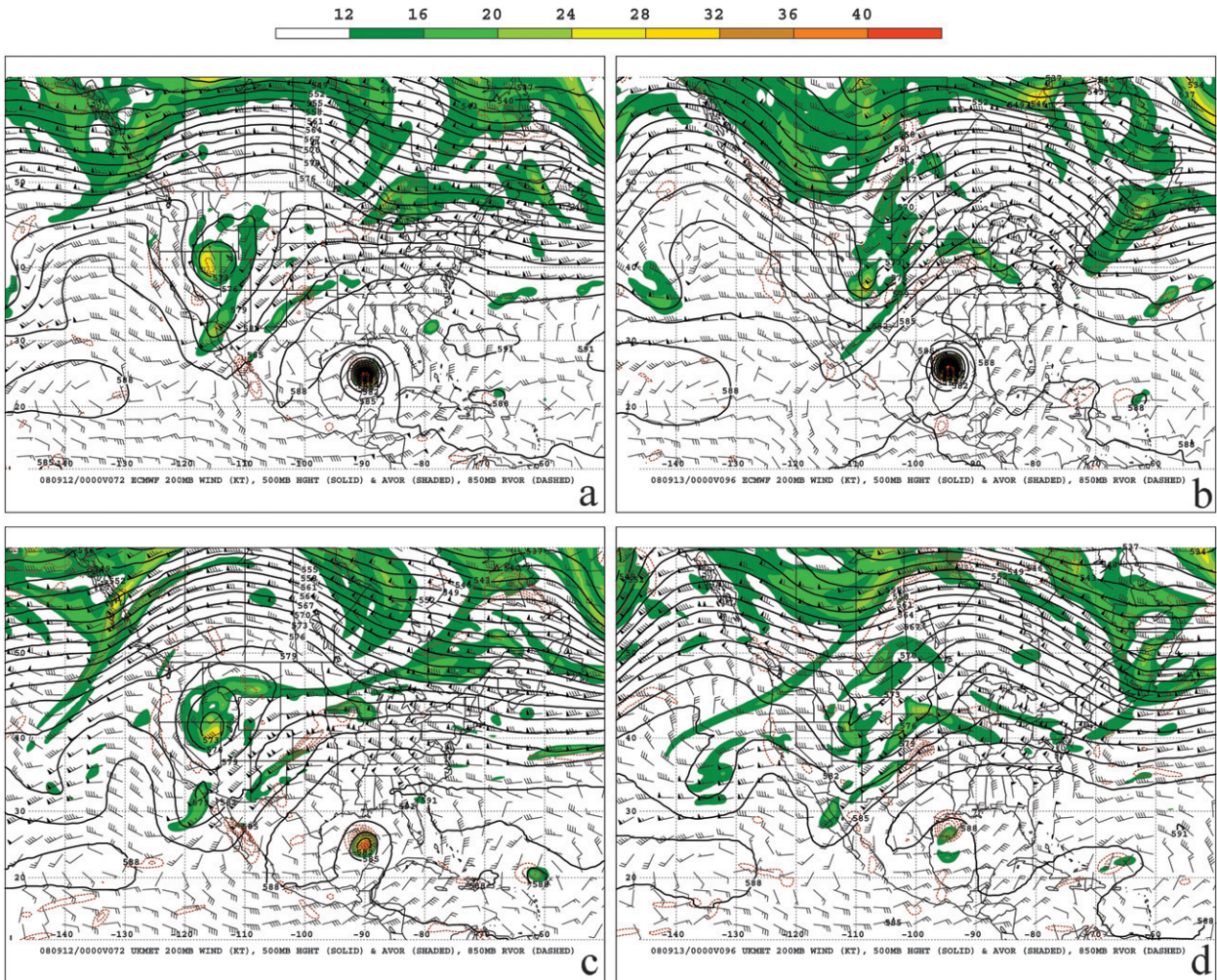
FIG. 5. As in Fig. 2, but for ECMWF forecast from 0000 UTC 9 Sep 2008 cycle valid at (a) 0000 UTC 12 Sep and (b) 0000 UTC 13 Sep 2008; (c),(d) as in (a),(b) but from UKMET forecast from the 0000 UTC 9 Sep 2008 cycle.

Ike and a second cluster farther south showing landfall in southern Texas (Fig. 7). However, the envelope of solutions shows a range of landfall locations from northern Mexico (member 10) to near the Mississippi–Alabama border (member 18). Given that the spread in the 2008 version of the GEFS ensemble is due only to initial condition uncertainty and its growth through the forecast period, the variability in the initial analysis of the subtropical ridge north of Ike will be examined.

The ensemble members show considerable variation in their analysis of the western extent of the ridge, as represented by the 591-dam 500-hPa geopotential height contour (Fig. 8a). One cluster of members shows the contour terminating over northern Florida, and another cluster of members is similar to the GFS analysis or slightly weaker. Interestingly, the analyzed height contour from members 10 and 18, which showed the largest departures from the observed track of Ike, show only slight

