Error Climatology of the 80-Wave Medium-Range Forecast Model

DAVID R. WALKER*

Climate Prediction Center, National Meteorological Center, NWS/NOAA, Washington, D.C.

ROBERT E. DAVIS

Department of Environmental Sciences, University of Virginia, Charlottesville, Virginia

(Manuscript received 7 December 1993, in final form 27 February 1995)

ABSTRACT

A climatology of the once-daily (0000 UTC) 1000-hPa error fields of the National Meteorological Center's 80-wave Medium-Range Forecast (MRF) model is studied. An analysis of the error field has been conducted over the contiguous United States and over the Northern Hemisphere from 20° to 80°N for three warm and four cool seasons (9 September 1987 to 6 March 1991). Temporal and spatial mean error fields over various integration lengths are presented.

The skill, as measured by the anomaly correlation, has not significantly changed over the lifetime of the 80-wave MRF model. Anomaly correlation values at 1000 hPa and 300 hPa show that the model is retaining useful information about the anomalies in the height field out to about one week. A reduction in the model biases may reflect an improvement in model physics (longwave radiational calculations, etc.). The cool and warm seasons have distinctly different spatial error patterns. The 1000-hPa warm season shows spurious height falls over the southwestern United States that grow with increasing integration length. The 1000-hPa cool season underestimates the intensity of low pressure systems over and east of Hudson Bay and overestimates their strength over the Pacific Northwest.

Principal components analysis of the 429-variable error covariance matrices for the cool and warm seasons identifies 6 orthogonal variables that explain over 60% of the original error variance. MRF model problems appear to be related to problems the model has with simulating the atmosphere's interaction with orographic features (Alberta and Colorado Rockies), storm tracks and baroclinic zones (Gulf Stream region and United States—Canadian border), and persistent atmospheric features (Hudson Bay low, eastern Pacific subtropical high, and desert Southwest heat low).

1. Introduction

Medium-range weather forecasting is largely dependent upon integration of numerical prediction models on timescales of 3–10 days into the future. In the 1980s, the skill of global models improved substantially for integrations beyond three days (Simmons 1986; Tracton et al. 1989; Kalnay et al. 1990). Phenomena of interest to forecasters are synoptic-scale low and high pressure systems, upper-level cutoff lows, blocking highs, and storm tracks. All of these features are produced in the medium-range numerical models. However, incorrectly identifying the position and intensity of these features can mislead the forecaster and produce undesirable results (Livingston and Schaefer 1990). Unfortunately, as model integration lengths increase, the location of these features drift from reality. The end result may produce similar features in both the model and the verification atmospheres, but the location of these features are often out-of-phase with each other. One of the largest errors occurs when the model improperly accounts for the speed of major features, resulting in a model low pressure system located where the verification atmosphere has a high pressure system, and vice versa (Akyildiz 1985; Sanders 1992). Correction of this phase error is critical if these models are to be used deterministically.

Preliminary studies of the mean errors and biases in the National Meteorological Center's 80-wave Medium-Range Forecast (MRF) model have been conducted by Caplan and White (1989) for two different winter seasons each with approximately 90 individual model integrations. Sanders (1992) studied cyclone behavior in the 80-wave MRF, the United Kingdom Meteorological Office (UKMO), and the European Centre for Medium-Range Weather Forecasts (ECMWF) models out to 5 days of integration for 3 winter months in 1988/89. Sanders found that position errors in the location of cyclones in the analysis fields were about 150 km in the National Meteorological Center's (NMC) Global Data Assimila-

* Additional affiliation: SMART, Inc. Silver Spring, Maryland.

Corresponding author address: Dr. David R. Walker, 5200 Auth Road, RM 606, Camp Springs, MD 20746.

© 1995 American Meteorological Society
tion System (GDAS). These position errors continued to grow, almost linearly with time, at a rate of approximately 125 km day$^{-1}$. Grumm (1993) studied 18 months of forecasts from the 1200 UTC aviation (AVN) run of the MRF model that is the same numerical weather prediction model as the MRF but with a shorter data cutoff time. Grumm found that the AVN model was more likely to miss the development of a cyclone than to forecast spurious cyclones and that the model frequently predicted well-developed cyclonic circulations out to 72 h of model integration.

We will diagnose different aspects of the climatological mean error and the variability of that error and show its relationship to synoptic-scale disturbances. Unlike previous studies, a much longer record of MRF model forecasts will be considered. The forecast model in question, the temporal division of the dataset, and the spatial domains used are described in section 2. Time series of the spatial mean bias and the anomaly correlation over the period of record are given (section 3a). Statistical quantities are used to help identify the model's departure from deterministic predictability with increasing integration length (section 3b). Error growth of the 1000-hPa (1 hPa = 1 mb) height field is then shown over the contiguous United States (section 4). Multivariate statistics using principal components analysis (PCA) is employed to determine the dominant modes of spatial variation of the 1000-hPa error (section 5). These modes are then related to patterns found on daily weather maps (section 6).

2. Model forecasts

a. MRF model

The 80-wave MRF is a spectral model with a horizontal triangular truncation of T80 corresponding to a (Gaussian) grid size of 1.5° or 160 km. The 80-wave MRF model has 18 unequally spaced sigma levels in the vertical. Detailed descriptions of this model and the assimilation system are presented in Kanamitsu (1989). The operational time period of the 80-wave MRF model spans four winter and three summer seasons (9 September 1987 to 6 March 1991) using 0000 UTC analyses. The R-40 MRF model was retired in September 1987 and the T-80 MRF model was retired in March 1991.

The model configuration operating in 1994 has a triangular truncation of T126 corresponding to an equivalent horizontal resolution of 105 km. The vertical resolution of the T126 MRF model is 25 unequally spaced sigma levels. Kanamitsu et al. (1991) describe improvements in the model's orography, representation of marine stratus, introduction of a mass correction time scheme, and adjustments to gravity wave drag parameterization. Adjustments to the NMC Global Data Assimilation and forecast system, which have intractable effects upon the MRF model performance, occur almost continuously.

b. Seasonal divisions

Division of the available 80-wave runs into 3-month seasons [December–January–February (DJF), March–April–May (MAM), June–July–August (JJA), and September–October–November (SON)] yields about 300 cases for each season. When mean fields of the 1000-hPa height field are compared for each season (not shown), it is apparent that the autumn mean pattern closely resembles the winter mean pattern and that the spring and summer mean patterns are also quite similar. We can increase the number of cases in each group by dividing the year into a cool and a warm season. The 1000-hPa height field during the cool season is dominated by the Icelandic and Aleutian lows while the warm season is dominated by the oceanic subtropical highs. The Siberian high is stronger during the cool season than during the warm.

The start of the cool season is defined based on Kalnicky's (1987) study that applied factor analysis to the daily frequencies of Dzerdzeevskii's Northern Hemisphere extratropical latitude circulation types. Kalnicky found that the winterlike meridional pattern began around 23 October in most years. We choose this date as the start of our cool season and allow it to end 6 months later on 22 April of each year. The number of cases available for each season varies with integration length because some analyses used for model verification are missing and some individual forecasts are truncated before reaching date 10.

c. Spatial resolution

The only model fields available for this study are stored as pressure coefficients of spherical harmonics truncated at wavenumber 12 with a corresponding horizontal rhomboidal truncation of R12 corresponding to a (Gaussian) grid size of about 10.0° or 1100 km. At this resolution, the large-scale features of the height field can be defined anywhere on the globe with 169 pressure coefficients (0–12 zonal and 0–12 meridional). Spherical harmonic pressure coefficient files, truncated at R12, are converted into geopotential heights on a 2.5° latitude/longitude Mercator projection. The regional domain (13 × 33 grid nodes) extends from 25° to 55°N and 140° to 60°W on a 2.5° × 2.5° grid mesh (Fig. 1a) and the hemispheric domain (7 × 36 grid nodes) stretches from 20° to 80°N and 0°E to 10°W on a 10° × 10° grid mesh (Fig. 1b).

