An Assessment of the Quality of Forecast Trajectories

BARBARA J. B. STUNDER

NOAA Air Resources Laboratory, Silver Spring, Maryland

(Manuscript received 17 July 1995, in final form 17 January 1996)

ABSTRACT

Forecast and “analysis” (reference) trajectories were computed from six sites over North America at three altitudes (500, 1000, and 1500 m above ground) twice a day for a one-year period using Nested Grid Model wind fields. The reference meteorology was a series of short-term forecasts. Absolute error (distance between reference and forecast trajectory), relative error (absolute error divided by forecast trajectory travel distance), and the angle between the reference and forecast trajectory were also computed. The mean relative error for all the forecast trajectories for a travel time of 36 h is about 35%; the 90th percentile of the relative error is about 65%. The forecast is slightly biased to the left of the reference early in the forecast period. Absolute error and travel distance both are larger in winter than summer, so that the relative error is generally constant throughout the year. Differences in mean error among the three starting altitudes, among the six origin sites, and between the two origin times are insignificant when compared to the variation in errors for a collection of trajectories at a given origin. The forecast trajectories were objectively classified through a cluster analysis, which groups trajectories by direction and travel distance. For all clusters, by season, origin site, and altitude, differences between the minimum and maximum cluster-mean relative errors were about a factor of 2–3. Individual forecast trajectories composing clusters with the minimum relative error (about 20%) tended to originate within stronger, steady flow either ahead of or behind a cold front. Maximum relative error (about 45%) was associated with forecast trajectories originating in regions of generally slow wind fields such as under a high pressure system or near stationary or slowly moving fronts.

1. Introduction

Air parcel trajectories have been extensively used over the past decade or so to study atmospheric transport because they are relatively simple to compute. Backward trajectories from a receptor are commonly used to identify air pollution source regions. Forecast trajectories have been used to position aircraft for tracer sampling (e.g., Stunder and Draxler 1989; Kahl et al. 1991) and to predict the downwind location of volcanic ash (Heffter et al. 1990). However, an individual trajectory gives only a general description of the flow field because it does not account for atmospheric processes such as vertical and horizontal mixing and diffusion. Kahl et al. (1989) applied a method described by Merrill et al. (1985) to assess the representativeness of individual trajectories by computing an array of trajectories in a region centered over the trajectory origin of interest. If the resulting set of trajectories were similar, then the center trajectory probably gave a fairly accurate depiction of the flow.

Dispersion models, which include atmospheric mixing processes and may have the same advection algo-

Corresponding author address: Barbara J. B. Stunder, NOAA Air Resources Laboratory (R/E/AR), SSMC3, Room 3151, 1315 East West Highway, Silver Spring, MD 20910.
E-mail: barbara@arrisc.ssmc.noaa.gov
tracer measurements are often difficult because of limited spatial density and temporal averaging associated with most measurements. Reference trajectories may also be computed using rawinsonde observations (e.g., Draxler 1991). An alternative approach is to use meteorological model data as the reference. Short-term forecast data with the best available resolution have been used in sensitivity studies of trajectory calculations (e.g., Rolph and Draxler 1990). Analyzed winds output from a meteorological model such as the National Centers for Environmental Prediction (NCEP, formerly the National Meteorological Center) Global Data Assimilation System [GDAS, available from the National Climatic Data Center (NCDC, TD-6140)] may also be used to compute trajectories. An investigation of the differences between "reference" trajectories and actual trajectories is beyond the scope of the present study.

Among many factors contributing to the forecast trajectory error, the quality of the meteorological forecast makes perhaps the largest contribution. Other factors include how well the gridded data represent (in terms of the horizontal and vertical grid spacing and the temporal frequency) the state of the atmosphere, the horizontal advection scheme, and the vertical transport algorithm or assumption.

Few forecast trajectory evaluations have been reported. Maryon and Heasman (1988) studied the accuracy of forecast trajectories using a one-year collection of trajectories and found mean absolute errors (distance between reference and forecast) of about 250 km at +36 h. Monthly mean absolute error and trajectory travel distance were both generally low in the summer and high in the autumn and winter. Errors were larger when forecast trajectories originated later in the forecast. Haagenson et al. (1990) found higher forecast trajectory error (400 km day$^{-1}$), which they attributed to an overprediction of winds, as compared to the analysis trajectory error (200 km day$^{-1}$) for a collection of winter trajectories based on tracer gas measurements.

