The Behavior of Synoptic-Scale Errors in the Eta Model

KIMBERLY L. ELMORE, DAVID M. SCHULTZ, AND MICHAEL E. BALDWIN*

Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, and NOAA/National Severe Storms Laboratory, Norman, Oklahoma

ABSTRACT

A previous study of the mean spatial bias errors associated with operational forecast models motivated an examination of the mechanisms responsible for these biases. One hypothesis for the cause of these errors is that mobile synoptic-scale phenomena are partially responsible. This paper explores this hypothesis using 24-h forecasts from the operational Eta Model and an experimental version of the Eta run with Kain–Fritsch convection (EtaKF).

For a sample of 44 well-defined upper-level short-wave troughs arriving on the west coast of the United States, 70% were underforecast (as measured by the 500-hPa geopotential height), a likely result of being undersampled by the observational network. For a different sample of 45 troughs that could be tracked easily across the country, consecutive model runs showed that the height errors associated with 44% of the troughs generally decreased in time, 11% increased in time, 18% had relatively steady errors, 2% were uninitialized entering the West Coast, and 24% exhibited some other kind of behavior. Thus, landfalling short-wave troughs were typically underforecast (positive errors, heights too high), but these errors tended to decrease as they moved across the United States, likely a result of being better initialized as the troughs became influenced by more upper-air data. Nevertheless, some errors in short-wave troughs were not corrected as they fell under the influence of supposedly increased data amount and quality. These results indirectly show the effect that the amount and quality of observational data has on the synoptic-scale errors in the models. On the other hand, long-wave ridges tended to be underforecast (negative errors, heights too low) over a much larger horizontal extent.

These results are confirmed in a more systematic manner over the entire dataset by segregating the model output at each grid point by the sign of the 500-hPa relative vorticity. Although errors at grid points with positive relative vorticity are small but positive in the western United States, the errors become large and negative farther east. Errors at grid points with negative relative vorticity, on the other hand, are generally negative across the United States. A large negative bias observed in the Eta and EtaKF over the southeast United States is believed to be due to an error in the longwave radiation scheme interacting with water vapor and clouds. This study shows that model errors may be related to the synoptic-scale flow, and even large-scale features such as long-wave troughs can be associated with significant large-scale height errors.

1. Introduction

Forecast errors from numerical weather prediction models occur because initial conditions may be uncertain (e.g., especially over data-sparse areas) and because the models may not be ideal representations of the real atmosphere (e.g., imperfect model parameter-izations). Recognizing the causes of model errors, however, is not trivial. Generalizations about forecast errors have arisen, often without empirical evidence to support them. Some examples include the following:

- Short-wave troughs over the ocean are typically poorly forecast in numerical weather prediction models due to a lack of observational data.
- Once onshore with more abundant upper-air data (aircraft winds and radiosonde data), such errors typically decrease.
- Large-scale features are better handled by the models because they are better resolved by observational systems and because the models better represent large-scale dynamical processes.

* Current affiliation: Department of Earth and Atmospheric Sciences, Purdue University, West Lafayette, Indiana.

Corresponding author address: Dr. David M. Schultz, NOAA/NSSL/FRDD, Room 4360, 120 David L. Boren Blvd., Norman, OK 73072.
E-mail: david.schultz@noaa.gov

These statements make some physical sense, but are they valid? To our knowledge, few studies, if any, have addressed these questions. The goal of this paper is to provide some insight into the validity or nonvalidity of these statements.

As in Elmore et al. (2006), this study evaluates output fields from the operational Eta Model (hereafter referred to simply as Eta; Black 1994) run at the National Centers for Environmental Prediction (NCEP) and a version of the Eta run locally at the National Severe Storms Laboratory (NSSL) with Kain–Fritsch convective parameterization (EtaKF; Kain et al. 2001). The EtaKF model differed in four significant ways from the Eta. First, the EtaKF used the Kain–Fritsch convective parameterization scheme (Kain and Fritsch 1990, 1993; Kain 2004) instead of the Betts–Miller–Janjic scheme (Betts 1986; Betts and Miller 1986; Janjic 1994). Second, the EtaKF used a different shallow convective scheme (Baldwin et al. 2002). Third, the EtaKF used a fourth-order diffusion scheme (Kain et al. 2001) instead of a second-order scheme. Fourth, to obtain results in a timely manner in an operational setting, the EtaKF domain was smaller than the Eta, and the 24-h lateral boundary conditions for the EtaKF were derived from the 36-h forecast of a previous Eta run (rather than the 24-h forecast from the comparable Eta run, which would have not been available at NSSL in time). Errors are defined as the forecast field minus the initial analysis from the model. Using model initial fields (as opposed to radiosonde data) to verify model forecast fields has the advantage of being able to subtract gridded fields easily and obviating the need to interpolate radiosonde data to model grid points. Using the analysis for verification has drawbacks, however (e.g., White et al. 1999, p. 90). For example, in locations where observations are scarce or nonexistent, the verifying analysis is highly biased by the previous forecast, as will be seen in section 3 of this paper. In data-rich areas such as the continental United States, however, the nature of the 4D Eta data assimilation system (Black 1994) tends to mitigate this effect.

Following the methodology in Elmore et al. (2006), mean spatial bias errors for fields like geopotential height (hereafter, height) and wind are constructed for the 14-month period 26 January 2001 to 31 March 2002, a period where both Eta and EtaKF output were readily available to us. For example, the average 24-h forecast, 500-hPa height errors for all available 0000 UTC initializations over this 14-month period are shown in Fig. 1. During this period, the Eta Model tends to have negative height errors (e.g., forecast fields too low compared to the verifying model analy-
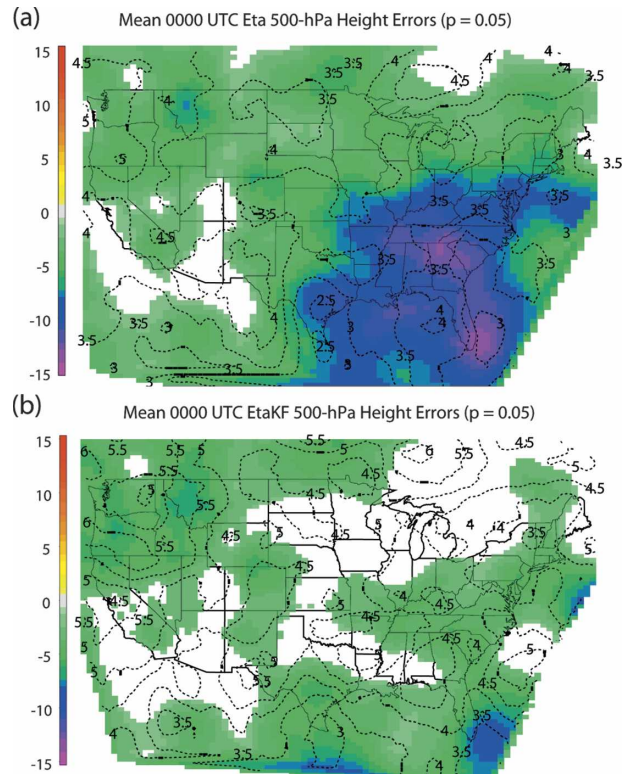

FIG. 1. The 500-hPa geopotential height 24-h forecast errors (m, colored) and the thickness of the 95% confidence interval (m, dashed lines) for the (a) 0000 UTC Eta and (b) 0000 UTC EtaKF for the 14-month dataset used in this study. Uncolored areas are not significant at the 95% level. See Elmore et al. (2006) for further information about the statistical techniques.

sis) across the United States, with minima less than -10 m over the southeastern United States (Fig. 1a). The EtaKF shows a tendency for less bias over most of the United States, particularly over the Southeast (Fig. 1b). (The version of the Eta evaluated in this study has been superceded by several significant model updates, and so the spatial bias errors shown here may not be applicable to the present version of the Eta.)