The effect of the spectral truncation on a meteorological field is demonstrated in Figs. 2a–d. Each panel of Figs. 2a–d shows full meridional wavenumber information (12 wave) with a varying number of zonal waves. The 0-wavenumber (not shown) is a zonal average at each latitude. Increasing the spectral resolution in the zonal direction produces crude features found in the NMC analysis (Fig. 2d). The 12-wave representation identifies the main features of the pressure field and compares favorably with the verification analyses.
3. Error patterns over the Northern Hemisphere
   a. Error over the period of record

   Figure 3 is a bar chart of the spatial mean bias over the hemispheric domain for each day-10 model integration at 500 and 1000 hPa. The bias is simply the summation of the forecast minus verification heights for each of the 252 grid points over the hemispheric domain. If the forecast values are consistently lower than the verification values, then the bias will have a negative sign. Days with missing forecast or verification maps are set to zero for display purposes. The biases show a marked change during 1990, especially at the 500-hPa level. Mean spatial biases are smallest during the 1990/91 winter season. This reduction in the mean bias may be attributed to improvements in the radiation scheme in the 80-wave MRF model (Caplan and White 1989) or could be associated with improvements in the flux of moisture in the model’s tropical atmosphere (van den Dool et al. 1993). There is a trend evident in the time series of the 500-hPa biases of about 10 m yr$^{-1}$. No trend is found at 1000 hPa.

   The anomaly correlation (AC) (defined in appendix A) is plotted for the Northern Hemisphere at 500 and 1000 hPa (Fig. 4). The climatology used in calculating the AC is the 15-yr mean height field for each season (DJF, MAM, JJA, SON) courtesy of the Climate Analysis Center. Unlike the spatial mean biases, the ACs show less variability over the period of record. Anomaly correlations were noticeably depressed in July 1988. Persistence anomalies (not shown) were also low in July 1988, indicating frequent regime changes over the 10-day time period. Close inspection of the time series reveals large fluctuations in the ACs on a 2–3-week timescale. A change in the AC values over the life of the 80-wave MRF model cannot be detected in the time series. No trend was found in the AC at 1000 or 500 hPa. Apparently, the model’s performance has not been greatly affected by improvements in the model’s physics over the time period from September 1987 to March 1991. Walker (1994) studied time series of ACs and teleconnection indices over the regional domain shown in Fig. 1a. A greater variability in the AC scores was found over the smaller domain with periods of very high skill followed by periods of occasionally very low skill.

b. Error over integration length

   In an effort to thoroughly assess model error, a regression model is developed according to the following equation:

   \[ E(o|f) = a + bf, \]

   where \( E(o|f) \) is the expected value of the verification height \( o \), given a particular forecast height \( f \), and \( a \) and \( b \) are least squares estimates of the regression coefficients. Each model integration produces 10 forecast
fields (days 1, 2, . . . , 10) that are compared with their corresponding verification fields. For each individual model integration, 252 points in the verification field are regressed upon 252 points in the forecast field. (Complete descriptions of the spatial statistics used in this analysis are presented in appendix A.)

Summary statistics of error as a function of integration length are presented at 1000 hPa (Table 1) and 500 hPa (Table 2). The number of error maps available for study N depends on the integration length, because some forecasts end before reaching day 10 and some verification analyses are missing. Generally, as the integration length increases, the scatter about the regression line grows, the regression slope b declines, and the intercept a increases. The error in the forecast depends systematically on f, and when b ≠ 1 the forecast is considered to be conditionally biased. That is, the error in the estimation of the verification height is conditioned on the value of the forecast height. The b term at the 1000-hPa level reflects a loss of predictability of the model forecast with an explained variance of 0.25 by day 9. The intercept a is a measure of the model's unconditional bias and is dependent on the value of the forecast. The a term grows rapidly, especially at the 500-hPa pressure level.

The bias reaches its lowest value on day 6 at 1000 hPa and then grows slowly negative. The biases are larger at the 500-hPa level and increase with integration length. The MRF forecast of the 1000- (500) hPa height field over the Northern Hemisphere appears to be "useful," as defined by Hollingsworth et al. (1980), out to about 6 (8) days. They considered the forecast to be useful as long as the root-mean-square error of the model's forecast height field was less than the standard deviation of the climatological height field. The mean square error (MSE) of the forecast does not reach the climatological level until after day 6 at 1000 hPa and after day 8 at 500 hPa. The AC stays above +0.6 beyond day 5 at 1000 hPa and beyond day 6 at 500 hPa, which indicates that the model is retaining useful information about the anomalies in the height field out to about 1 week.

The actual skill scores (SS) do not drop below zero until day 7 at 1000 hPa and until day 9 during the cool half of the year at 500 hPa. The last four columns in Tables 1 and 2 are the decomposition of SS into four component terms (appendix A) and show significant improvements when compared with the 40-wave MRF model results (Murphy and Epstein 1989). The mean height anomaly values of the forecast and the verification fields are relatively close to each other, which is quantified in the unconditional bias term (UNCO), and stay relatively small at both pressure levels. The unconditional term is proportional to the difference
between the mean anomaly of the forecast minus the mean anomaly of the verification field. Tables 1 and 2 show that the conditional bias (COND) greatly contributes to the loss of skill in the 80-wave MRF model, as also noted by Murphy and Epstein (1989, p. 580): “these biases, and the corresponding loss of skill, can be interpreted as the penalty associated with retaining meteorological features in the geopotential height field when such features are not predictable.”

4. Spatial error patterns over the contiguous United States

a. Cool season 1000 hPa

Seasonal mean error fields over the regional domain at 1000 hPa for day 4 and day 10 during the cool and warm seasons are shown in Figs. 5 and 6. During the cool season at 1000 hPa, a positive height error centered over the intermountain region appears within 24 h of model integration. This error center shifts southeastward with increasing integration length and has its largest positive amplitude 24 h after the initial state. This may indicate the “stray” phenomenon identified by Thiebaux and Morone (1990) in their study of the MRF model. Stray is a problem found in the initial-
### Table 1. The 1000-hPa model error statistics as a function of integration length for the 80- wave MRF model.
(See appendix A for the definition of terms.)

The expected value of the verification height $o$, given a particular forecast height $f$, is represented by $E(o|f) = a + bf$, where $a$ and $b$ are the regression coefficients; IL is the integration length in days, and $N$ is the number of available forecasts.