differences in their initial analysis of the ridge north of the hurricane (Fig. 8a). This suggests that only slight differences in the analysis of the midlevel ridge north of Ike could result in large differences in the eventual track of the cyclone. By 0000 UTC 12 September, the difference in the location of the 591-dam contour among the ensemble members has increased dramatically (Fig. 8b), particularly between members 10 and 18, with member 10 taking the contour as far west as central Louisiana, while member 18 shows a much weaker ridge and height contour that remains east of Florida. There is considerable spread among the remainder of the ensemble members in their orientation and strength of the midlevel ridge, with several of the members and the ensemble mean showing a weaker ridge centered farther east than the deterministic GFS. When compared to the GFS FNL analysis valid at this time (Fig. 4a), none of the ensemble members shows the midlevel anticyclone as
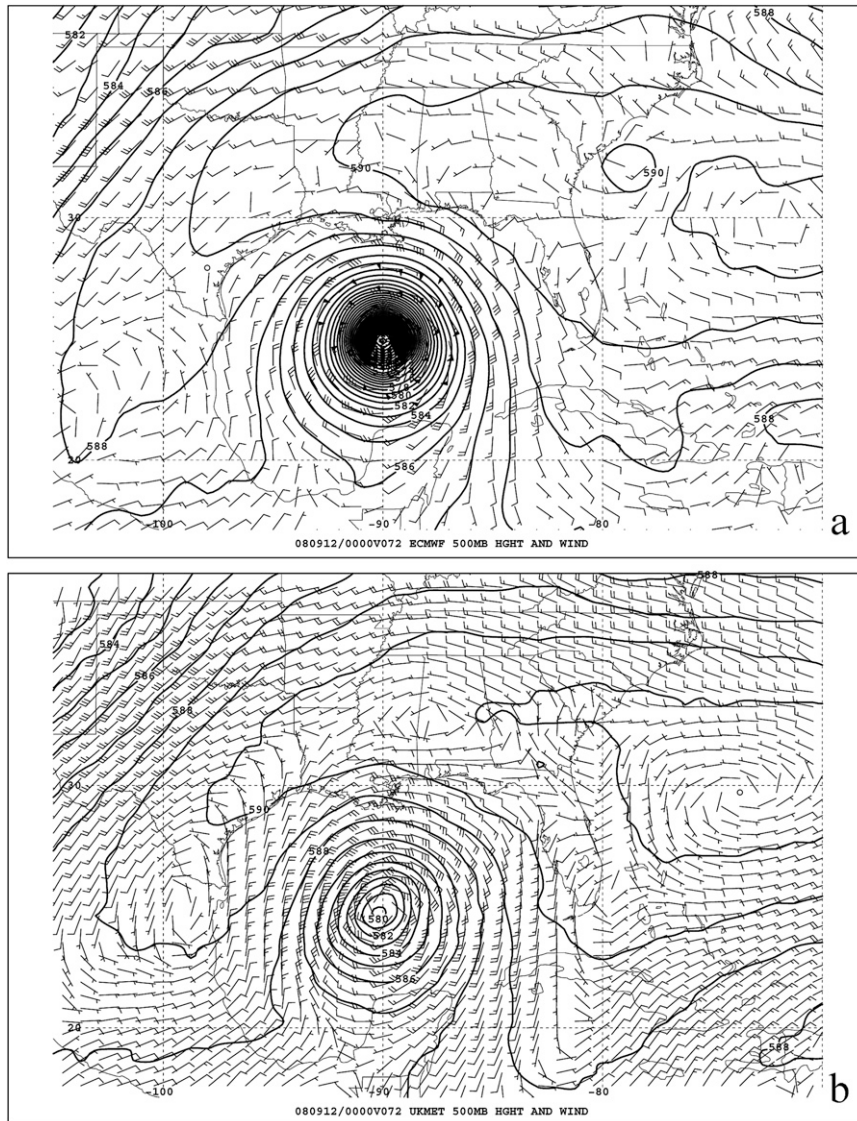
FIG. 6. As in Fig. 4, but for (a) 72-h ECMWF and (b) 72-h UKMET forecasts valid at 0000 UTC
12 Sep 2008.

far north as analyzed, but at least a few members show a more northwest-to-southeast orientation of the height contour, as depicted in the analysis.

These differences in the orientation and strength of the midlevel ridge in the GEFS ensemble members, along with the sizeable spread in those members' track of Ike, suggest that only slight differences in the initial analysis and evolution of the midlevel ridge north of Ike resulted in large differences in the eventual track of the cyclone. For comparison, in the ECMWF Ensemble Prediction System, many of the members that forecast a more southerly track possessed a midlevel ridge that is stronger 24 h into the forecast than those in the members that eventually recurve Ike (not shown). This characteristic is

exaggerated further at 72 h, confirming that the ultimate track of Ike is influenced by the ridge. In contrast, no clear correlation between the track of Ike and the western U.S. trough is evident at 24 h, although at 72 h the trough (now inland) is part of a more amplified pattern for those members that take Ike on a more northward track.

5. Sensitivity experiments

a. Choice of initial perturbations

The features to be perturbed in the GFS analysis at 0000 UTC 9 September 2008 were chosen subjectively, based primarily on the synoptic reasoning of section 3 and the performance of the operational models and
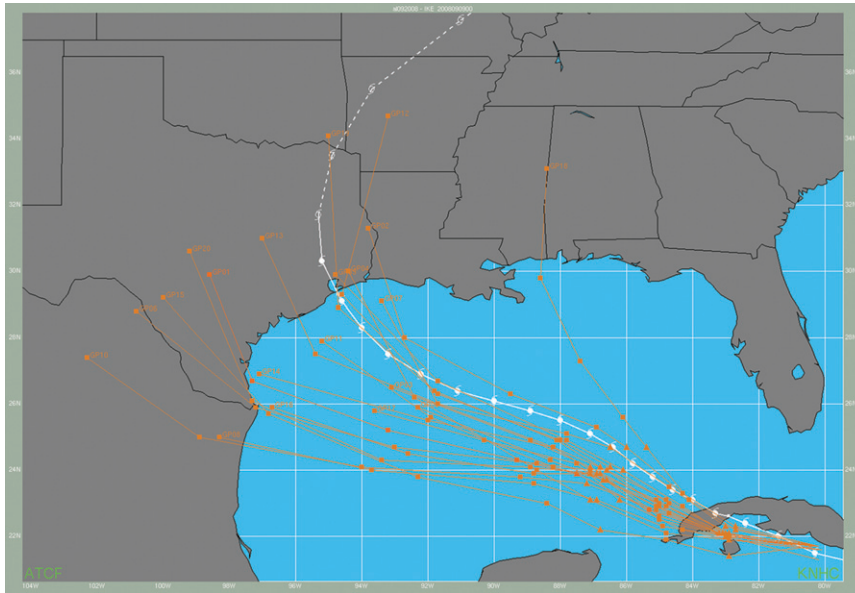
FIG. 7. Forecast tracks of Ike from all 20 GEFS perturbed members (orange) from the 0000 UTC 9 Sep 2008 cycle, with member 10 farthest to the left and member 18 farthest to the right of the best track of Ike (white).

ensemble prediction systems described in section 4. The locations of the synthetic temperature ''observations'' assimilated into the GSI for each experiment e1–e8 are illustrated in Fig. 9, and the magnitude and location of the observations prescribed in each experiment are listed in Table 3. A series of systematic experiments was performed to determine the location and magnitude of temperature perturbations, and a subset of all the experiments is shown here. In all assimilation cases, the wind fields are adjusted via information in the background error covariance matrix and the dynamic balance constraint applied in the GSI. For example, the assimilation of an observation at 300 hPa that is 4 K warmer than the analysis leads to a consistent, moderate lowering of the height pattern through the middle and lower troposphere, in the regions that are expected to be important for steering the tropical cyclone (Fig. 10). This logic is applied in experiment e1, in which the purpose is to weaken the midlevel ridge directly north of Ike and thereby induce a northward shift of the ridge in the 2–3-day forecasts, creating a stronger northward component in Ike's track than was seen in the operational models. In contrast to experiment e1, an upper-tropospheric observation 4 K cooler than the operational analysis is prescribed in experiment e2, in order to amplify the midtropospheric ridge over Alaska and thereby amplify the meridional pattern downstream. In experiments e3, e4, and e5, the goal is also to modify the amplitude and phase of the pattern across the contiguous United States, by deepening the associated midtropospheric short-wave trough in each

case. For experiment e6, it is suggested that the insertion of a midlevel warm perturbation in the outflow ridge associated with Tropical Storm Lowell would act to amplify the upper-level ridge downstream, thereby modifying the track of Ike (R. McTaggart-Cowan 2010, personal communication). The final two experiments are based on strengthening a tropical cyclone by warming the core in the middle levels, thereby yielding a deeper cyclone that is more consistent with observations than the model analysis. In experiment e7, a deepening of Lowell is proposed, in order to amplify the subtropical pattern downstream. Such an upscale transfer of energy from the meso- to synoptic scales may occur within 1–2 days (Zhang et al. 2003). Finally, in experiment e8, the deepening of Ike may possibly modify its interaction with its immediate environment. The number of synthetic observations assimilated serially in each experiment is selected to be consistent with the scale of the feature being perturbed, with six observations chosen for each of the two synoptic ridges (e1, e2), four for the short-wave troughs (e3, e4, e5) and outflow ridge (e6), and one observation for the Tropical Cyclones Lowell and Ike (e7, e8).