These persistent error fields over a 14-month period pose some interesting questions. What is the role of synoptic-scale variability on these mean error fields? Are these mean errors due to errors in troughs, errors in ridges, or both? What is the source of these errors: imperfect initial conditions or an imperfect model? The research presented in this paper does not provide complete answers to these questions. Instead, we present some intriguing examples of these forecast errors and, in the process, show some different approaches for model verification to begin to understand the physical nature of these model errors.
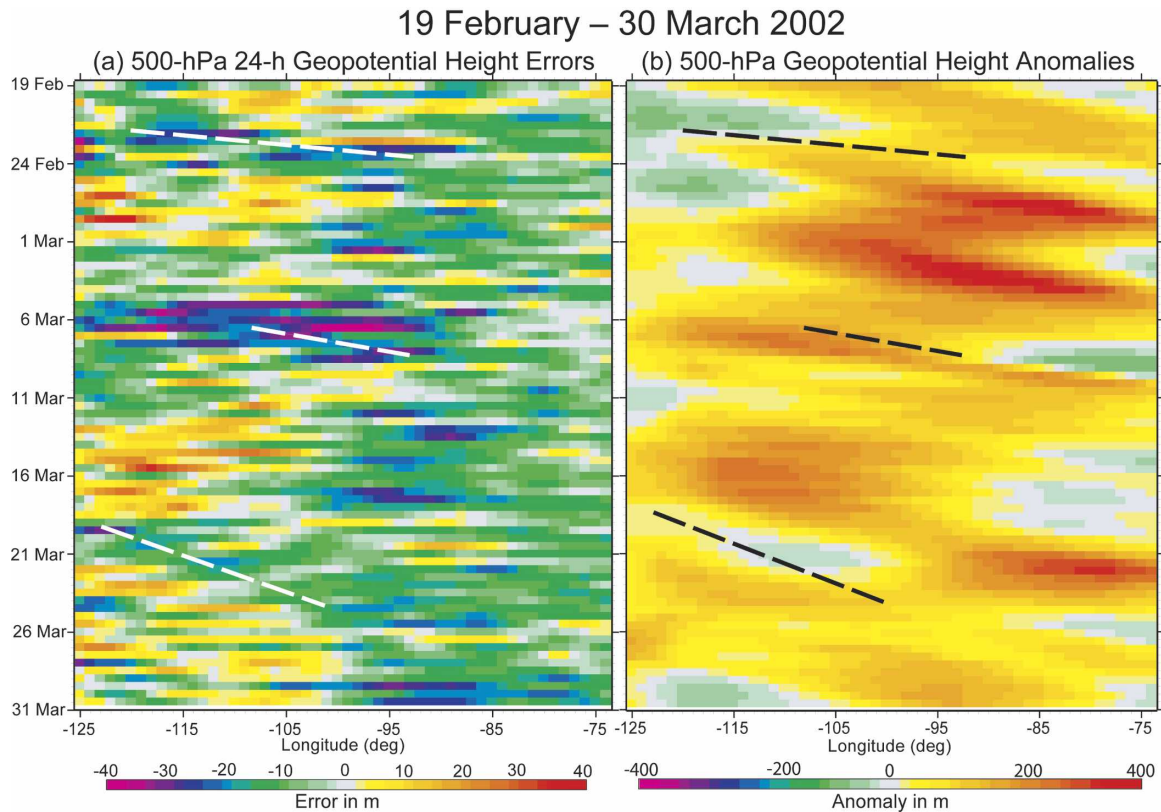
FIG. 2. The Hovmöller (1948) diagram between 19 Feb and 30 Mar 2002 for the 0000 and 1200 UTC Eta Model initializations for (a) 500-hPa 24-h geopotential height errors (error = 24-h forecast − initial analysis) and (b) 500-hPa geopotential height anomalies (calculated as the anomaly from the mean height at each grid point over the entire 14-month dataset). Dashed lines represent examples of some eastward-moving errors.

2. Tracking errors associated with short-wave troughs and long-wave ridges in time

One simple way to illustrate forecast errors is using a Hovmöller (1948) diagram of the 500-hPa height errors averaged over 30°–49°N constructed from both 0000 and 1200 UTC 24-h Eta forecast errors. Figure 2 shows these errors over a 40-day period from 19 February to 30 March 2002. This period is selected because it was one of the longest periods of uninterrupted Eta Model data (e.g., no missing model output) during the 14 months examined by Elmore et al. (2006). This sample data is typical of other periods of data in the 14-month dataset not discussed here.

On average during this period, positive errors tend to occur over the western United States and negative errors tend to occur over the eastern United States (Fig. 2a). The negative height errors over the eastern United States are consistent with the large spatial bias errors at 500 hPa for the whole 14-month period (Fig. 1a). Occasionally, as during 12–18 March, a large positive error persists over the western United States for several days,

perhaps even retrograding a bit. In contrast, some large negative errors persist and remain stationary for several days (e.g., 4–7 March). Figure 2a also suggests some slow movement of this large negative error eastward. Other times, smaller errors move much more quickly across the domain. A particularly long-lived negative height error on 21–23 February moved from 120° to 80°W (3500 km) in 48 h—an average speed of 20 m s^{-1}, the typical speed of mobile short-wave troughs and ridges (e.g., Hess and Wagner 1948; Schultz and Doswell 2000).