<table>
<thead>
<tr>
<th>IL</th>
<th>N</th>
<th>BIAS</th>
<th>AC</th>
<th>a</th>
<th>b</th>
<th>MSE</th>
<th>SS</th>
<th>PS</th>
<th>VVA</th>
<th>COND</th>
<th>UNCO</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>671</td>
<td>1.58</td>
<td>0.98</td>
<td>2.8</td>
<td>0.99</td>
<td>260.0</td>
<td>0.96</td>
<td>0.93</td>
<td>0.026</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>2</td>
<td>671</td>
<td>1.80</td>
<td>0.94</td>
<td>6.9</td>
<td>0.96</td>
<td>781.0</td>
<td>0.88</td>
<td>0.87</td>
<td>0.026</td>
<td>0.008</td>
<td>0.002</td>
</tr>
<tr>
<td>3</td>
<td>671</td>
<td>1.29</td>
<td>0.89</td>
<td>13.0</td>
<td>0.90</td>
<td>1498.0</td>
<td>0.77</td>
<td>0.77</td>
<td>0.026</td>
<td>0.023</td>
<td>0.004</td>
</tr>
<tr>
<td>4</td>
<td>671</td>
<td>1.01</td>
<td>0.80</td>
<td>20.8</td>
<td>0.84</td>
<td>2605.0</td>
<td>0.60</td>
<td>0.64</td>
<td>0.026</td>
<td>0.056</td>
<td>0.006</td>
</tr>
<tr>
<td>5</td>
<td>671</td>
<td>0.53</td>
<td>0.71</td>
<td>29.9</td>
<td>0.76</td>
<td>3907.0</td>
<td>0.40</td>
<td>0.50</td>
<td>0.026</td>
<td>0.117</td>
<td>0.010</td>
</tr>
<tr>
<td>6</td>
<td>671</td>
<td>-0.33</td>
<td>0.61</td>
<td>39.2</td>
<td>0.67</td>
<td>5210.0</td>
<td>0.20</td>
<td>0.39</td>
<td>0.026</td>
<td>0.200</td>
<td>0.014</td>
</tr>
<tr>
<td>7</td>
<td>666</td>
<td>-1.51</td>
<td>0.52</td>
<td>47.6</td>
<td>0.59</td>
<td>6512.0</td>
<td>0.00</td>
<td>0.29</td>
<td>0.026</td>
<td>0.294</td>
<td>0.019</td>
</tr>
<tr>
<td>8</td>
<td>666</td>
<td>-2.65</td>
<td>0.44</td>
<td>54.5</td>
<td>0.53</td>
<td>7583.0</td>
<td>-0.16</td>
<td>0.22</td>
<td>0.026</td>
<td>0.384</td>
<td>0.025</td>
</tr>
<tr>
<td>9</td>
<td>666</td>
<td>-3.45</td>
<td>0.36</td>
<td>60.7</td>
<td>0.47</td>
<td>8531.0</td>
<td>-0.31</td>
<td>0.17</td>
<td>0.026</td>
<td>0.476</td>
<td>0.031</td>
</tr>
<tr>
<td>10</td>
<td>666</td>
<td>-4.25</td>
<td>0.31</td>
<td>65.3</td>
<td>0.43</td>
<td>9182.0</td>
<td>-0.41</td>
<td>0.14</td>
<td>0.026</td>
<td>0.544</td>
<td>0.035</td>
</tr>
</tbody>
</table>

MSE of climatology = 6512.

<table>
<thead>
<tr>
<th>IL</th>
<th>N</th>
<th>BIAS</th>
<th>AC</th>
<th>a</th>
<th>b</th>
<th>MSE</th>
<th>SS</th>
<th>PS</th>
<th>VVA</th>
<th>COND</th>
<th>UNCO</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>573</td>
<td>1.35</td>
<td>0.97</td>
<td>2.7</td>
<td>0.99</td>
<td>190.0</td>
<td>0.94</td>
<td>0.91</td>
<td>0.033</td>
<td>0.003</td>
<td>0.002</td>
</tr>
<tr>
<td>2</td>
<td>570</td>
<td>1.90</td>
<td>0.92</td>
<td>7.1</td>
<td>0.95</td>
<td>538.0</td>
<td>0.83</td>
<td>0.82</td>
<td>0.033</td>
<td>0.014</td>
<td>0.005</td>
</tr>
<tr>
<td>3</td>
<td>568</td>
<td>2.10</td>
<td>0.84</td>
<td>12.9</td>
<td>0.90</td>
<td>1013.0</td>
<td>0.68</td>
<td>0.69</td>
<td>0.032</td>
<td>0.039</td>
<td>0.007</td>
</tr>
<tr>
<td>4</td>
<td>567</td>
<td>2.31</td>
<td>0.74</td>
<td>21.5</td>
<td>0.82</td>
<td>1646.0</td>
<td>0.48</td>
<td>0.54</td>
<td>0.032</td>
<td>0.087</td>
<td>0.010</td>
</tr>
<tr>
<td>5</td>
<td>566</td>
<td>2.28</td>
<td>0.62</td>
<td>31.5</td>
<td>0.73</td>
<td>2406.0</td>
<td>0.24</td>
<td>0.39</td>
<td>0.032</td>
<td>0.167</td>
<td>0.015</td>
</tr>
<tr>
<td>6</td>
<td>565</td>
<td>2.04</td>
<td>0.51</td>
<td>40.4</td>
<td>0.64</td>
<td>3134.0</td>
<td>0.01</td>
<td>0.27</td>
<td>0.032</td>
<td>0.276</td>
<td>0.019</td>
</tr>
<tr>
<td>7</td>
<td>554</td>
<td>1.56</td>
<td>0.41</td>
<td>47.7</td>
<td>0.57</td>
<td>3799.0</td>
<td>-0.20</td>
<td>0.19</td>
<td>0.032</td>
<td>0.397</td>
<td>0.023</td>
</tr>
<tr>
<td>8</td>
<td>553</td>
<td>1.03</td>
<td>0.33</td>
<td>53.3</td>
<td>0.52</td>
<td>4306.0</td>
<td>-0.36</td>
<td>0.14</td>
<td>0.031</td>
<td>0.507</td>
<td>0.026</td>
</tr>
<tr>
<td>9</td>
<td>551</td>
<td>0.54</td>
<td>0.27</td>
<td>57.5</td>
<td>0.47</td>
<td>4717.0</td>
<td>-0.49</td>
<td>0.11</td>
<td>0.032</td>
<td>0.603</td>
<td>0.030</td>
</tr>
<tr>
<td>10</td>
<td>551</td>
<td>-1.0</td>
<td>0.23</td>
<td>60.1</td>
<td>0.44</td>
<td>5066.0</td>
<td>-0.60</td>
<td>0.09</td>
<td>0.032</td>
<td>0.690</td>
<td>0.033</td>
</tr>
</tbody>
</table>

MSE of climatology = 3166.

along the lee of the Alberta Rockies. The model fails to properly simulate this process and therefore produces model lows that are deeper than those found in reality.

### b. Warm season 1000 hPa

The largest error in the 1000-hPa height field during the warm season is associated with the model’s misrepresentation of the climatological heat low located in the southwest portion of the United States (Figs. 5b and 6b). This negative error in the height field reaches $-24$ m by day 4 and grows to $-40$ m by day 10. The height fields in the MRF model are much lower than the observed height fields. An inspection of individual model integrations reveal that the MRF model characteristically overdevelops the strength of the heat lows over the southwest United States. Conversely, the strength of the high pressure systems that occur in the verification atmosphere over this area are underestimated in the MRF model.

The strength of the subtropical high over the western Atlantic is overestimated by 10 m in the day-4 forecast. This center weakens and loses its identity between day 4 and day 10. The model has a positive height error at the position of the climatological Hudson Bay polar low. Positive height errors are also found over the eastern Pacific, which have their largest magnitude at day 7 (not shown).