It is important to note that these perturbations were not made in an effort to try and correct the GSI analysis, but to investigate the sensitivity of the track forecast of Ike to the features perturbed. Additionally, these experiments only reveal the sensitivity to a temperature perturbation in the selected areas, and changes in the track of Ike in the experiments may also be related to compensating processes in the GFS forecast model.
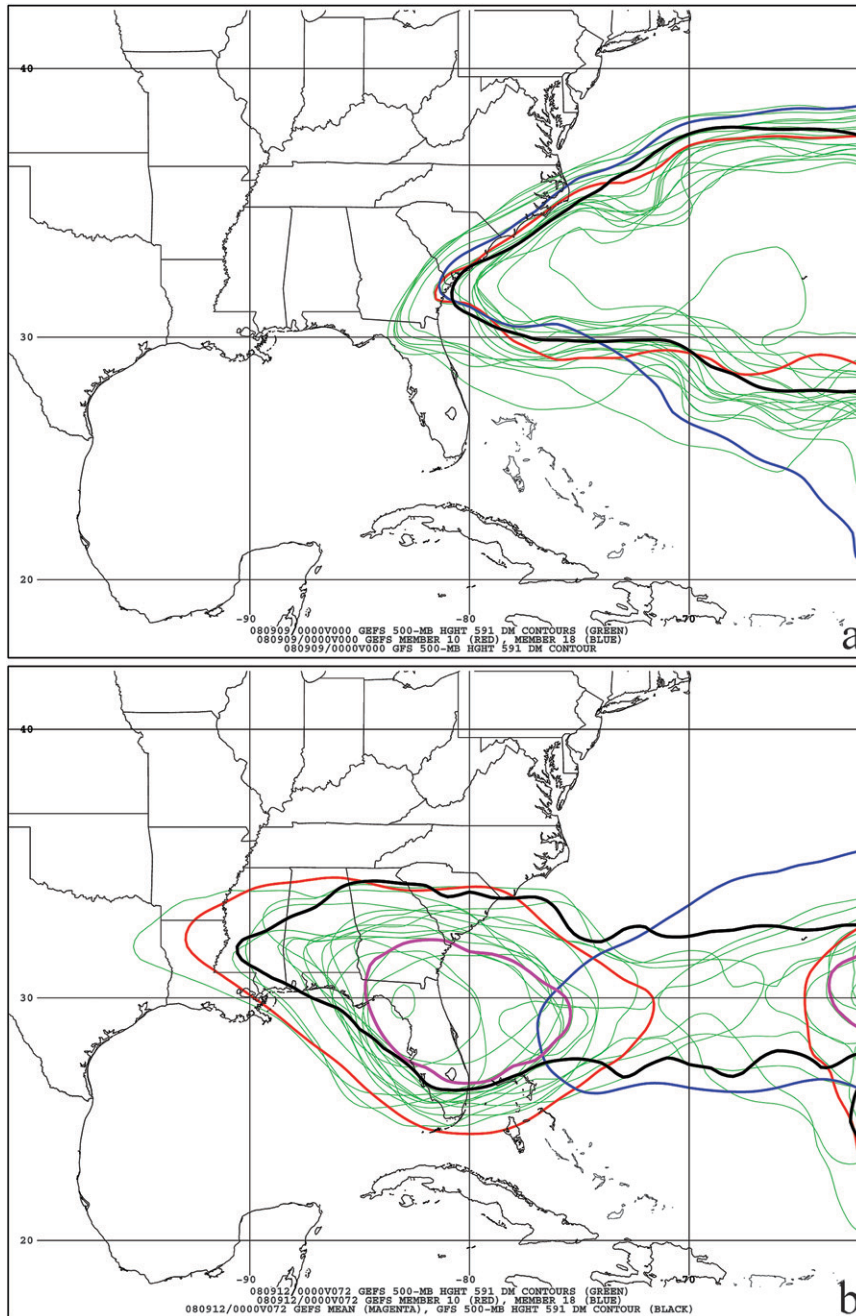
FIG. 8. (a) Analysis of 591-dam 500-hPa geopotential height contour from GEFS ensemble member 10 (red), member 18 (blue), all other GEFS perturbed members (green), and GFS (black), valid at 0000 UTC 9 Sep 2008. (b) As in (a), but for 72-h forecast valid at 0000 UTC 12 Sep 2008 with the addition of the GEFS ensemble mean (magenta).

b. GFS track forecasts

After creation of the modified GSI analysis for each experiment, the GFS model is integrated forward 120 h, yielding eight forecasts of Ike valid between 0000 UTC 9 September and 0000 UTC 14 September 2008 (Fig. 11).

First, it is evident that several of the initial perturbations made a minimal difference to the track forecast when compared with the operational forecast. Among these cases are the remote ridge over Alaska (e2), the short wave near the Great Lakes (e5), the warm sector of Lowell (e6), and the deepening of Lowell and Ike (e7,
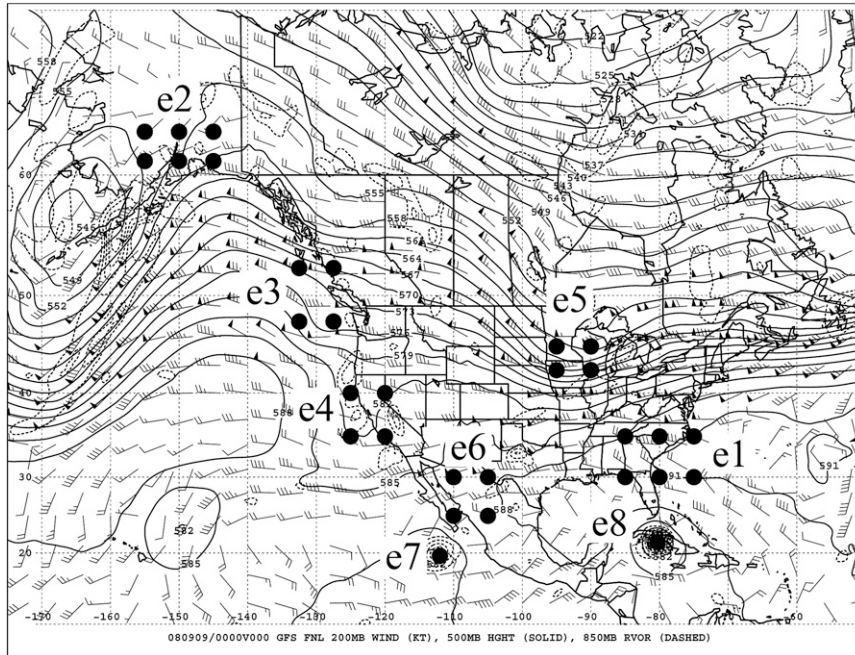
FIG. 9. GFS analysis of 200-hPa wind (barbs, kt), 500-hPa geopotential height (solid contours, dam), and 850-hPa relative vorticity (positive values contoured every $4 \times 10^5 \text{ s}^{-1}$) valid at 0000 UTC 9 Sep 2008. Dots represent the approximate locations of observations for GFS perturbation experiments e1–e8.