Our investigations show that many of these errors tend to be associated with short-wave troughs (e.g., Fig. 2b). For example, the short-wave trough arriving on the west coast of the United States on 0000 UTC 8 March, and traveling across the United States until 1200 UTC 10 March, was repeatedly underforecast in intensity by +30 to +60 m (Fig. 3). In addition, the trough tended to exhibit a phase error with greater positive errors upstream of the trough axis and greater negative errors downstream of the trough axis (e.g., Figs. 3b,c), suggesting the model was too fast at moving the trough
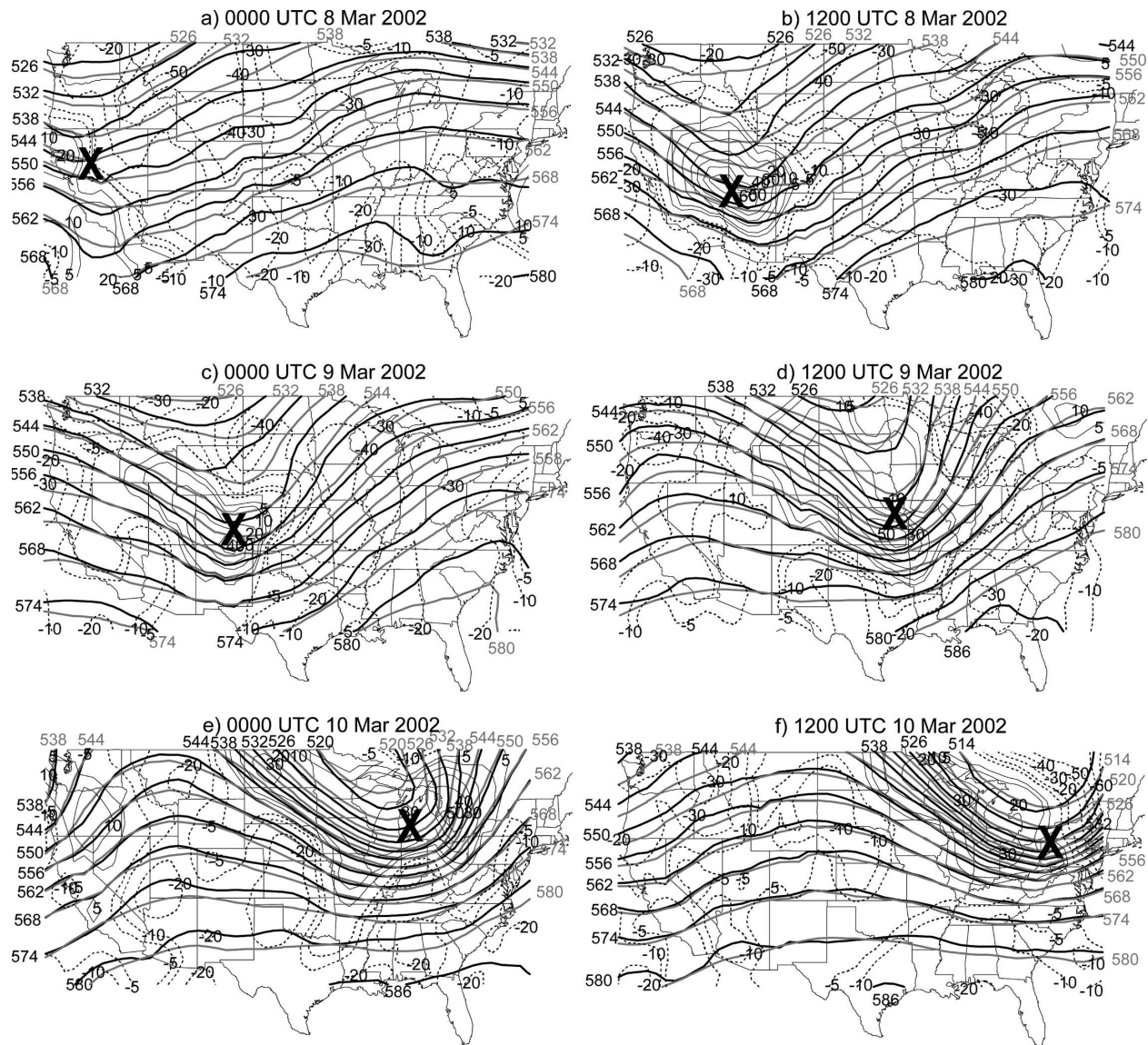
FIG. 3. The 500-hPa geopotential height 24-h Eta forecast (every 6 dam, gray solid lines), verifying analysis (every 6 dam, thick black solid lines), and errors (every 10 m, except for ±5 m; positive errors are thin black solid lines, negative errors are black dotted lines). Boldface "X" marks the short-wave trough centroid. (a) 0000 UTC 8 Mar, (b) 1200 UTC 8 Mar, (c) 0000 UTC 9 Mar, (d) 1200 UTC 9 Mar, (e) 0000 UTC 10 Mar, and (f) 1200 UTC 10 Mar 2002.

eastward. On the other hand, the short-wave trough in Oregon on 1200 UTC 23 Feb 2002 possessed a +50-m height error, followed by a +20-m error 12 h later, and eventually became a broad −10-m error as the trough weakened in the confluent flow downstream (Fig. 4).

In fact, as with the short-wave troughs in Figs. 3 and 4, most troughs are underforecast. A sampling of three 37–44-day periods of uninterrupted Eta Model data (4 March–18 April 2001, 21 April–28 May 2001, and 19 February–30 March 2002) indicates 31 (70%) of all 44 well-defined short-wave troughs arriving on the west

coast of the United States during these periods were underforecast. A different sample of 45 troughs that could be tracked easily across the country indicates that 20 (44%) had the magnitude of the errors decreasing in time, 5 (11%) increased in time, 8 (18%) had relatively steady errors, 1 (2%) was uninitialized entering the West Coast, and 11 (24%) exhibited some other kind of behavior. Thus, short-wave troughs are typically underforecast at landfall, but the forecasts tend to improve as the troughs move across the United States. Nevertheless, not all short-wave troughs are corrected as they
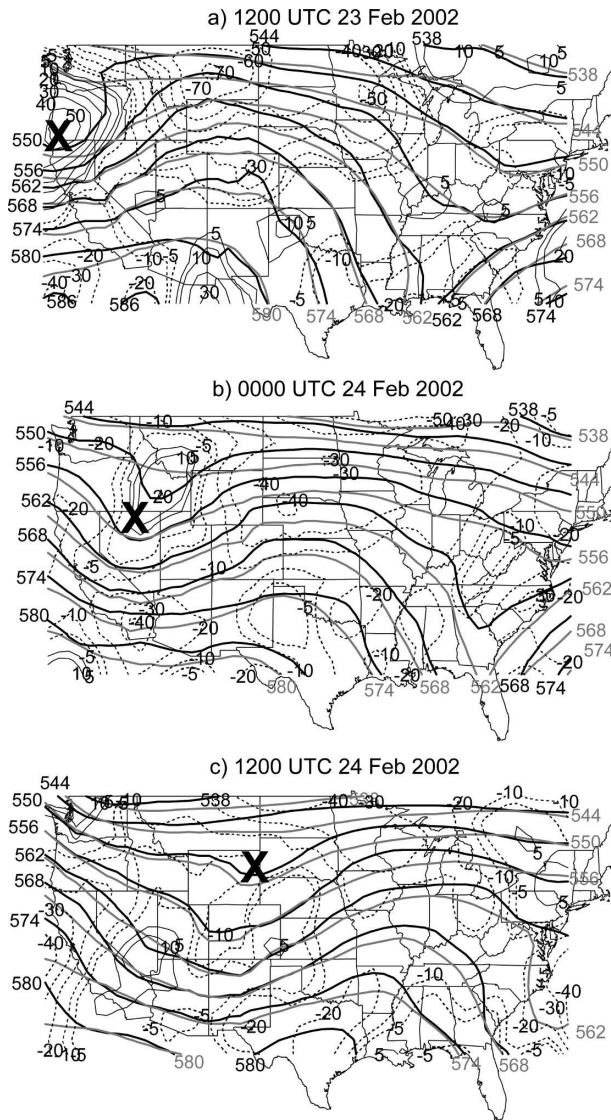
FIG. 4. Same as in Fig. 3, but for (a) 1200 UTC 23 Feb, (b) 0000 UTC 24 Feb, and (c) 1200 UTC 24 Feb 2002.

FIG. 5. Same as in Fig. 3, but for 0000 UTC 6 Mar 2002.

encounter increased data amount and quality over the United States.