### c. Cool season 500 hPa

The 500-hPa height field error for the cool season (Figs. 7a and 8a) exhibits four error centers: one associated with the southward extension of the Hudson Bay polar vortex and neighboring Icelandic low, a second along the east coast of the United States, a third off the Pacific Northwest coast, and a fourth southwest of the California coast. The magnitude of the error is small during the first 3 days of integration but grows rapidly from day 4 through day 10. Caplan and White (1989) have made reference to the negative error found along the east coast of the United States. The MRF model characteristically overdevelops upper-level troughs in this region, producing an overall negative bias.

The error associated with the Hudson Bay polar vortex and neighboring Icelandic low is positive, from day 7 (not shown) through day 10, indicating that the model is underestimating the strength of these features. The third center off the northwest coast grows rapidly in magnitude between day 4 ($-24$ m) and day 10 ($-80$ m). This negative error is caused by the MRF model overestimating the strength of
the upper-level low pressure centers and is consistent with the overdevelopment of surface lows in this region (Sanders 1992). The fourth center off the southern California coast at day 4 loses its identity by day 10 (Fig. 8a) and may be caused by the model's fictitious development and persistence of cutoff lows in this region.

d. Warm season 500 hPa

The MRF model has a negative bias across virtually the entire domain during the warm season at 500 hPa (Figs. 7b and 8b). This negative bias is consistent with the daily model biases shown in Fig. 3. These biases are largest over the low-latitude oceanic regions. Analysis of the mean day-10 verification map (not presented) shows a reduction in the amplitude of the planetary-scale ridge–trough pattern of North America. Heights in the eastern North American trough are higher than those in the corresponding verification field, and the heights in the western North American ridge are lower than those found in the verification field.

5. PCA of 1000-hPa model error
   a. Rationale for PCA method

   Much of the error found in a forecast weather map is associated with a model's inability to correctly forecast the location and magnitude of the observed synoptic-scale weather patterns. Any position or strength error in the model will occur simultaneously over many grid points. A fictitious southward displacement of a synoptic-scale low pressure system will produce negative height errors to the south of the storm track and positive height errors to the north. This improper representation of such a low will produce height errors across a large number of grid points.

   We wish to study the spatial variability of the MRF model error fields over the 48 contiguous United States and surrounding waters (see Fig. 1b). Given a 1000-hPa forecast map and an accompanying verification map, we can create a 1000-hPa error map over this region by subtracting the verification heights from the forecast heights at each grid node. Across this region, the error maps will display positive and negative error patterns hundreds of kilometers in the north–south and east–west directions. A quick examination of any day will show differences in the location and magnitude of 1000-hPa height anomalies between the forecast and the verification analyses that correspond to errors in the model's prediction of high and low pressure systems over the region.

   Regions with positive errors are associated with higher heights in the model forecast than in the corresponding verification analyses. This error pattern can
be caused by the model predicting high pressure where the verification has low pressure (an error in location) or the model can underpredict the intensity of a low pressure system or overpredict the magnitude of a high pressure system (errors in magnitude). Regions with negative errors are associated with lower heights in the model forecast than in the corresponding verification analyses. The negative error patterns are also associated with errors in model location and magnitude of low and high pressure systems.

In our dataset of nearly 1000 MRF model runs, each with 10 integration lengths (day 1, day 2, . . . , day 10), approximately 10 000 different spatial error patterns are available for analysis. Each error pattern is defined at 429 grid points in the region of interest (Fig. 1b). This produces more than 4 million MRF model error values of the height field on the 1000-hPa surface. Temporal and spatial averaging of these error values, as introduced in sections 3 and 4 of this paper, simplifies the analysis of the model error but gives us no information about the spatial variability of the error.

PCA is a statistical technique used to find mutually orthogonal linear combinations of the original variables that capture most of the variability in a dataset (Manly 1986). A large part of the variance can usually be captured by the first few principal components without much loss of information. A highly intercorrelated data-

![Day 4 Error at 1000-hPa for Cool Season (Oct23-Apr22)](image)

**Fig. 5.** Mean 1000-hPa error field (meters) at day 4 for the (a) cool season and (b) warm season.

taset of $p$ variables can be reduced to one consisting of $q$ mutually orthogonal new variables. This reduction in variables provides motivation for the use of PCA as a method of analyzing the model error. PCA is used in this study to identify the primary modes of spatial variability in the MRF model error fields. The error patterns are spatially coherent, and when the error is large positive at one grid point, it is almost always large and positive at the surrounding grid points. In principle, the variation of the error at one point is comparable to the variation of the error at the neighboring points.

We define a matrix where the columns are the error values at each grid point in a spatial domain (Fig. 1b) and the rows are the error values for each map time in a season. Summation of each individual column yields a time-averaged error value for each grid point. Summation of each individual row yields a domain-averaged error for each map time. Such error matrices are defined for each integration length for the cool and warm seasons. PCA is applied to these error matrices where $p = 429$ variables or grid points in Fig. 1b and $n = 629$ (489) observations or map times in the cool (warm) season.

In a spatial PCA such as this one (in which the variables are locations in space), maps of component *loadings* are produced to depict each PC. The magnitude of a loading indicates the importance of that par-

![Day 10 Error at 1000-hPa for Cool Season (Oct23-Apr22)](image)

**Fig. 6.** Mean 1000-hPa error field (meters) at day 10 for the (a) cool season and (b) warm season.
c. Scree surfaces

Three-dimensional scree diagrams from nonrotated 1000-hPa error correlation matrices for the warm and the cool seasons, respectively, are shown in Figs. 9a,b. Eigenvalues (see appendix B) of the 1000-hPa error covariance matrices grow rapidly with increasing integration length, making intercomparison difficult. The eigenvalues of the 1000-hPa error correlation matrices vary less with integration length since each element is normalized by the product of the variable standard deviations. In the scree test (Catell 1966), principal components are arranged in descending order along the abscissa with explained variance or eigenvalue along the ordinate. The purpose of the scree test is to help select the number of new orthogonal variables (principal components) to retain in the analysis. Since the particular variable on that component. For example, if high positive loadings are present in the northeastern United States on the first PC (the one explaining the most variance), this indicates that the largest portion of the total error variance in the dataset is explained by errors over the northeastern United States. If high positive loadings are present over the southwestern United States and negative loadings are present over the Pacific Northwest, this would indicate that when height errors are positive in the Southwest they are negative in the Northwest, and vice versa.

b. Related studies

Many atmospheric scientists have used similar techniques to identify the dominant modes of spatial and temporal variability in atmospheric datasets (e.g., Kutzbach 1970; Rogers 1981; Wallace and Gutzler 1981; Volmer et al. 1984; Clinet and Martin 1992; Hewitson and Crane 1992). Arpe and Klinker (1986) have used empirical orthogonal functions (EOFs) to characterize the model forecast fields, but there are few examples in the scientific literature of PCA being applied directly to error fields in medium-range numerical models. Branstator et al. (1993) developed an empirical orthogonal function index from a covariance matrix of the 500-hPa height fields in ECMWF forecasts truncates at zonal wavenumber 6 on a hemispheric scale for 11 90-day winter seasons. They found that instantaneous forecasts integrated to day 10 could be effectively decomposed into components that were either easy or difficult to predict. Using their EOF decompositions, spatial filters were designed that remove the poorly forecasted components, subsequently improving the skill scores (as measured by an anomaly correlation) of medium-range forecasts.