e8). All these tracks take Ike on a virtually straight path just north of due west, along a latitude too far south to allow recurvature. A modest northward deflection is evident when the short wave over British Columbia (e3) is deepened through the middle troposphere. For the deepened short wave farther south over California and offshore (e4), the change to the track is more substantial, with the landfall location approximately halfway between the operational forecast and the best track. It is noticeable that the track is shifted slightly north of the nonrecurving cases after 48 h, with a larger gain in latitude between 48 and 72 h. In this case, the forecast of Ike makes landfall after 108 h and is approaching recurvature at 120 h, approximately 1 day too late. The experiment that produces the largest northward deflection is the weakening of the westward extension of the subtropical ridge over the southeastern United States (e1). The gain in latitude and the relatively fast forward motion of Ike combine to produce landfall in less than 96 h, followed immediately by recurvature through the contiguous United States.

The perturbation to the ridge over Alaska (e2) produces a minimal impact on the forecast due to the influence (i.e., the difference between operational and perturbed GFS runs) remaining confined to the large-scale midlatitude pattern and not reaching Ike before

+72 h. By this time, the perturbed model run has already moved Ike a considerable distance westward, on a near-identical track to the operational GFS, and any modifications to Ike's immediate environment thereafter only produce minor changes in the track prior to landfall. For the deepening of the midlatitude short wave near the Great Lakes (e5), the influence propagates downstream toward the Atlantic Ocean, with weak amplification of the pattern upstream that does not significantly modify the ridge to the north of Ike. The amplification of the ridge in the warm sector to the northeast of Lowell (e6) again produces modest amplitude and phase changes to the large-scale midlatitude geopotential height pattern, with no coherent modification in the immediate vicinity of Ike. When Lowell itself is deepened, the influence is again tiny in the immediate environment of Ike. The deepening of Ike itself (e8) does not modify its environment appreciably, and the fact that Ike was already a well-developed deep cyclone in the operational GFS suggests that it was already being steered by the deep-layer flow. A wide range of other numerical experiments have been performed, for temperature observations warmer and colder than the operational GFS analysis, at different tropospheric levels at locations ranging from far upstream to the near environment of Ike. In all of these cases (not

TABLE 3. Summary of GFS perturbation experiments, identifying the types and locations of synoptic features in which the perturbations are made, the numbers of synthetic observations assimilated for each experiment, the differences between the observed temperature and the GFS analysis, and the levels at which these observations are assimilated.

Expt	Synoptic feature	Location	No.	Strength (K)	Level (hPa)
e1	Midlevel ridge A	Due north of Ike	6	4	300
e2	Upstream ridge	Southern AK	6	−4	300
e3	Short-wave trough D	BC	4	4	300
e4	Short-wave trough E	CA	4	4	300
e5	Short-wave trough B	W of Great Lakes	4	4	300
e6	Outflow ridge from TS Lowell	Northern Mexico	4	4	500
e7	Tropical Storm Lowell G	Off Baja California	1	5	500
e8	Hurricane Ike	Cuba	1	5	500
f1	Ridge e1 + short-wave e4	SW and SE United States	10	4	300

shown), the modification to the track of Ike was also found to be minimal.

It remains to be determined why the perturbations to the subtropical ridge (e1) and the California short wave (e4) produced a considerable change in track. Before doing so, it is worth noting that the track in the first 24 h of the forecast is almost identically due west in all cases e1–e8, regardless of the later evolution (Fig. 11). Given that all the perturbations (except e8) are remote, and that the GSI analysis increments produced by a single observation are localized (Kleist et al. 2009b and Fig. 10), it is not expected that the effects of the remote perturbations would modify the track of Ike within a day. However, the difference between the westward initial motion vector in all the GFS experiments and the initial west-northwestward motion in the best track is striking. One may speculate that this difference in steering motion is due to an inaccurate first-guess field and/or errors in a particularly difficult initialization of Ike over Cuba. On the other hand, the perturbation to Ike itself (e8) demonstrates little sensitivity, and there are no clearly discernible differences between these short-range forecasts and the GFS analysis at these times. Additionally, the National Oceanic and Atmospheric Administration (NOAA) Gulfstream IV (G-IV) aircraft had conducted synoptic surveillance missions every 12 h between 0000 UTC 7 September and 1200 UTC 8 September 2008, targeting the vicinity of Ike and the ridge to the north where possible. Up through this time, it would have been expected that the immediate environment of Ike was well initialized due to the volume of dropwindsonde data and the relatively close proximity of the North American rawinsonde network. It is worth noting that the increase in track error on 9 September occurred in a period when NOAA G-IV synoptic surveillance flights did not occur; however, it is difficult to determine if these data would have helped improve the forecast of Ike. It is possible that the lack of dropwindsonde data for the 0000 UTC 9 September cycle is related to the inaccurate initial motion

of Ike in the GFS at 0000 UTC 9 September 2008, but that determination is beyond the scope of this paper. However, it is important to note that the forecast tracks in all eight experiments propagate sufficiently northward between 12 and 24 h, leaving all simulations within tens of kilometers from the best track at 24 h. From here on, we will focus on the modifications to the track of Ike beyond 1 day, paying particular attention to the gain in latitude between 24 and 96 h, which is the primary factor in the subsequent recurvature or nonrecurvature around the time of its landfall.

By the design of perturbation e1, in which the midlevel ridge to the north of Ike is weakened at 0000 UTC 9 September 2008 (Fig. 12a), a signature of the weakened ridge remains to the north-northwest of Ike in the 24- (Fig. 12b) and 36-h (not shown) forecasts. The TC in the perturbed forecast is therefore able to gain more latitude than the operational forecast during this time. By 48 h, the ridge north of Ike becomes more amplified than the operational GFS, producing a track that now possesses a stronger westward component (Fig. 12c). Over the western United States at this time, a modest rise in the height field is evident over a large area. At 72 h (Fig. 12d), the rapid forward motion of Ike in the perturbed forecast has allowed it to begin interacting with the approaching short-wave trough to the northwest, thereby turning Ike more northward and making landfall only 12 h later (84 h), before recurving. We suggest that the gain in latitude between 24 and 36 h due to the weakened ridge led to an improved track forecast, albeit with landfall 12 h too early.