Unlike short-wave troughs, in which errors could decrease, increase, or show other behaviors crossing the country, long-wave ridges tend to show much more consistent error behavior. Our investigation of the 14 months of model output shows that errors of −10 to −40 m tend to occur over a large area underneath the ridge (e.g., Figs. 3, 4, and 5), indicating that ridges tend to be underforecast. The large spatial scale of the errors associated with the ridge strongly suggests that initial condition errors are unlikely to be the cause. The different nature of these errors associated with troughs and ridges suggests that further investigation into these errors is warranted.
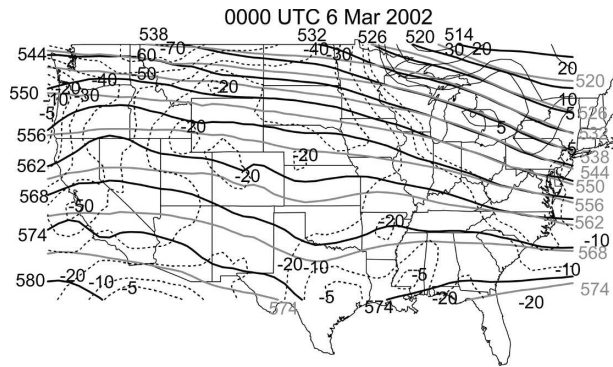
3. Error differences between troughs and ridges

To distinguish between troughs and ridges over the 14 months of model output, 500-hPa height errors within the Eta and the EtaKF are divided into two groups. Errors associated with gridpoint relative vorticity from the verifying analysis greater than zero are defined as errors within troughs, whereas errors associated with gridpoint relative vorticity less than 0 are defined as errors within ridges. This procedure is performed on both the 0000 and 1200 UTC runs of both the Eta and EtaKF for the entire dataset.

The mean errors associated with troughs and ridges are −3.2 and −8.6 m, respectively, for the 0000 UTC Eta runs and −2.3 and −6.0 m, respectively, for the 1200 UTC Eta runs. Based on a permutation test with 5000 replicates, these differences are significant at $p = 0.0004$ for both 0000 and 1200 UTC. Hence, on average, ridges suffer a stronger negative bias than troughs.

In addition, the root-mean-square (rms) errors behave differently for both troughs and ridges in the Eta. For 0000 UTC, the rms error for troughs is 16.0 and 15.0 m for ridges. Likewise, for 1200 UTC, rms errors are 16.8 m for troughs and 15.5 m for ridges. Similar to the mean errors described above, the difference in rms errors for troughs and ridges are significant at $p = 0.0004$. These values are slightly smaller than the 20-m grid-averaged rms error from the 24-h Eta forecasts from the 0000 UTC runs from January–March 1996 over the western United States found by White et al. (1999, their Fig. 7b). Thus, troughs suffer a larger rms error than ridges.

The difference between the mean errors and rms errors between troughs and ridges indicates that the errors associated with the troughs possess much more variability—large positive and negative errors associated with troughs cancel each other to produce a smaller mean error than that for ridges, albeit a larger rms error than that for ridges. Such a conclusion is consistent with our
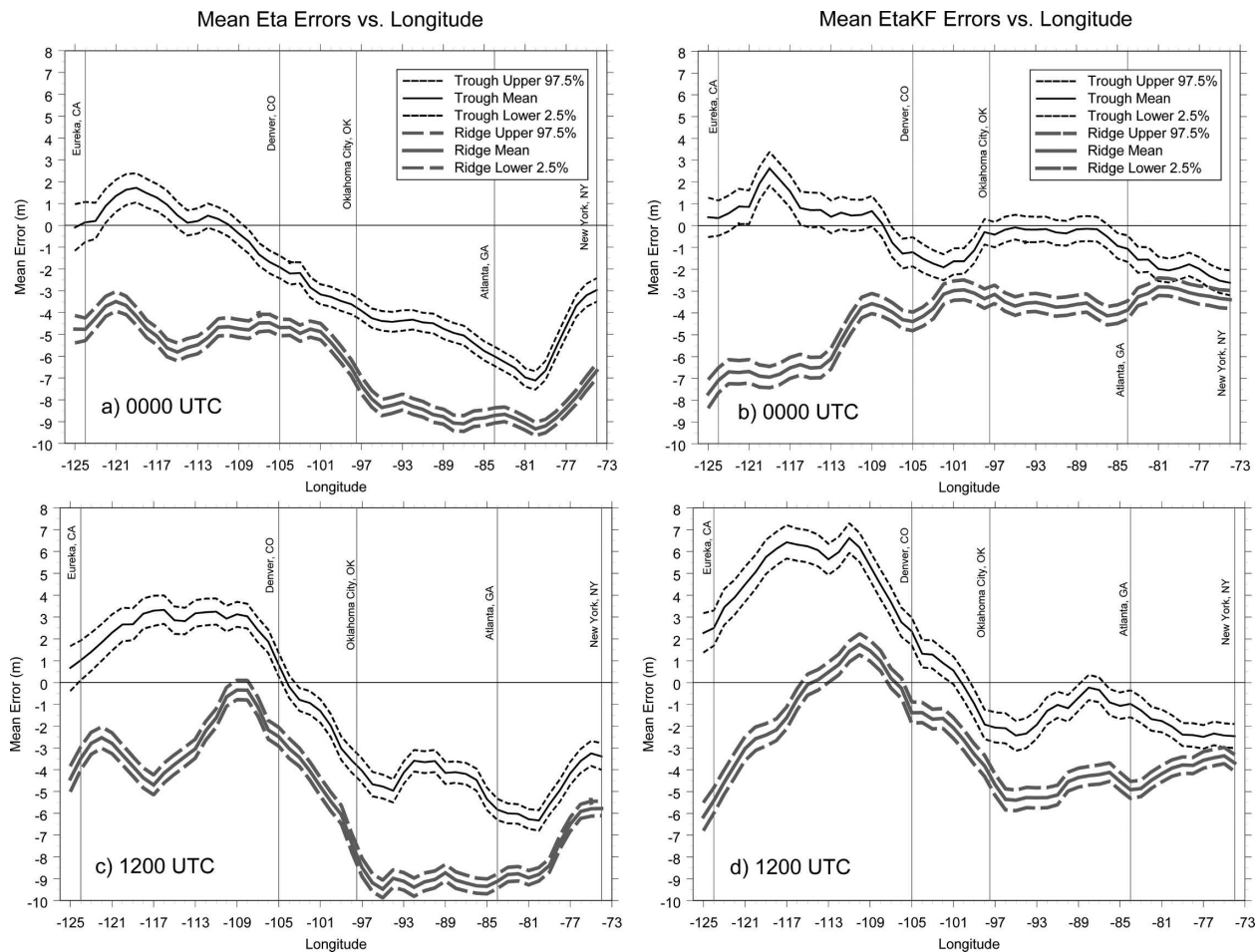
FIG. 6. Mean 500-hPa geopotential height 24-h forecast mean bias error (m) as a function of longitude for the (a) 0000 UTC Eta, (b) 0000 UTC EtaKF, (c) 1200 UTC Eta, and (d) 1200 UTC EtaKF forecasts for all grid points where the relative vorticity is positive (troughs, solid black lines) and where the relative vorticity is negative (ridges, solid gray lines). Dashed lines on either side of the solid lines represent the 95% confidence interval derived from a permutation test with 5000 replicates. Vertical lines represent approximate longitudes of some cities for reference.

manual observations of errors associated with troughs and ridges discussed in the previous section.

More insight comes when these errors are examined as a function of longitude (Fig. 6). Troughs in the Eta for both 0000 and 1200 UTC tend to have a small, but positive, mean bias error in the western United States (consistent with the results from section 2 for landfalling troughs), but the bias error becomes more negative to the east (Figs. 6a,c). Mean bias errors for ridges are always negative (consistent with the results of section 2) and become more negative to the east at about the same rate as troughs (Figs. 6a,c). Confidence intervals for the errors associated with troughs decrease moving eastward across the United States, but the confidence intervals for errors associated with ridges are relatively constant across the country (Figs. 6a,c). Thus, the variation in errors appears to be higher for troughs than for

ridges, especially in the western United States when troughs first arrive onshore. Errors in the troughs and ridges in the western United States tend to be more positive at 0000 than at 1200 UTC, but, in the eastern United States, the errors are slightly more negative at 0000 than at 1200 UTC, consistent with Fig. 3 in Elmore et al. (2006).