Fig. 7. Mean 500-hPa error field (meters) at day 4 for the (a) cool season and (b) warm season.

Fig. 8. Mean 500-hPa error field (meters) at day 10 for the (a) cool season and (b) warm season.
PCs are ordered according to the portion of the total variance explained, the first few PCs typically explain most of the variance. Those PCs located on the tail (or scree) of the curve can be ignored since they contribute little to the total variance explained. These diagrams, derived from error correlation matrices, show how the variance explained by each component is distributed across the 10 integration lengths. (The shape of the scree surfaces indicate that the eigenvalues decline rapidly until about component 6, when the values reach asymptotically small levels.) These diagrams suggest that the first five PCs on the regional scale contain significant amounts of information about the spatial distribution of error.

The scree surfaces show an interesting property across the 10 integration lengths. The eigenvalue for the first principal component declines from day 1 through day 4 and then increases from day 7 through day 10. This error mode is located over the Great Plains region for all integration lengths for the nonrotated principal components. A detailed study of individual forecast and verification maps associated with this principal component reveals two characteristic MRF errors. Large positive errors occur when the model places high pressure where the verification atmosphere places low pressure, and large negative errors occur when the model places low pressure where the verification atmosphere has high pressure. During the cool season the most significant error center weakens and becomes dissociated with the underlying orography over the first 5 days of model integration. The number of fictitious lows in this area during the warm season rises with increasing integration length. The increase in the magnitude of the first principal component may be related to these fictitious lows.

d. Explained variances

Covariance rather than correlation matrices of the error fields are selected for spatial PCA since all of the 429 variables (each grid point in the regional domain) are measured in the same units (geopotential height). The percentage of variance explained by the first 10 PCs over the regional domain is shown for the warm and cool seasons in Tables 3 and 4, respectively. At day 3, the first five principal components explain 60% (62%) of the variance during the warm (cool) season, and at day 10 the first five principal components explain 70% (71%) of the variance during the warm (cool) season. The first five principal components, individually, explain at least 5% of the variance in the 1000-hPa error field over the regional domain.

Inspection of the first 10 principal components indicates that the 1000-hPa error is distributed across many spatial modes. We should be able to assign physical explanations for the largest error modes. Reoccurring errors in the simulation of atmospheric features should account for much of the variance in the error field. One would expect to find regions where the principal

<table>
<thead>
<tr>
<th>PC/IL</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>24</td>
<td>24</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>20</td>
<td>20</td>
<td>21</td>
<td>21</td>
<td>22</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>13</td>
<td>13</td>
<td>14</td>
<td>13</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>10</td>
<td>12</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>13</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>11</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Variance explained by each component after VARIMAX rotation

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>11</td>
<td>18</td>
<td>10</td>
<td>11</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>12</td>
<td>14</td>
<td>12</td>
<td>14</td>
<td>10</td>
<td>15</td>
<td>14</td>
<td>9</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>14</td>
<td>12</td>
<td>11</td>
<td>11</td>
<td>9</td>
<td>13</td>
<td>14</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>9</td>
<td>7</td>
<td>13</td>
<td>13</td>
<td>9</td>
<td>13</td>
<td>9</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
<td>11</td>
<td>7</td>
<td>13</td>
<td>12</td>
<td>15</td>
<td>14</td>
<td>12</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>10</td>
<td>12</td>
<td>8</td>
<td>12</td>
<td>12</td>
<td>9</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
</tbody>
</table>
components dominate the spatial error pattern. Each principal component represents different patterns in the variance of the 1000-hPa error fields. Referring to appendix B we see that the variance of each principal component is the summation of the 429 original variables (grid points) each multiplied by a constant. The value of the constants, for each principal component, varies across the 429 variables (grid points). A “loading map” for each principal component is produced by plotting these constants on the 429 grid points (original variables). Figures 10a–p display the location and magnitude of each of the principal components.

Buell (1979) believed that the domain shape had an implicit effect on the resulting component loadings. For a square domain, such as this one, Buell indicated that the first component would have positive loadings over the entire domain, the second component would have positive and negative centers to the north and south, the third component would have a similar bipolar in the east-west direction, etc. Buell recommended rotating the original PCs to avoid this problem. Our unrotated PCA solution had these characteristics (Walker 1994). The principal component 1 pattern looked like Buell pattern 1, with a weak center covering much of the domain. The principal component 2 pattern tended to show a bipolar north-south flip-flop (Buell pattern 2). The principal component 3 pattern tended to look bipolar in the east-west direction (Buell pattern 3), and the principal component 4 pattern had a plus-minus-plus shape to it that looked like Buell pattern 4. We therefore chose to use our rotated solutions in this presentation, since these depicted regional modes of error variance that would be most useful to forecasters. However, it is important to acknowledge that there is considerable debate on the value and appropriateness of rotating principal components, and we refer interested readers to Legates (1991) and to a subsequent comment and reply (Richman 1993; Legates 1993).

Richman (1986) and others argue that spatial patterns over a generally rectangular domain may have little meaning unless they are first rotated. The VARIMAX orthogonal rotation is used to rotate the first six principal component axes over the regional domain in such a way that the variance explained by each of the axes is maximized. This rotation of the PCs aids in the interpretation of each of the error patterns. The explained variance after rotation is shown at the bottom of Tables 3 and 4 for both seasons. The amount of variance explained by each PC is exchanged between the other PCs. Hewitson and Crane (1992) use this type of orthogonal rotation on sea level pressure maps over an area similar to this regional domain. In their study, each component also tends to show a single center located in a different part of the spatial domain.

6. Error patterns of the largest principal components

a. Analysis procedure for each error pattern

Loading maps for the cool and warm seasons over the 10 integration lengths in the regional domain were

FIG. 10. Mean of the ten 1000-hPa that score highly on Great Plains pattern (a) forecast maps and (b) verification maps.
examined to identify recurring patterns that were tied to physical processes not correctly simulated by the MRF model. Inspection of over 100 maps (10 PCs by 10 integration lengths) lead to the identification of seven error patterns, among the first 5 PCs, that occur in both seasons and two additional patterns present in the cool season (Table 5). Loading maps for the rotated PCs, at mostly day 7, are shown in Figs. 10a–p. Day 7 loading maps are selected because they fall in the middle of the 3–10-day time period; the largest error modes are found across most of the integration lengths. The Great Plains (GP) and Alberta Rockies (AR) error patterns are present through all integration lengths. The Great Lakes (GL), East Coast (EC), and West Coast (WC) error patterns are found at most integration lengths. The New Brunswick (NB), California coast (CC), Manitoba (MA), and Great Basin (GB) error patterns are more seasonal and occur less frequently.

b. Interpretation of the individual spatial error patterns

Principal component scores represent the projection of a point from a principal component axis onto the original axes. The coordinates of any point on the component axes are equal to the coordinates of that point measured on the original axes multiplied by a matrix of constants (which contains the eigenvectors). For each component, the scores have a mean of zero and a variance equal to the eigenvalue (which is proportional to the variance explained by the PC). In this analysis, a component score is calculated for each map time (day) in each PCA. Thus, a day with a large positive score (relative to the other scores) on component 1 indicates that the error pattern present on that day is highly correlated with the positive component one loading map.