Boundary layer trajectory errors of about 25% of the travel distance are common when various "analysis" trajectories based upon tracer measurements and gridded analyzed data of various resolutions are compared (Draxler 1991; Rolph and Draxler 1990; Haagenson et al. 1990). This is equivalent to an absolute error of about 300 km after a travel time of 36 h with a 10 m s$^{-1}$ wind. Typically, the absolute error increases nearly linearly with time. Rolph and Draxler (1990), who found errors of around 15% at +36 h (travel time) with 6-h data, mention many studies showing relationships between trajectory error and the temporal and spatial resolution of the meteorology data. Draxler (1991) showed that trajectories based on four-per-day rawinsondes, rather than the standard two-per-day, gave better results when fronts or low pressure systems were present. Doty and Perkey (1993) examined trajectories in the vicinity of an intense extratropical cyclone using data of various temporal resolutions from 15 min through 12 h. They found small errors (75 km or less at +36 h) with the high-resolution data (15 min to 1 h) and much higher errors (100–500 km) with 3-h data. Regarding the spatial density of data, McQueen and Draxler (1994) found less error with a higher data density.

Trajectory error also may be related to synoptic conditions. Rolph and Draxler (1990) found nearly constant relative error (absolute distance error divided by travel distance), but larger absolute error during cyclonic conditions with strong winds as compared to anticyclonic conditions. McQueen and Draxler (1994) found low relative error (15%) with persistently strong flow. Heffter et al. (1990) and Haagenson et al. (1990) also suggest that relative error is inversely proportional to wind speed.

Trajectory error due to the horizontal advection scheme is generally small compared to the overall error (e.g., Maryon and Heasman 1988). However, Seibert (1993) says that with small uncertainties in the wind field, the numerical error arising from the advection scheme makes a relatively larger contribution to the overall error. Seibert also notes that with current computer capabilities, the numerical error arising from the advection scheme should be limited to about one order of magnitude less than the total uncertainty; this may be accomplished with a sufficiently small time step (typically about 0.5 h). An investigation of numerical error is beyond the scope of the present study.

The purpose of this study is to estimate the forecast trajectory error by comparing trajectories calculated using two meteorology datasets, a forecast and an approximation to the analysis, and to relate the error to synoptic patterns. Knowledge of typical errors should help the interpretation of operational forecasts. The trajectory model and the meteorology data are described in section 2. The error statistics are defined in section 3 and the overall results presented in section 4. Results from cluster analysis, which groups similar trajectories, are given in section 5. Finally, section 6 contains the summary and conclusions.

2. Trajectory model

Three-dimensional trajectories were calculated using the Hybrid Single-Particle Lagrangian Integrated Trajectories Model (HY-SPLITT, Draxler 1992) with a 1-h advection time step (the smallest time step allowed). The model advection of a particle from its initial position $X_0$ to a downwind position $X_2$ over a time step $\Delta t$ is computed as

$$X_2 = X_0 + 0.5(V_0 + V_1)\Delta t,$$

where $V_0$ is the horizontal wind vector at position $X_0$ and $V_1$ is the wind vector at position $X_1 = X_0 + V_0\Delta t$. If the particle is between two meteorological data levels, the horizontal advection is computed on those two
levels, with the final horizontal position at pressure level $P_0$ being the distance weighted average of the two $X_2$ values. The new vertical position of the particle is

$$P_2 = P_0 + 0.5(W_0 + W_1)\Delta t,$$

where $P$ is pressure, $W$ is vertical velocity, and the subscripts are the same as above.

Forecast and analysis (reference) 48-h trajectories were computed twice a day (beginning at 0000 UTC and 1200 UTC) for the period June 1993–May 1994 originating at three levels (500, 1000, and 1500 m above ground) from six sites chosen on a geographic grid over North America (Fig. 1). The sites are identified as A on the Montana–Wyoming border, B in Minnesota, C in Ontario, D in Arizona, E in Oklahoma, and F on the North Carolina–South Carolina border. Gridded meteorological data fields used to create both the forecast and reference trajectories are from the NCEP Nested Grid Model (NGM; Phillips 1979; Hoke et al. 1989).