Mean errors in the 1200 UTC runs for the EtaKF (Fig. 6d) are similar in shape to those of the Eta (Fig. 6c), although more positive for troughs and ridges across nearly all longitudes. Mean errors in the 0000 UTC runs for the EtaKF (Fig. 6b), however, show much larger differences longitudinally than those of the Eta (Fig. 6a). Specifically, both troughs and ridges in the Eta are more strongly negative in the eastern United States (Figs. 6a,c) than those in the EtaKF (Figs. 6b,d), consistent with the results from Fig. 1.
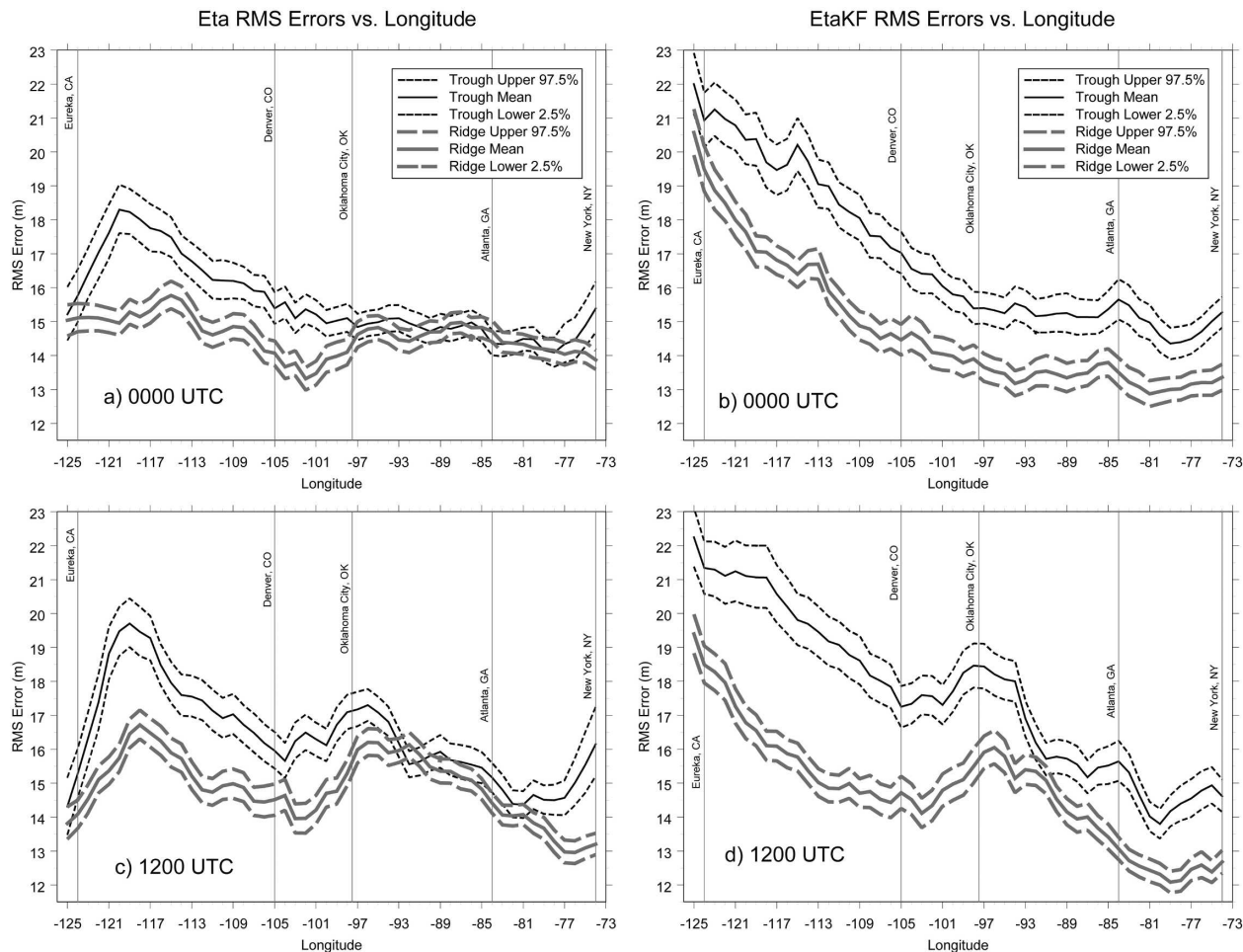
FIG. 7. Same as in Fig. 6, but for 24-h forecast rms error (m).

The variation in rms error with longitude also yields insight into how errors behave across the United States (Fig. 7). The rms errors plotted as a function of longitude for both 0000 and 1200 UTC Eta runs are similar for troughs and ridges across the United States, except in the west (Figs. 7a,c). The larger rms errors for troughs in the west indicate that landfalling troughs contain more large errors than do ridges, as suggested in section 2. However, the magnitude of the errors in troughs decreases significantly from west to east across the United States, while the magnitudes do not change as much across the United States for ridges, also suggesting some validity to the results in section 2.

A comparison between the rms errors in the Eta and EtaKF shows the effect that the western boundary condition has on the results (Fig. 7). In both the 0000 and 1200 UTC EtaKF runs, the rms errors are largest off the West Coast (Figs. 7b,d), as opposed to the lower values offshore in the Eta (Figs. 7a,c). Because the EtaKF uses old Eta Model forecasts to update the lateral boundary

conditions (section 1), errors in the 36-h Eta forecast compared to the 24-h EtaKF forecast are higher at the western edge of the EtaKF model domain (which is very near the western coast of the United States) compared to the western edge of the Eta domain (which is much farther west over the Pacific Ocean). Thus, rms errors in the EtaKF are largest near the West Coast and decrease eastward (Figs. 7b,d). This result is consistent with some of the problems discussed by Warner et al. (1997) in their tutorial on the importance of lateral boundary conditions in limited-area forecast models.