Each principal component pattern represents regions where the spatial variance in the 1000-hPa error field is large. We should be able to identify improperly handled synoptic-scale features in the MRF model that are contributing to these large errors. The error patterns found in each forecast of the 629 cool season or 489 warm season cases will project (or score) on one or more of the different principal components. The individual cases that score highly on each principal component have been gathered and their characteristic MRF model errors are summarized below.

Error patterns associated with the highest and lowest scores for the principal components are not shown, because they have shapes similar to the loading maps and do not provide much additional information. Forecast and verification maps of the 10 highest and 10 lowest scores for the PCs give insight into the causes

<table>
<thead>
<tr>
<th>Component</th>
<th>Center of action</th>
<th>Cool season percent of time found within first five PCs</th>
<th>Warm season percent of time found within first five PCs</th>
</tr>
</thead>
<tbody>
<tr>
<td>GP</td>
<td>Great Plains</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>AR</td>
<td>Alberta Rockies</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>GL</td>
<td>Great Lakes</td>
<td>100</td>
<td>90</td>
</tr>
<tr>
<td>EC</td>
<td>East Coast</td>
<td>60</td>
<td>80</td>
</tr>
<tr>
<td>WC</td>
<td>West Coast</td>
<td>40</td>
<td>80</td>
</tr>
<tr>
<td>NB</td>
<td>New Brunswick</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td>CC</td>
<td>California coast</td>
<td>40</td>
<td>10</td>
</tr>
<tr>
<td>MA</td>
<td>Manitoba</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>GB</td>
<td>Great Basin</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 5. Distribution of error patterns across the first five principal components by each integration length for rotated PCs over the regional domain.

<table>
<thead>
<tr>
<th>PC/IL</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cool season</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GP</td>
<td>AR</td>
<td>GP</td>
<td>GP</td>
<td>GP</td>
<td>GP</td>
<td>GP</td>
<td>GP</td>
<td>GP</td>
<td>AR</td>
<td>GP</td>
</tr>
<tr>
<td>AR</td>
<td>GP</td>
<td>AR</td>
<td>AR</td>
<td>MA</td>
<td>AR</td>
<td>AR</td>
<td>AR</td>
<td>AR</td>
<td>GP</td>
<td>AR</td>
</tr>
<tr>
<td>GB</td>
<td>EC</td>
<td>EC</td>
<td>EC</td>
<td>AR</td>
<td>MA</td>
<td>CC</td>
<td>GL</td>
<td>WC</td>
<td>EC</td>
<td>WC</td>
</tr>
<tr>
<td>GL</td>
<td>CC</td>
<td>GL</td>
<td>GL</td>
<td>GL</td>
<td>GL</td>
<td>EC</td>
<td>GL</td>
<td>GL</td>
<td>GL</td>
<td>GL</td>
</tr>
<tr>
<td>CC</td>
<td>GL</td>
<td>NB</td>
<td>NB</td>
<td>NB</td>
<td>WC</td>
<td>WC</td>
<td>EC</td>
<td>CC</td>
<td>EC</td>
<td>EC</td>
</tr>
<tr>
<td>Warm season</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GP</td>
<td>AR</td>
<td>GP</td>
<td>GP</td>
<td>GP</td>
<td>AR</td>
<td>GP</td>
<td>GP</td>
<td>GP</td>
<td>GP</td>
<td>GP</td>
</tr>
<tr>
<td>EC</td>
<td>GL</td>
<td>GP</td>
<td>GP</td>
<td>GP</td>
<td>AR</td>
<td>GP</td>
<td>GP</td>
<td>GP</td>
<td>GP</td>
<td>GP</td>
</tr>
<tr>
<td>CC</td>
<td>AR</td>
<td>GL</td>
<td>GL</td>
<td>WC</td>
<td>WC</td>
<td>GP</td>
<td>WC</td>
<td>WC</td>
<td>AR</td>
<td>EC</td>
</tr>
<tr>
<td>GL</td>
<td>EC</td>
<td>WC</td>
<td>WC</td>
<td>GL</td>
<td>EC</td>
<td>WC</td>
<td>GL</td>
<td>EC</td>
<td>WC</td>
<td>WC</td>
</tr>
<tr>
<td>AR</td>
<td>NB</td>
<td>EC</td>
<td>NC</td>
<td>GL</td>
<td>NB</td>
<td>GL</td>
<td>NB</td>
<td>GL</td>
<td>GL</td>
<td>GL</td>
</tr>
</tbody>
</table>

Unauthenticated | Downloaded 09/25/23 11:44 AM UTC
of MRF model error patterns. Generally, high scores are associated with the model, placing high pressure where the verification atmosphere has low pressure, and low scores are associated with the model, placing low pressure where the verification atmosphere has high pressure. This is the case regardless of season, integration length, or PC.

Composite forecast and verification maps are only shown for the Great Plains pattern during the cool season. Individual forecast and verification maps have been studied but will not be shown. Low (high) composite scores for the forecast (verification) maps are very similar to high (low) composite scores for the verification (forecast) maps. The strength of the pressure systems in the verification maps are almost always greater than those found in the corresponding forecast maps. For example, Fig. 10a has a very similar shape to Fig. 11b, but the 1000-mb composite verification low is 40 m deeper than the corresponding composite forecast low. This discrepancy in the strength of the pressure systems is contributing to the overall magnitude of each error pattern.

1) GREAT PLAINS PATTERN (GP)

The Great Plains error pattern (Figs. 12a, b) is centered near the Texas–Oklahoma region in the cool season for all integration lengths. The error mode shifts north and west (closer to the lee of the Rocky Mountains) for day 1, day 2, and day 3 during the warm season, and then the pattern returns to the Texas area for the other integration lengths. This error center is located downstream from the climatological heat low over the desert Southwest during the warm season. Low scores for the GP pattern are commonly associated with the model’s fictitious displacement of the climatological heat low (also present but weaker during the cool season) northward into the Great Basin. High scores for the Great Plains pattern often result from the fictitious displacement of the Pacific subtropical high eastward into the Great Basin. A study of the forecast and verification maps that load highly positive on the GP pattern reveals that in these situations the model is missing cyclogenesis observed in the atmosphere.

2) ALBERTA ROCKIES PATTERN (AR)

The Alberta Rockies error pattern (Figs. 12c, d) migrates from the Washington coast in the first 3 days to the lee of the Alberta Rockies for day 4 through day 10. This is the region in which the verification atmosphere dissipates lows approaching from the Gulf of Alaska and redvelops them along the lee of the Alberta Rockies. The MRF model has difficulty simulating this dissipation and redevelopment process (Caplan and White 1989). The warm season exhibits a similar shift in the location of this center, and it is shifted closer to the lee of the Alberta Rockies. Individual cases with low scores on the AR pattern are often associated with the model’s inability to adequately dissipate low pressure systems as they cross the Rocky Mountains. The polar anticyclone often found in this region dissipates and moves eastward too quickly in the model atmosphere. The low pressure systems in the verification atmosphere over the Gulf of Alaska are nearly 60 m deeper than those found in the model atmosphere. High scores of the AR pattern are often associated with an absence of a model-developed low pressure trough in the Canadian plains.

3) GREAT LAKES PATTERN (GL)

The Great Lakes error pattern (Figs. 12e, f) for the cool season is shifted about 5° farther south and east than during the warm season. Storms track across this region every month of the year. The error center meanders on day 1 and day 3 during the cool season; otherwise, it is a very consistent feature throughout the different integration lengths. Low and high scores of the GL pattern during the cool season show forecast low pressure systems developing in the model and not in the verification analyses, and vice versa.