For perturbation e4, the influence of the strengthened trough over California on the forecast of Ike is not yet noticeable after 24 h (Fig. 13a). However, it is already evident that the original local perturbation has modified the large-scale pattern, with a general lowering of heights over the central United States. Over the next day, this lowering of heights becomes more pronounced. While Ike gains latitude for this reason, it does not achieve sufficient latitude to enable recurvature around
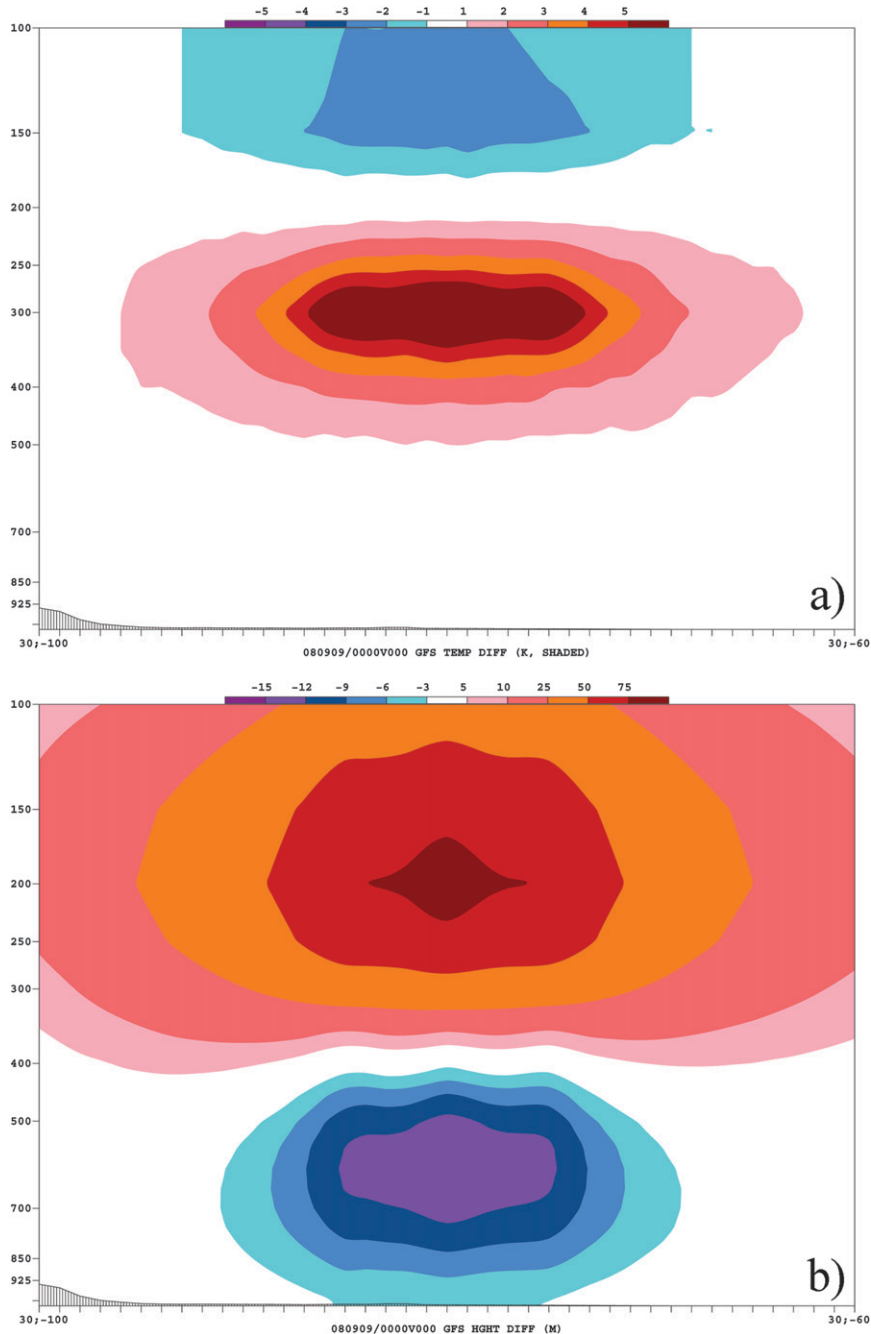
FIG. 10. (a) Temperature perturbation (K) for experiment e1 along a cross section from 30°N, 100°W to 30°N, 60°W. (b) As in (a), but for geopotential height (m).

the time of landfall (Fig. 11). The forward motion in the e4 forecast is considerably slower than that of the e1 forecast, given a weaker ridge to the north at 48 h (contrasting Figs. 12c and 13b). There is no clearly discernible change in the location or amplitude of the short wave over California that interacts with Ike, suggesting that the change in the long-wave pattern is mainly responsible for

the change to Ike's motion. For the forecast initialized using perturbation e3, a weaker modification to the same long-wave pattern as that of e4 is produced, resulting in a smaller northward deflection of the track.

The numerical experiments described so far comprise the assimilation of synthetic observations in one local region, associated with making perturbations to one distinct
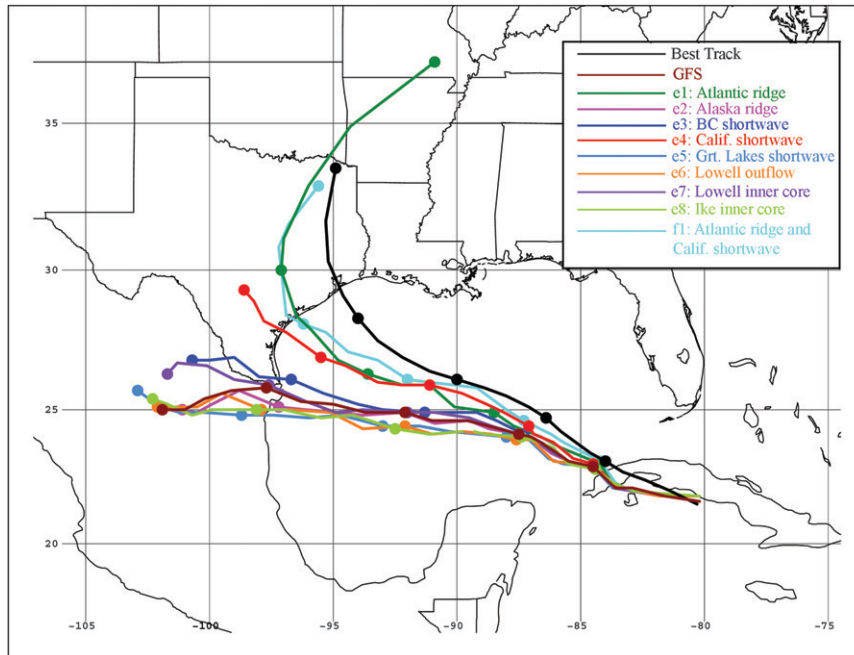
FIG. 11. The 5-day GFS track forecasts of Hurricane Ike initialized at 0000 UTC 9 Sep 2008, for the operational forecast (maroon), the eight perturbed analyses e1–e8, and the analysis f1 from the combined perturbations of the ridge north of Ike and the short wave off of California (see legend). The observed best track of Ike is shown in black. The dots along the tracks correspond to positions every 24 h during the forecast period.