a. Forecast trajectories

The forecast trajectories were computed operationally after the latest NGM forecast data were received from NCEP. The data extended from the initialization time (0000 UTC or 1200 UTC) to +48 h at 6-h intervals. There were five sigma ($\sigma$) levels from near the surface ($\sigma = 0.982$) to near 800 hPa ($\sigma = 0.785$) and the horizontal grid spacing was 190.5 km ($\sigma = P/P_0$, where $P$ is the pressure and $P_0$ is the surface pressure at the model terrain). The number of forecast trajectories for the year and by season are summarized in Table 1. Winter is December–February, spring is March–May, summer is June–August, and autumn is September–November. There are no trajectories for days when the forecast NGM data were not received properly from NCEP or there was some other problem with the model run. Another complication with the forecast trajectory database is that about half the trajectories terminated prematurely (i.e., before 48 h) when they reached one of the meteorological grid boundaries, most commonly the top. Those trajectories tended to originate from the sites in the mountains (A and D) or from the higher altitudes at all the sites. The percentage of forecast trajectories reaching each grid boundary over the year are summarized in Table 2.

b. Reference trajectories

The reference trajectories were calculated using NGM data (the “NGM archive,” Rolph et al. 1993; Draxler 1992) archived at NCDC (TD-6140). The archive contains three-dimensional fields (the three wind components, temperature, and specific humidity) and single-level fields (mean sea level pressure, precipitation, surface fluxes, etc.). The data are a series of 2-h forecasts from 2 to 12 h after the NGM initialization time. The 12-h forecast fields from the previous NGM run are used instead of the analysis because the precipitation and other diagnostic fields are available for each forecast hour, but not at the initialization time. There are 10 sigma levels, the same five as for the forecast data, plus five more extending up to near 400 hPa ($\sigma = 0.434$). The horizontal grid resolution, 182.9 km, is slightly smaller than the resolution of the forecast data because forecast fields are mapped to a different grid than the archive. The data, a series of high temporal resolution short-term forecast data, approximate the analyses. Rolph and Draxler (1990) and Maryon and Heasman (1988) have also considered short-term forecasts to be the reference. Analyzed data from NCEP’s GDAS (see section 1) are not used because of their coarse spatial (381 km) and vertical (mandatory pressure levels) resolution.

It is important to understand the meteorology used to calculate the trajectories because the forecast trajectory error (see section 3) is primarily a result of the difference between the forecast and reference fields. A comparison of the times of meteorological model fields for example 48-h forecast and reference trajectories beginning at 0000 UTC are given in Fig. 2. Note that the trajectories used in this study begin at 1200 UTC as well. At 0000 UTC day 15, the data for the forecast trajectory are the current forecast initialization (the forecast hour is +0), but the data for the reference trajectory are a 12-h forecast from the previous NGM run (the trajectory error introduced by this difference is discussed in section 4a). Persistence determines the fields to use until the next field is available. At 0600 and 1200 UTC day 15, fields for the two trajectories are from the same NGM run and forecast hour, but
unless the two trajectory endpoints are in the same position, different data values are used in the calculations. At 1800 UTC day 15, the forecast field is the 18-h forecast from the 0000 UTC day 15 NGM run, but the reference field is a 6-h forecast from the 1200 UTC day 15 NGM run. Similarly, at 0000 UTC day 17, the forecast field is the 48-h forecast from the 0000 UTC day 15 NGM run, but the reference field is a 12-h forecast from the 1200 UTC day 16 NGM run.

From the set of reference trajectories calculated for the year, some terminated prematurely when they reached the grid boundaries, particularly those starting from sites C, E, and F, because the reference lateral grid boundaries are much closer to the sites than the forecast grid boundaries (see Fig. 1). The percentage of reference trajectories reaching each grid boundary over the year are summarized in Table 2. About 10% of the reference trajectories reached the eastern boundary. Fewer reference than forecast trajectories terminate at the model top due to the greater vertical extent of the reference data. Over 80% of the reference trajectories have 48-h durations.

3. Error statistics

The basis of the trajectory error computation is the distance between corresponding forecast and reference trajectory endpoints. The distance $D$ (km) between two points on the earth identified by latitude $\theta$ and longitude $\lambda$ is defined as

$$D = 2R \left[ \sin^2 \left( \frac{\theta_f - \theta_a}{2} \right) + \cos \theta_f \cos \theta_a \sin^2 \left( \frac{\lambda_f - \lambda_a}{2} \right) \right]^{0.5},$$

where $R$, the radius of the earth, is 6371.2 km and subscripts $f$ and $a$ refer to the two points, here the forecast and reference (analysis) trajectory endpoints, respectively (A. D. Taylor 1992, personal communication). Here $D$ is a positive number. The angle between the forecast and reference trajectory is also noted to determine if the forecast is biased.