In contrast, rms errors in the Eta are quite small near the West Coast and increase onshore, reaching a maximum at about 118°W or roughly the longitude of Spokane, Washington, and Los Angeles, California (Figs. 7a,c). The behavior of these errors is associated with the large domain of the Eta. With little data over the ocean, the initial conditions over the ocean in the Eta depend strongly on the 6-h forecast of the previous run. Thus, rms errors over the ocean (west of roughly
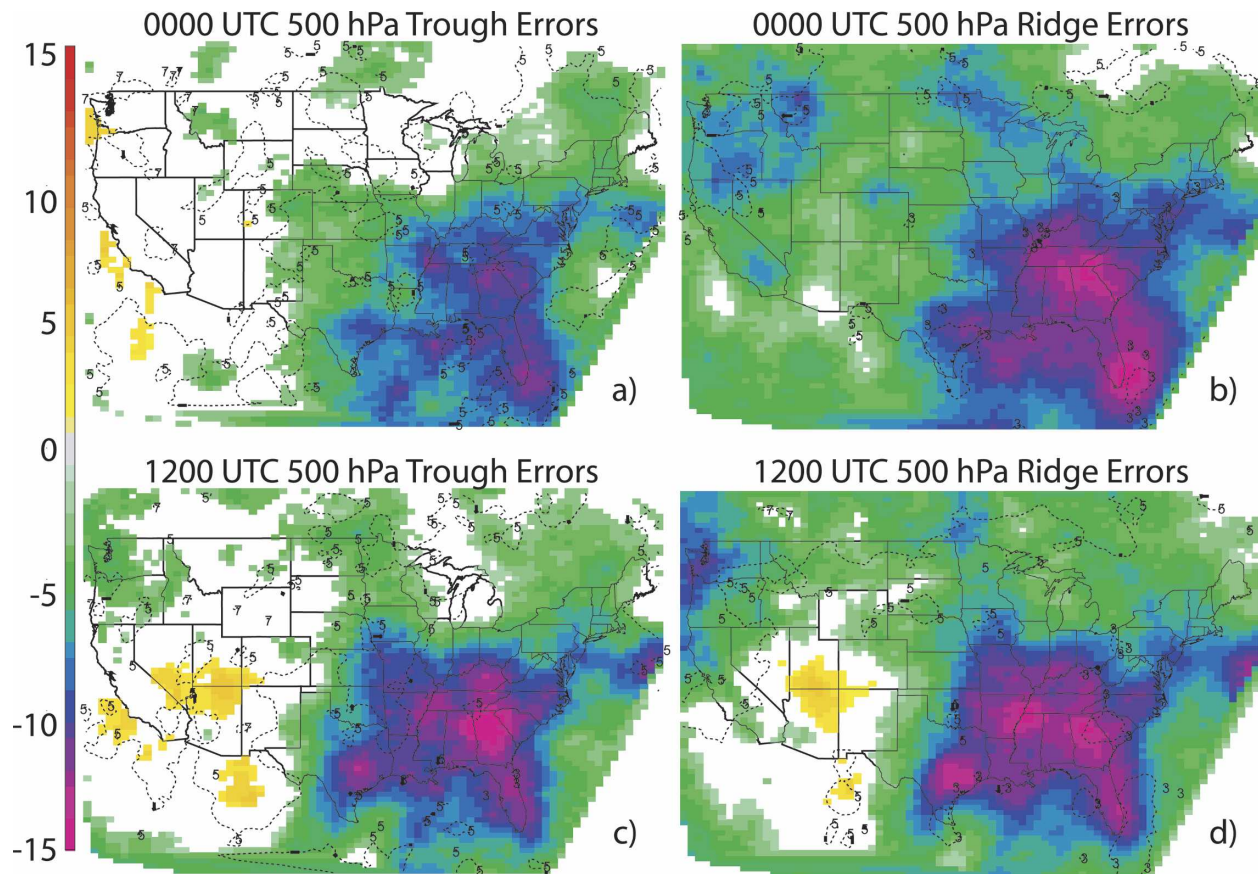
FIG. 8. The 500-hPa geopotential height 24-h Eta forecast errors (m, colored) and the thickness of the 95% confidence interval (m, dashed lines). (a) The 0000 UTC forecasts for all grid points where the relative vorticity is positive (troughs), (b) 0000 UTC forecasts for all grid points where the relative vorticity is negative (ridges), (c) 1200 UTC forecasts for all grid points where the relative vorticity is positive (troughs), and (d) 1200 UTC forecasts for all grid points where the relative vorticity is negative (ridges). Uncolored areas are not significant at the 95% level. See Elmore et al. (2006) for further information about the statistical techniques.

125°W) are small. Once onshore, the presence of more observed data to be assimilated into the model potentially introduces large changes to the previous 6-h forecast because the model reality does not match the atmospheric reality. Thus, errors are largest just inland, peaking around 118°W (Figs. 7a,c). In addition, model data with this newly updated information are translated inland. Given a 20 m s^{-1} speed, mobile synoptic features would move about 1700 km in 24 h (20° longitude at 40°N), which would be just about the size of the large rms errors (125°–105°W, or the distance from Eureka, California, to Denver, Colorado). Thus, mobile short-wave troughs at the longitude of Denver in a 24-h forecast moving at 20 m s^{-1} would have been initialized just at the west coast of the United States. Such rms errors over the eastern United States in the Eta and the EtaKF tend to be lower than those over the western United States (Fig. 7), likely due to the relative abundance of upstream data. Once weather systems move

offshore over the Atlantic Ocean (starting roughly east of 80°W), height errors increase due to the lack of verifying data, for the Eta (Figs. 7a,c) and, to a lesser extent, the EtaKF (Figs. 7b,d).

The spatial bias error patterns further reinforce these characteristics (Fig. 8). As suggested by Fig. 6 and the results of section 2, bias errors in ridges are nearly everywhere negative and significant (Figs. 8b,d). The variability of the errors for troughs, however, results in spatial errors without statistical significance in much of the west, but negative bias errors east of the Rocky Mountains (Figs. 8a,c). This characteristic is expected in light of the behavior of the rms errors as a function of longitude (Figs. 7a,c). Though the 1200 UTC runs show a region of significant positive bias errors near the Four Corners area for both troughs and ridges (Figs. 8c,d), the 0000 UTC runs do not show this feature as much (Figs. 8a,b). As in Fig. 3 of Elmore et al. (2006), negative bias errors are more severe in the southeast
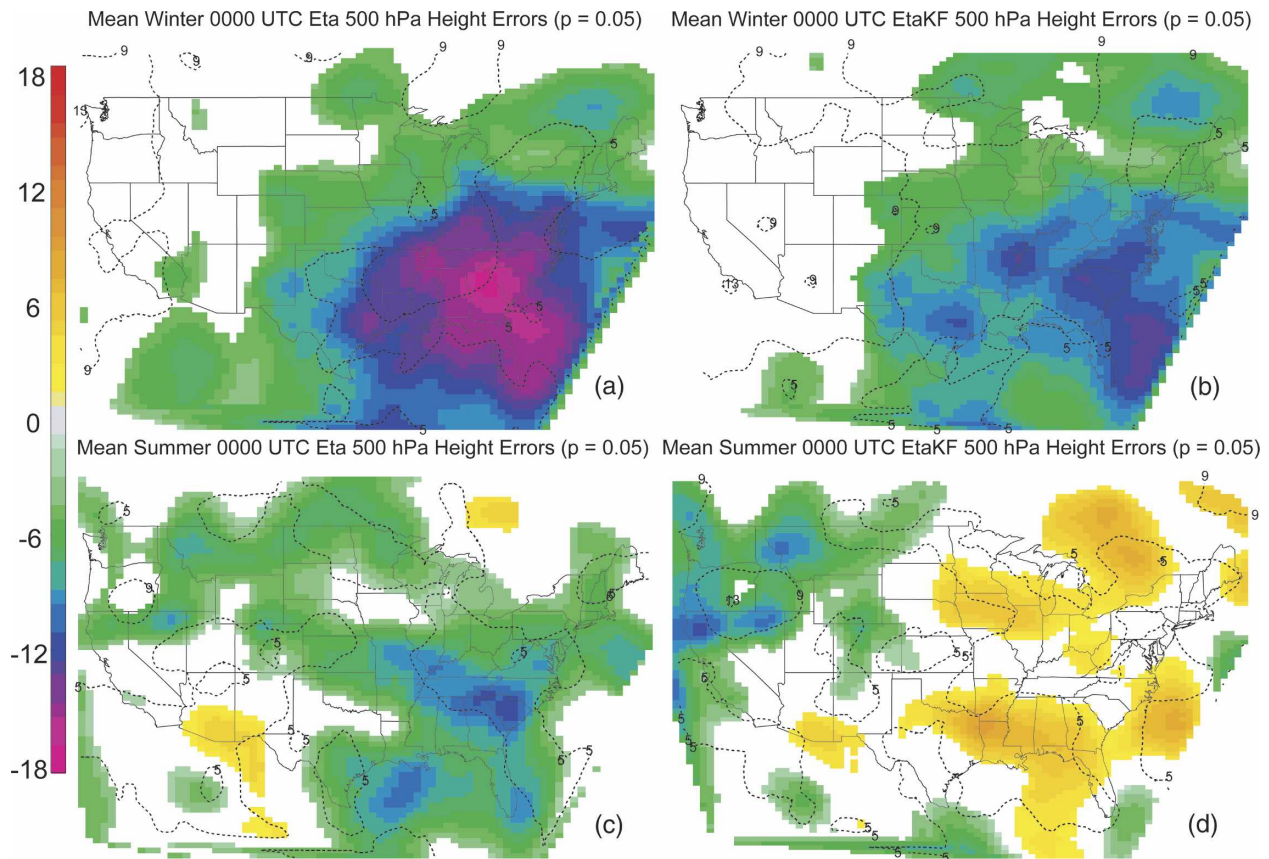
FIG. 9. The 500-hPa geopotential height 24-h forecast errors (m, colored) and the thickness of the 95% confidence interval (m, dashed lines). (a) All Eta forecasts in winter, (b) all EtaKF forecasts in winter, (c) all Eta forecasts in summer, and (d) all EtaKF forecasts in summer. Uncolored areas are not significant at the 95% level. See Elmore et al. (2006) for further information about the statistical techniques.