The Atlantic subtropical high appears to play a more important role during the warm season. The warm
Fig. 12. Loading maps of largest principal components after VARIMAX rotation for (a) Great Plains cool, (b) Great Plains warm, (c) Alberta Rockies cool, (d) Alberta Rockies warm, (e) Great Lakes cool, (f) Great Lakes warm, (g) East Coast cool, (h) East Coast warm, (i) West Coast cool, (j) West Coast warm, (k) New Brunswick cool, (l) New Brunswick warm, (m) California coast cool, (n) California coast warm, (o) Manitoba cool, and (p) Great Basin cool.
Fig. 12. (Continued)
season loading map exhibits a strong bimodal pattern (Fig. 12f). This indicates that the direction of the errors varies significantly between the Great Lakes and the Bermuda high. Thus, when positive height errors are present over the Great Lakes, negative errors occur simultaneously over the Bermuda high, and vice versa.

4) EAST COAST PATTERN (EC)

The East Coast error pattern (Figs. 12g,h) has one large positive center off the mid-Atlantic coast for all integration lengths during the cool season. The error pattern shifts north and west during the warm season. The warm season shows, with the higher integration lengths, an additional weaker negative center west of the Great Lakes. Large positive errors associated with the underdevelopment of coastal lows in the model atmosphere are accompanied by the underdevelopment of highs over the Great Lakes. The location of cyclogenesis along the East Coast shifts with season for the East Coast error pattern. The cool season focuses its cyclogenetic activity over the warmer ocean waters. The warm season focuses its cyclogenetic activity more over the eastern Great Lakes. The model has problems near these climatological locations of cyclogenesis. The model often correctly develops a low pressure system but the development occurs a day or two later than the development found in the verification fields. The model also characteristically underestimates the strength of the Atlantic subtropical high during the warm season.

5) WEST COAST PATTERN (WC)

The West Coast error pattern (Figs. 12i,j) is centered off the Washington–Oregon coast for all integration lengths. This is in a region where numerous low pressure systems cross the Pacific and approach the mountainous western United States. Improper dissipation of model lows is often a problem for the model in this area. The MRF model tends to show centers increasingly deeper with integration length over the eastern Pacific (Sanders 1992). Low scores for the WC pattern occur with low pressure in the model atmosphere when the verification atmosphere has high pressure. The fictitious eastward displacement of the Pacific subtropical high occurs with high scores of the WC pattern. The model frequently shows low pressure development in the Gulf of Alaska but it often occurs later than that found in the verification maps.

6) NEW BRUNSWICK PATTERN (NB)

The New Brunswick pattern (Figs. 12k,l) is seen within the first five PCs from day 2 through day 7. This loading pattern has a wavelike structure across the regional domain during the cool season and is associated with rapidly moving low pressure systems in the westerlies. The loading pattern is confined to the East Coast during the warm season when the westerlies are weaker. Low scores of the NB pattern are associated with the slow and weak development of low pressure in the model atmosphere as systems move toward the Canadian Maritimes. High scores of the NB pattern are often the result of the model completely missing the development of low pressure systems off the Canadian Maritimes.

7) CALIFORNIA COAST PATTERN (CC)

The California coast pattern (Figs. 12m,n) is identified within the first five PCs primarily during the cool season. This error pattern is located in the region of the climatological Pacific subtropical high. The MRF model has a tendency to overdevelop and displace the subtropical high in its simulations. The CC pattern for day-1 forecast shows strikingly similar patterns for both high and low scores. This pattern has a loss of height in the Pacific subtropical high of approximately 20 m after only 24 h of model integration. High score maps show the fictitious displacement of the subtropical high into the Great Basin region.

8) MANITOBA PATTERN (MA)

The Manitoba error pattern (Fig. 12o) is identified during the cool season for integration lengths day 3 and day 4 only. Low scores of the Manitoba pattern during the cool season for day-4 forecasts show that the model is overdeveloping low pressure centers along the lee of the Alberta Rocky Mountains and fictitiously displacing the cold-core anticyclones southward into the Mississippi Valley. High scores for the Manitoba pattern appears to be associated with the overdevelopment of low pressure in the Gulf of Alaska at the expense of low pressure over Manitoba.

9) GREAT BASIN PATTERN (GB)

The Great Basin error pattern (Fig. 12p) is found only for day-1 forecasts. Individual forecast maps that have low scores of the Great Basin pattern fictitiously deepen of the climatological desert Southwest heat low after 24 h of model integration. High scores of the Great Basin pattern show the unrealistic displacement of the Pacific subtropical high eastward into the Great Basin region.

c. Summary

A careful study of the forecast and verification maps that contribute significantly to the different spatial error modes has been undertaken. Identification of the specific behavioral problems in the 80-wave MRF model that produce these spatial error patterns is difficult. Distinguishing the difference between “initial condition” and “model-dependent” errors is beyond the scope of this paper. The weather forecaster, who uses medium-range numerical guidance, should recognize
that the largest model spatial errors will most likely be related to the model's improper representation of the location and intensity of the synoptic-scale systems. This analysis demonstrates that the largest error modes are found over all 10 integration lengths, for both seasons. These model spatial errors are associated with the very features the forecaster is attempting to resolve. The model guidance on the 3–10-day timescale must be consulted cautiously.

7. Discussion and conclusions

The demand for deterministic forecasts, over the 3–10-day timescale, continues to grow. The operational medium-range forecast models, which produce estimates of temperature and precipitation, have become major players in producing such forecasts. This paper attempts to summarize the error statistics of the 80-wave version of the NMC MRF model. Comparison of this model with the earlier 40-wave MRF model shows a measurable but small improvement in model performance.

The mean biases over the Northern Hemisphere have decreased somewhat over time at the 500-hPa level. Potential skill, as defined by the square of the anomaly correlation, has not risen significantly over this 3.5-yr period, suggesting that the day-to-day skill of the model has not changed much over the period of record. Tables 1a,b and 2a,b show that skill scores are marginal in the 6–10-day time range. Limited faith should be placed in the validity of the model output at these integration lengths. The prediction of the height field at 12-wave resolution has higher skill scores at the 500-hPa level than closer to the earth's surface. Figures 4a,b show large week-to-week variability in model skill at day 10. Subsynoptic-scale features at this length of model integration should not be viewed deterministically.

Maps of the time-mean 1000-hPa height errors over the United States at day 4 and day 10 show distinctly different seasonal patterns cool and warm seasons. The 1000-hPa height errors in the warm season are centered over the southwestern United States, while the largest error centers during the cool season are in the general proximity of the climatological Icelandic and Aleutian polar lows. The height error centers at the 500- and 1000-hPa levels shift positions in the cool season but are more stationary during the warm season. Figures 5b and 6b caution the forecaster to expect magnitude errors over the desert Southwest during the cool season.

Principal components analysis has been performed directly on the 1000-hPa MRF model error fields. The dominate modes of error have been identified over the contiguous United States. Results from this analysis suggest that the error patterns are spatially coherent. Greater than 60% of the spatial error variance over this domain can be explained by the first five principal components. The forecast and verification maps associated with each error mode confirm that the errors are associated with the improper positioning of model forecasted synoptic-scale features. Figures 10a,b and 11a,b indicate that position errors account for the largest variance in the 1000-hPa height fields. When forecasters are viewing model features they should assume there is an error in the features' locations. The forecaster should consider adjusting the position of the model feature using whatever information is available. Even if there is no position error in the model feature, the forecaster should watch for errors in the magnitude of the model systems.