feature. However, it is likely that errors in the analysis are due to difficulties in analyzing multiple synoptic features. To examine this, several additional experiments are performed with synthetic observations being assimilated within two features. In the majority of cases, the forecast of Ike's track is not modified significantly beyond what was seen by perturbing the western Atlantic ridge by itself (not shown). The most distinctive pair of perturbations, denoted f1 here, is due to the combination of e1 and e4, namely the simultaneous assimilation of synthetic warm upper-tropospheric observations in the western Atlantic ridge and the short wave over California. The track forecast produced by this combination yields an improvement over those produced by weakening the ridge and strengthening the short wave individually, although, as in all experiments described in this section, the track remained south of the best track (Fig. 11). The combined assimilation f1 leads to a greater gain in latitude by 2 days, and the perturbation of the short wave acts to slow down the forward motion of Ike compared with the simulation based solely on the Atlantic ridge perturbation (e1). This change in track is primarily due to the large-scale downstream influence from the short wave, which produces a considerably weaker ridge to the north of Ike compared with e1 (Fig. 14). Although the track still possesses too much of a westward component [resulting in a 135 n mi

(250 km) track forecast error at 4 days], the timing of the onset of recurvature is more accurate than that produced by either the Atlantic ridge perturbation (recurvature too fast) or the California short wave (recurvature too slow).

6. Conclusions

Operational NWP models and the official NHC forecast showed larger track forecast error for Hurricane Ike on 9–10 September 2008 relative to previous days (Table 1; Fig. 1). Much of the track model guidance showed Ike moving generally due westward across the Gulf of Mexico and making landfall along the central or southern Texas coast (Fig. 1), well to the south of the observed best track as Ike ultimately recurved after it made landfall near Galveston Bay. This large increase in track forecast error during the critical prewatch time frame, an important one for emergency management preparations, motivated this study into what caused this systematic track error in much of the model guidance.

A comparison of deterministic model forecasts from the GFS, ECMWF, and UKMET models from the 0000 UTC 9 September model cycle to GFS FNL analyses identified several possible synoptic-scale sources of error in the track of Ike. There was substantial spread in the analyzed western extent of the 591-dam 500-hPa geopotential height
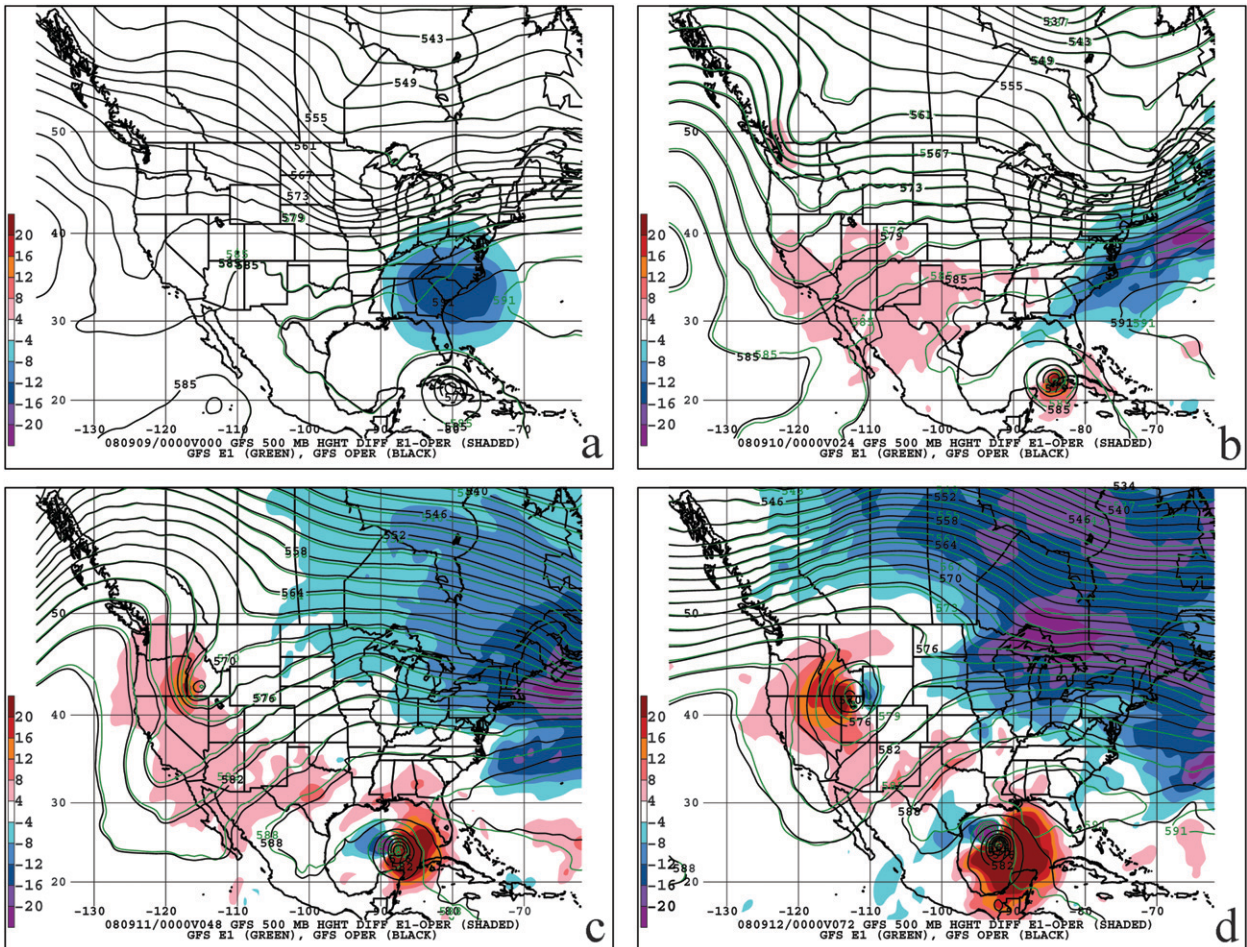
FIG. 12. GFS 500-hPa geopotential height (contours, dam) from the operational run (black) and perturbation experiment e1 (green) initialized at 0000 UTC 9 Sep 2008, and difference field e1–operational (shaded, m) for (a) the analysis, and (b) 24-, (c) 48-, and (d) 72-h forecasts.

contour associated with the subtropical ridge north of Ike in both the GEFS and ECMWF ensemble systems at 0000 UTC 9 September. Corresponding forecasts from the ensemble members showed sensitivity of Ike's eventual track to the amplitude and orientation of the ridge during the forecast period. Also, large differences were noted in the evolution of the ridge in the GFS, ECMWF, and UKMET models from that cycle compared to the GFS FNL analysis. All three models showed a more east–west-oriented ridge axis along the Gulf coast states, consistent with a more westward track of Ike.