United States than anywhere else for both troughs and ridges (Fig. 8). The cause of this error in the Southeast is further explored in the next section.

4. Error differences between seasons

One hypothesis is that these large errors in the southeast United States are associated with the abundant low-level moisture found in the Southeast (e.g., Dodd 1965) and its vertical redistribution through the convective parameterization schemes. To test this hypothesis, the Eta and EtaKF model output was partitioned by summer (June, July, and August) and winter (December, January, and February). (Results partitioned by 0000/1200 UTC and by trough/ridge show similar spatial patterns to the errors in Fig. 9 and are not discussed here.) During winter, the spatial bias errors are similar between the Eta and EtaKF, although the maximum errors in the Eta over the southeast United States are about 50% larger (Figs. 9a,b). In summer, however, the

differences between the Eta and EtaKF are more substantial, with the Eta having negative height errors over the southeast United States and the EtaKF having positive height errors over the Southeast (Figs. 9c,d).

By partitioning the model output by model and season, we discover this negative bias error over the Southeast is largest during winter in the Eta (Fig. 9a). During summer, when convection is more active, the magnitude of this error in the Eta decreases (the bias error becomes less negative; Fig. 9c). Furthermore, the magnitude of this error is less during the winter in the EtaKF (Fig. 9b) than in the Eta (Fig. 9a). Finally, summer in the EtaKF shows the opposite error pattern, with mostly positive errors over the southeast United States (Fig. 9d).

These interesting results lead to the following hypothesis. A base negative error in the southeast United States appears in both the Eta and EtaKF. These large negative errors over the Southeast are likely *not* due to the convective schemes, which appear to be mitigating

these large negative errors to a greater degree during the summer. Furthermore, the presence of the negative errors in the EtaKF during winter (albeit reduced in amplitude; Fig. 9b) suggests that the different convective scheme in the EtaKF acts more strongly to mitigate this large negative error. In fact, during summer, the effect of this mitigation in the EtaKF may be *too large* as errors become positive in this region. Thus, when convective schemes are more active, the magnitude of this error in the southeast United States is reduced.

If the convective scheme is not responsible for these large negative errors, then what is? As previously mentioned, the large negative error in the Southeast is reminiscent of the region of abundant low-level moisture. The extreme sensitivity to low-level water vapor and clouds in the atmosphere and longwave radiation is well known. Perhaps errors in the longwave radiation scheme as it pertains to water vapor and clouds are responsible for these errors. Alternatively, the radiation scheme may not be at fault, but processes that transport water vapor into the southeast United States could be in error (e.g., surface evaporation over the ocean or land, horizontal advection). That both the Eta and EtaKF possess this error supports the argument that a mechanism common to both models (e.g., radiation scheme, initial conditions) is responsible. This argument is further advanced in the second half of the next section.

5. Possible causes of the errors and implications for numerical weather prediction

The sparse data over the ocean means that some short-wave troughs arrive on the West Coast without being depicted at all in the models. McMurdie and Mass (2004) found similar results with numerous examples of large initialization errors over the Pacific Ocean affecting landfalling weather systems. Colle et al. (2001) also found a tendency for model-predicted surface troughs over the eastern Pacific to be less intense (as measured by the sea level pressure and wind speed at trough passage) than observed at buoys. In addition, Hakim (2005) showed that the intensity of short-wave troughs tended to be underdone in four different operational forecast models. Likewise, the results from our study also suggest that errors in the models with landfalling weather systems are largest in the western United States, regardless of whether the short-wave troughs were tracked one at a time (section 2) or were analyzed en masse (section 3). These results were consistent for both 0000 and 1200 UTC runs of the Eta. (The EtaKF did not fit this pattern because of the different lateral boundary conditions, as discussed in section 3.) That

short-wave troughs experience their largest errors shortly after landfall indicates that the prevalence of satellite-derived radiances over the ocean alone is inadequate for accurate model initialization.

Although errors in the intensities of short-wave troughs tend to decrease as the troughs move eastward, problems remain with the intensity of some troughs that were not corrected as they encountered increased data amount and quality over the United States. This result is perhaps not surprising, given the average station spacing of the operational radiosonde network over the United States is about 400 km and the typical radius of short-wave troughs is 350–600 km (Hakim 2000).

Unlike the short-wave trough errors that were localized and perhaps easily explainable as due to the relative availability of upper-air data, the huge spatial scale and homogeneity of the errors associated with longwave ridges suggest that an error in the model physical parameterizations may have been to blame. Given that moist processes generally are absent underneath ridges, we speculate one likely candidate is longwave radiation. If the longwave radiation scheme were too active, the air underneath the ridge would cool too quickly, implying a cooler lower troposphere and underforecast ridges over a large area, much as observed. This physical linkage remains untested, although, interestingly, NCEP implemented a model change in July 2003, well after the model output studied in this paper. Specifically, the longwave radiation was updated hourly instead of every 2 h and other corrections were applied to the manner in which clouds interacted with radiation (more information available online at http://www.emc. ncep.noaa.gov/mmb/research/eta.log.html). These changes were intended to increase the cloud cover, resulting in less cooling in the lower troposphere (more information available online at http://wwwt.emc.ncep. noaa.gov/mmb/tpb.spring03/tpb.htm). These changes are the right sign to correct the underforecast 500-hPa geopotential heights in the southeast United States. In addition, even in the absence of clouds, a cold bias in the Eta has also been identified (more information available online at http://meted.ucar.edu/norlat/tencom/ p07.htm). Although evidence that the bias errors are larger in magnitude and scale in ridges than troughs suggests problems with the longwave radiation scheme and its interactions with water vapor and clouds, verification that the radiation scheme is directly responsible for such errors requires deeper investigation.