The major modes of spatial variation in the 1000-hPa height errors are associated with the climatologically persistent atmospheric circulation systems (Pacific and Atlantic subtropical highs, desert Southwest heat low), the orographic features within the domain (Alberta and Colorado mountains), and the active storm tracks regions (East Coast and the United States–Canadian border). The error patterns suggest that much of the model error is associated with the improper representation of the atmosphere's response to orographic forcing. The model low pressure systems play an important part in virtually all these major modes of 1000-hPa MRF error.

Grum (1993) suggests that if the aviation model forecasts a strong cyclone event out to 72 h, there is a high probability the cyclone will occur. The MRF model (which is the same as the aviation model with a different data cutoff) cannot predict the exact location of the low pressure systems but is successful in predicting the presence or absence of most lows near their verification locations. A forthcoming paper will show how information about the presence or absence of a low pressure system in a model forecast can be used to estimate the error associated with the low pressure system. Information about model error can then be used to adjust the model, enabling one to improve the correspondence between forecast and verification fields (Walker 1994).

Acknowledgments. We thank John (Clay) Davenport for the development of Fortran code to create PC McIDAS grid files for graphical display over the contiguous United States. We are in debt to Dr. Steven Tracton and the Climate Analysis Center at the National Meteorological Center for the pressure coefficient files of the 80-wave MRF model, and extend deep appreciation to Dr. Michael Garstang for a careful reading of an earlier version of this manuscript.

APPENDIX A

Spatial Statistics for Model Verification

\[ \langle \rangle \quad \text{Mean over } n \text{ grid points} \\
\hat{f}_i \quad \text{Forecast height at grid point } i \\
\hat{a}_i \quad \text{Verification height at grid point } i
\[ c_i = f_i - c_i \quad \text{Climatological height at grid point } i \]
\[ f'_i = f_i - c_i \quad \text{Forecast height anomaly at grid point } i \]
\[ o'_i = o_i - c_i \quad \text{Verification height anomaly at grid point } i \]

The mean value of the forecast height anomaly is
\[ \langle f' \rangle = \frac{1}{n} \sum_{i=1}^{n} (f_i - c_i). \]

The mean value of the verification height anomaly is
\[ \langle o' \rangle = \frac{1}{n} \sum_{i=1}^{n} (o_i - c_i). \]

The variance of the forecast height anomaly is
\[ s^2_{f'} = (\langle f'^2 \rangle - \langle f' \rangle^2). \]

The variance of the verification height anomaly is
\[ s^2_{o'} = (\langle o'^2 \rangle - \langle o' \rangle^2). \]

The covariance between the forecast and verification height anomalies is
\[ s_{f'o'} = \langle f'o' \rangle - \langle f' \rangle \langle o' \rangle. \]

The anomaly correlation is
\[ AC = r_{f'o'} = \frac{s_{f'o'}}{s_{f'} s_{o'}}. \]

The spatial bias of the forecast is
\[ BIAS = \frac{1}{n} \sum_{i=1}^{n} (f_i - o_i). \]

The mean-square error of the forecast is
\[ MSE = \frac{1}{n} \sum_{i=1}^{n} (f_i - o_i)^2. \]

The mean-square error of the climatology is
\[ MSC = \frac{1}{n} \sum_{i=1}^{n} (c_i - o_i)^2. \]

The skill score and Eq. (13) from Murphy and Epstein (1989) is
\[ SS = \left( 1 - \frac{MSE}{MSC} \right) = PS + VVA - COND - UNCO. \]

The potential skill is
\[ PS = \frac{(r_{f'o'})^2}{[1 + (\langle o' \rangle/s_{o'})^2]}. \]

The variation of the verification anomalies is
\[ VVA = \frac{[\langle o' \rangle/s_{o'}]^2}{[1 + (\langle o' \rangle/s_{o'})^2]}. \]

The conditional bias of the forecast is
\[ COND = \frac{[r_{f'o'} - (s_{f'}/s_{o'})]^2}{[1 + (\langle o' \rangle/s_{o'})^2]}. \]

The unconditional bias of the forecast is
\[ UNCO = \frac{[\langle f' \rangle - \langle o' \rangle]/s_{o'}]^2}{[1 + (\langle o' \rangle/s_{o'})^2]}. \]

**APPENDIX B**

**Principal Components Analysis**

A principal component analysis starts with data on \( p \) variables for \( n \) observations as indicated below:

\[
\begin{align*}
X_{11} & \quad X_{12} & \cdots & \quad X_{1p} \\
X_{21} & \quad X_{22} & \cdots & \quad X_{2p} \\
\cdots & \quad \cdots & \cdots & \quad \cdots \\
X_{n1} & \quad X_{n2} & \cdots & \quad X_{np} 
\end{align*}
\]

The mean of variable \( X_j \) is given by
\[ \bar{X}_j = \frac{1}{n} \sum_{i=1}^{n} X_{ij}. \]

The sample variance is
\[ S_{j} = \frac{1}{n-1} \sum_{i=1}^{n} (X_{ij} - \bar{X}_j)^2. \]

The sample covariance between variables \( j \) and \( k \) is defined as
\[ C_{jk} = \frac{1}{n-1} \sum_{i=1}^{n} (X_{ij} - \bar{X}_j) \ast (X_{ik} - \bar{X}_k). \]

The covariance matrix \( \mathbf{C} \) of \( p \) variables \( X_1, X_2, \ldots, X_p \) is
\[
\mathbf{C} = \begin{bmatrix}
c_{11} & c_{12} & \cdots & c_{1p} \\
c_{21} & c_{22} & \cdots & c_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
c_{p1} & c_{p2} & \cdots & c_{pp}
\end{bmatrix}
\]

The diagonal element \( c_{ii} \) is the variance of \( X_i \) and \( c_{ij} \) is the covariance of variables \( X_i \) and \( X_j \). The variances of the principal components are the eigenvalues of the matrix \( \mathbf{C} \). There are \( p \) of these, some of which may be approximately zero. Assuming that the eigenvalues are ordered as \( \lambda_1 > \lambda_2 > \ldots > \lambda_p \), then \( \lambda_i \) corresponds to the \( i \)th principal component
\[ Z_i = a_{i1} X_1 + a_{i2} X_2 + \cdots + a_{ip} X_p. \]

---

1 The 15-year Climate Diagnostics Data Base climatology (1979–1993) from NMC’s Climate Prediction Center.
The variance of $Z_i$ is equal to $\lambda_i$, and the constants $ai1, ai2, \ldots, aip$ are the elements of the corresponding eigenvector. An important property of the eigenvalues is that they add up to the sum of the diagonal element of $C$:

$$\lambda_1 + \lambda_2 + \cdots + \lambda_p = c11 + c22 + \cdots + cpp.$$ 

Since $cii$ is the variance of $X_i$ and $\lambda_i$ is the variance of $Z_i$; this means that the sum of the variances of the principal components is equal to the sum of the variances of the original variables. Therefore, the principal components account for all of the variation in the original data. SAS was used to perform the multivariate statistical analysis.

REFERENCES