To identify potential sources of initial condition sensitivity in the forecast of Ike, a series of perturbation experiments was performed using the GFS modeling system. Based on ideas derived from the operational model and ensemble forecasts, synthetic point "observations" of temperature, prescribed to be warmer or cooler than the operational GFS analyses, were selected

for assimilation into NCEP's Gridpoint Statistical Interpolation (GSI) scheme. Balanced perturbations to the analysis were produced, and the GFS was integrated forward 5 days from the new perturbed analysis. The largest northward deflection to the GFS forecast track of Ike arose from perturbations made within the western Atlantic ridge directly to the north of Ike, suggesting that an improved model forecast may have resulted if the ridge had been initialized to be weaker. Another improvement to the track of Ike was obtained by strengthening an upstream short-wave trough over California, which modified the large-scale pattern sufficiently after a day to begin weakening the midlevel ridge north of Ike. A combination of a weakened Atlantic ridge and a strengthened trough over California improved the timing of the track forecast. The introduction of initial perturbations in other regions of potential synoptic importance yielded minor changes to the forecast track,
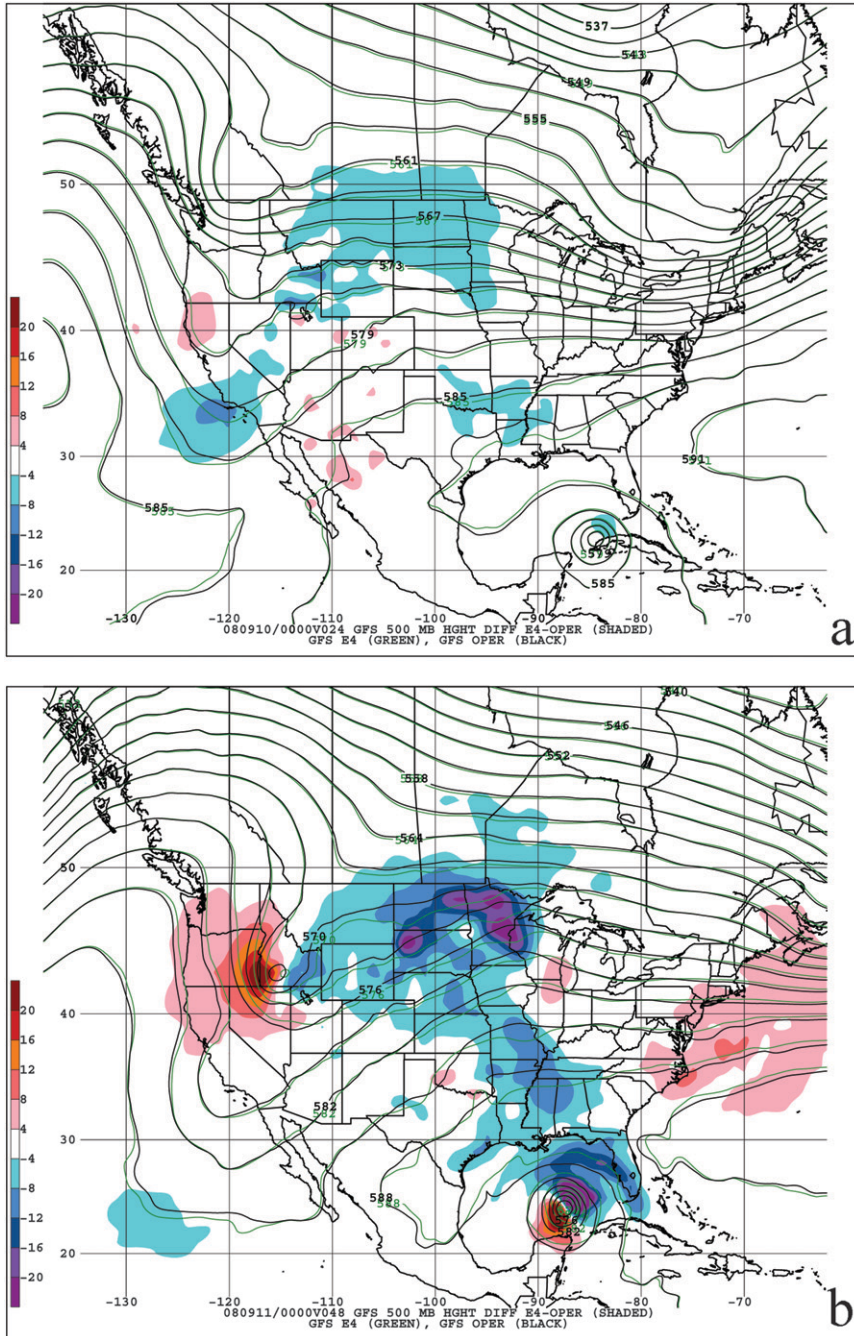
FIG. 13. As in Fig. 12, but for operational and perturbation experiment e4 for (a) 24- and (b) 48-h forecasts.

suggesting that the GFS forecast sensitivity of Hurricane Ike was limited to specific atmospheric features.

Careful interpretation of these results is necessary. Given that the forecast changes are based on analysis increments made by assimilating synthetic point observations of temperature in the three-dimensional variational GSI data assimilation scheme, we would not conclude