Although 10–40-m errors in the amplitude of the long-wave ridges may not appear impressive, Miguez-Macho and Paegle (2000, 2001) showed that obtaining good initial analyses of the long-wave pattern in a rela-

tively low-resolution global model was essential to realizing accurate forecasts, suggesting that large-scale errors in the initial state play a more prominent role in downscale uncertainty growth than some singular vector analyses imply. Roman et al. (2004) showed that errors in the intial state of the long waves (wavenumbers 0–15) accounted for 50% more error growth in 14-day global model forecasts than from errors in the initial state of the short waves (wavenumbers 16–42). Errico et al. (2002) confirmed that errors in the 500-hPa height field for zonal wavenumber 6 dominate over errors at other wavelengths. Clearly, these results show that getting the large scale correct is more important for global forecasting than getting the small scales correct. How these results apply for a limited-area model like the Eta on shorter time scales remains an open question, however.

6. Discussion

Errors in numerical weather forecasts are attributed to either errors in the specification of the initial conditions or imperfections in the model itself. This study explored the nature of bias errors observed in two operational models, the Eta and the EtaKF, using both manual and automated techniques to identify and understand the errors. The results from the two methodologies complement each other and provide consistency in interpreting the results. This study found that spatial bias errors in the models were due to errors in the initial and lateral boundary conditions, as well as suspected errors in the model itself.

Specifically, this study found that the Eta had a tendency to underforecast the intensity of the short-wave troughs. An obvious explanation may be that the troughs were undersampled by observational data, even after arriving onshore in the United States. An alternate explanation may be that the Eta Model possessed too much numerical diffusion (e.g., Baldwin and Wandishin 2002), damping the intensity of even well-initialized troughs and ridges over time, although comparisons between the Eta and EtaKF (which has less numerical diffusion) do not show any evidence of this process being responsible. Further experimentation with the models is required to be more specific as to the nature of these forecast errors.

Whereas the short-wave troughs tend to possess eastward-moving small-scale errors that vary in sign and intensity from run to run, long-wave ridges are consistently underforecast. This result suggests model initialization is unlikely to be the major factor in the errors. We speculate that model error, specifically longwave radiation, played a role instead. We also argue the per-

sistent negative bias over the southeast United States (e.g., Fig. 1) may be due to such errors in the radiation scheme, errors with water vapor and clouds, or both.

This article is not the first to relate the movement of forecast error fields in models to mobile synoptic-scale features. Fritsch et al. (2000, their Fig. 9) presented an example of the movement of forecast errors in 700-hPa temperature over the continental United States, finding continuity in the errors from one model run to the next, as we did. Fritsch et al. (2000) questioned whether the weighting of past model cycles in the formation of an initial analysis was too strong, such that the addition of new observational data was unable to overcome previous errors. The inability of the Eta to correct many of the forecast errors in the short-wave troughs from run to run (e.g., section 2 of this paper) further supports Fritsch et al.'s (2000) point.

Clearly, understanding the nature of these forecast and model errors requires more than cursory consideration. Very detailed dissection of model behavior may prove more fruitful in light of field-significant spatial bias errors (e.g., Elmore et al. 2006). Thus, looking at field-significant errors may help provide an understanding of how initialization errors or errors in the physical processes combine to generate the spatial bias errors.

REFERENCES

Baldwin, M. E., and M. S. Wandishin, 2002: Determining the resolved spatial scales of Eta model precipitation forecasts. Preprints, *19th Conf. on Weather Analysis and Forecasting,* San Antonio, TX, Amer. Meteor. Soc., 85–88.

——, J. S. Kain, and M. P. Kay, 2002: Properties of the convection scheme in NCEP's Eta model that affect forecast sounding interpretation. *Wea. Forecasting,* **17,** 1063–1079.

Betts, A. K., 1986: A new convective adjustment scheme. Part I: Observational and theoretical basis. *Quart. J. Roy. Meteor. Soc.,* **112,** 677–691.

——, and M. Miller, 1986: A new convective adjustment scheme. Part II: Single column tests using GATE wave, BOMEX and arctic air-mass data sets. *Quart. J. Roy. Meteor. Soc.,* **112,** 693–709.

Black, T. L., 1994: The new NMC mesoscale Eta model: Description and forecast examples. *Wea. Forecasting,* **9,** 265–278.

Colle, B. A., C. F. Mass, and D. Ovens, 2001: Evaluation of the

timing and strength of MM5 and Eta surface trough passages over the eastern Pacific. *Wea. Forecasting,* **16,** 553–572.

Dodd, A. V., 1965: Dew point distribution in the contiguous United States. *Mon. Wea. Rev.,* **93,** 113–122.

Elmore, K. L., M. E. Baldwin, and D. M. Schultz, 2006: Field significance revisited: Spatial bias errors in forecasts as applied to the Eta model. *Mon. Wea. Rev.,* **134,** 519–531.

Errico, R. M., R. Langland, and D. P. Baumhefner, 2002: The Workshop in Atmospheric Predictability. *Bull. Amer. Meteor. Soc.,* **83,** 1341–1343.

Fritsch, J. M., J. Hilliker, J. Ross, and R. L. Vislocky, 2000: Model consensus. *Wea. Forecasting,* **15,** 571–582.

Hakim, G. J., 2000: Climatology of coherent structures on the extratropical tropopause. *Mon. Wea. Rev.,* **128,** 385–406.

——, 2005: Vertical structure of midlatitude analysis and forecast errors. *Mon. Wea. Rev.,* **133,** 567–578.

Hess, S. L., and H. Wagner, 1948: Atmospheric waves in the northwestern United States. *J. Meteor.,* **5,** 1–19.

Hovmöller, E., 1948: North–south cross section of temperature, relative humidity, and wind in a well-marked zonal current over western Europe. *J. Meteor.,* **5,** 67–69.

Janjic, Z. I., 1994: The step-mountain Eta coordinate model: Further development of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.,* **122,** 927–945.

Kain, J. S., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.,* **43,** 170–181.

——, and J. M. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.,* **47,** 2784–2802.

——, and ——, 1993: Convective parameterization for mesoscale models: The Kain–Fritsch scheme. *The Representation of Cumulus Convection in Numerical Models, Meteor. Monogr.,* No. 24, Amer. Meteor. Soc., 165–170.

——, M. B. Baldwin, and S. J. Weiss, 2001: Utilizing the Eta model with two different convective parameterizations to predict convection initiation and evolution at the SPC. Preprints, *Ninth Conf. on Mesoscale Processes,* Ft. Lauderdale, FL, Amer. Meteor. Soc., 91–95.

McMurdie, L., and C. Mass, 2004: Major numerical forecast failures over the northeast Pacific. *Wea. Forecasting,* **19,** 338–357.

Miguez-Macho, G., and J. Paegle, 2000: Sensitivity of a global forecast model to initializations with reanalysis datasets. *Mon. Wea. Rev.,* **128,** 3879–3889.

——, and ——, 2001: Sensitivity of North American numerical weather prediction to initial state uncertainty in selected upstream subdomains. *Mon. Wea. Rev.,* **129,** 2005–2022.

Roman, J. C., G. Miguez-Macho, L. A. Byerle, and J. Paegle, 2004: Intercomparison of global research and operational forecasts. *Wea. Forecasting,* **19,** 534–551.

Schultz, D. M., and C. A. Doswell III, 2000: Analyzing and forecasting Rocky Mountain lee cyclogenesis often associated with strong winds. *Wea. Forecasting,* **15,** 152–173.

Warner, T. T., R. A. Peterson, and R. E. Treadon, 1997: A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction. *Bull. Amer. Meteor. Soc.,* **78,** 2599–2617.

White, B. G., J. Paegle, W. J. Steenburgh, J. D. Horel, R. T. Swanson, L. K. Cook, D. J. Onton, and J. C. Miles, 1999: Short-term forecast validation of six models. *Wea. Forecasting,* **14,** 84–108.