that errors in the initial conditions have been accurately quantified. We instead suggest that this diagnostic technique can be used to identify atmospheric features in which modifications may or may not be useful for improving the track forecast, using a similar philosophy to the numerical experiments proposed in Hoffman (2004). Given that both of the critical synoptic features
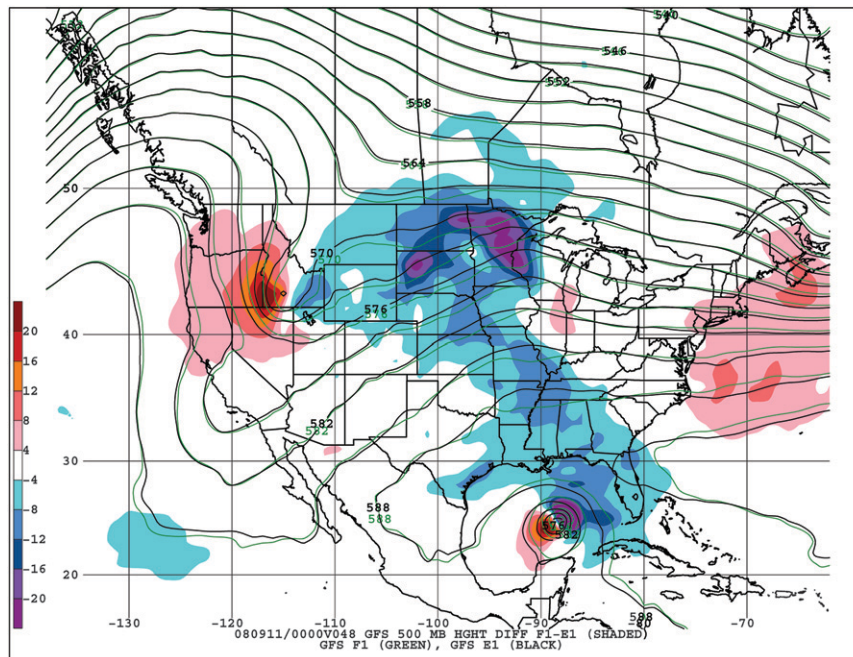
FIG. 14. As in Fig. 12, but for perturbation experiment e1 (black), perturbation experiment f1 (green), and the difference field f1 − e1 (shaded) for the 48-h forecast.

were located within or close to the North American radiosonde network for the 0000 UTC 9 September 2008 analysis cycle, hypotheses could be offered on optimal areas to assimilate observational data. These hypotheses could be tested by adding (removing) observational data from satellites, aircraft, rawinsondes or dropwindsondes to (from) the operational system; however, this capability was not available at the time of study. Additionally, comparisons of the results herein and those from widely used (but flawed) objective targeting strategies for tropical cyclones such as singular vectors (Peng and Reynolds 2006) or the ensemble transform Kalman filter (Majumdar et al. 2011) would yield further insights into the applicability of these strategies in an operational setting. It is also worth reemphasizing that the perturbation technique employed here only considers initial condition error, and does not account for errors inherent to the model physics.

The case study in this paper emphasizes that while numerical forecasts of tropical cyclone tracks have improved substantially over the past two decades, they remain susceptible to errors in the initial conditions. These errors may occur in a variety of features, including the tropical cyclone itself and its near and far environments. This study also suggests that information may be present in single-model ensemble systems that forecasters can use to gauge uncertainty in certain scenarios, despite the underperformance on average of single-model ensembles

in TC track forecasting compared to the multimodel consensus approach. Relevant information can include not only a range a possible forecast outcomes, but identification of critical features and areas of analysis uncertainty whose evolution can be monitored in observations and future model cycles.

In the future, superior forecasts may be accomplished via augmentation of the conventional observational network in areas of large initial condition uncertainty or growth, advances in data assimilation and tropical cyclone initialization, and improved model physics. This study also suggests that any forecast, based on consensus or other methods, should ideally be accompanied by an estimate of uncertainty that captures all potential sources of error, with weight given to those sources that are most likely to affect the forecast (such as the ridge to the north of Ike in this paper). If ensemble prediction systems comprising a large number of forecasts are able to provide a reliable range of solutions, the uncertainty in the forecast can then be communicated on a case-by-case basis. New on-demand capabilities to perturb models or select most likely ensemble solutions may be useful to improve our understanding of the forecast case in real time and reduce the range of uncertainty.

reviewers who provided helpful comments and suggestions to improve this manuscript.

REFERENCES

Berg, R., cited 2011: Tropical cyclone report: Hurricane Ike. [Available online at http://www.nhc.noaa.gov/pdf/TCR-AL092008_Ike_3May10.pdf.]

ECMWF, cited 2011: IFS documentation—Cy33r1 operational implementation 3 June 2008. Part V: Ensemble prediction system. [Available online at http://www.ecmwf.int/research/ifsdocs/CY33r1/ENSEMBLE/IFSPart5.pdf.]

Franklin, J. L., cited 2011: 2009 National Hurricane Center forecast verification report. [Available online at http://www.nhc.noaa.gov/verification/pdfs/Verification_2009.pdf.]

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

Henderson, J. M., G. M. Lackmann, and J. R. Gyakum, 1999: An analysis of Hurricane Opal's forecast track errors using quasigeostrophic potential vorticity inversion. *Mon. Wea. Rev.,* **127,** 292–307.

Hoffman, R. N., 2004: Controlling hurricanes. *Sci. Amer.,* **291,** 68–75.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, R. M. Errico, and R. Yang, 2009a: Improving incremental balance in the GSI 3DVAR analysis system. *Mon. Wea. Rev.,* **137,** 1046–1060.

——, ——, ——, ——, W. S. Wu, and S. Lord, 2009b: Introduction of the GSI into the NCEP Global Data Assimilation System. *Wea. Forecasting,* **24,** 1691–1705.

Liu, Q., T. Marchok, H.-L. Pan, M. Bender, and S. Lord, 2000: Improvements in hurricane initialization and forecasting at NCEP with global and regional GFDL models. National Weather Service Tech. Procedures Bull. 472, 7 pp. [Available online at http://www.nws.noaa.gov/om/tpb/472.pdf.]

Majumdar, S. J., S.-G. Chen, and C.-C. Wu, 2011: Characteristics of ensemble transform Kalman filter adaptive sampling guidance for tropical cyclones. *Quart. J. Roy. Meteor. Soc.,* **137B,** 503–520, doi:10.1002/qj.746.

Peng, M. S., and C. A. Reynolds, 2006: Sensitivity of tropical cyclone forecasts as revealed by singular vectors. *J. Atmos. Sci.,* **63,** 2508–2528.

Rappaport, E. N., and Coauthors, 2009: Advances and challenges at the National Hurricane Center. *Wea. Forecasting,* **24,** 395–419.

Sampson, C. R., and A. J. Schrader, 2000: The Automated Tropical Cyclone Forecasting System (version 3.2). *Bull. Amer. Meteor. Soc.,* **81,** 1231–1240.

——, J. S. Goerss, and H. C. Weber, 2006: Operational performance of a new barotropic model (WBAR) in the western North Pacific basin. *Wea. Forecasting,* **21,** 656–662.

Wei, M., Z. Toth, R. Wobus, and Y. Zhu, 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. *Tellus,* **60A,** 62–79.

Wu, C. C., and K. A. Emanuel, 1995a: Potential vorticity diagnostics of hurricane movement. Part I: A case study of Hurricane Bob (1991). *Mon. Wea. Rev.,* **123,** 69–92.

——, and ——, 1995b: Potential vorticity diagnostics of hurricane movement. Part II: Tropical Storm Ana (1991) and Hurricane Andrew (1992). *Mon. Wea. Rev.,* **123,** 93–109.

Zhang, F., C. Snyder, and R. Rotunno, 2003: Effects of moist convection on mesoscale predictability. *J. Atmos. Sci.,* **60,** 1173–1185.