The transport error statistics are similar to those used by Rolph and Draxler (1990) though the notation differs. The mean absolute transport error $E_A$ describes the mean separation between the forecast and reference trajectory at a given time $t$ and is given by

$$E_A(t) = \frac{1}{N} \sum_{i=1}^{N} D_i(t),$$

where $N$ is the number of trajectories and $D$ is the distance between the forecast and reference trajectory endpoint pairs.

The mean relative transport error $E_R$ describes the mean separation between forecast and reference trajectories with respect to the forecast trajectory length ($L_f$, also called the travel distance) at a given time and is given by

$$E_R(t) = \frac{1}{N} \sum_{i=1}^{N} \left[ \frac{D_i(t)}{L_f(t)} \right],$$

where $L_f$ is the sum of the distances between successive forecast trajectory endpoints. The forecast trajectory is used for the reference distance because the forecast trajectory is under evaluation.

All trajectories are included in this analysis, even those with duration less than 48 h. This means that at later times into the forecast, a smaller number of trajectories are used for averaging. As noted earlier, many forecast trajectories reach the top of the grid and some of the reference trajectories reach the eastern boundary. The forecast trajectories reaching the top of the grid ($\sigma = 0.785$) are distributed throughout the year with somewhat more occurring in the summer. About twice as many reference trajectories reach the eastern boundary during the winter than during the summer, presumably because of typically stronger westerly winds. Hence, at any given travel time, the mean errors pertain to trajectories at sigma levels less than 0.785 and within the horizontal domain of the reference grid. Statistical results for less than a season are not computed.

An alternative to the separation distance statistic is the relative time-mean error (RTE). The time-mean error (TE) of the $i$th individual forecast-reference trajectory pair is defined as

<table>
<thead>
<tr>
<th>Table 1. Forecast trajectory information.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Winter</td>
</tr>
<tr>
<td>Number of days</td>
</tr>
<tr>
<td>90</td>
</tr>
<tr>
<td>Number of possible trajectories*</td>
</tr>
<tr>
<td>3240</td>
</tr>
<tr>
<td>Number of actual trajectories</td>
</tr>
<tr>
<td>2850</td>
</tr>
</tbody>
</table>

* Note there are 36 trajectories per day (6 sites × 3 levels × 2 trajectories).

<table>
<thead>
<tr>
<th>Table 2. Percentage of trajectories for the year reaching each grid boundary within 48 h.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Forecast</td>
</tr>
<tr>
<td>Reference</td>
</tr>
</tbody>
</table>
\[
TE_i = \frac{1}{M} \sum_{j=24}^{j=48} \left[ \frac{D_i(t)}{L_i(t)} \right], \tag{6}
\]

where \( M \) (generally 25) successive hourly endpoint pairs are used in the average. The lower limit of the average is +24 h because the relative error for the trajectories in this study tends to stabilize after +18 h or so (see section 4). If either one of a paired reference or forecast trajectory terminates at some time before +48 h, that time replaces 48 h as the upper limit to the average. The statistic RTE is then simply the mean of the TEIs of a set of \( N \) trajectory pairs.

4. Overall results

The results will first be discussed in terms of annual averages, then by season, both for all the data and by origin site and altitude.

a. Annual

The absolute error \( E_A \) of all the forecast trajectories increases with travel time at a rate of about 200 km day\(^{-1}\), a rate of increase that is somewhat smaller than that found by Maryon and Heasman (1988, see Fig. 3). However, the growth rate is larger than that found by Rolph and Draxler (1990) for analyzed meteorology data of comparable temporal and spatial resolution because \( E_A \) in this analysis includes the additional error introduced by the forecast. The Maryon and Heasman absolute error is less than that in the present study because their reference meteorology data were composed of a series of initialization and 6-h forecasts, resulting in no forecast trajectory error until the meteorology data differed at +12 h. In the present study, the reference and forecast trajectories usually differed after one trajectory calculation time step because of the differences in the meteorology fields at the beginning of the calculation (Fig. 2). The relative error \( E_R \) of all the trajectories (Fig. 4) is initially quite high (not shown) when the absolute errors are typically comparable to or larger than the travel distance, but the error stabilizes near 0.34 after about 18 h.

To assess the effect of using the NGM archive data with 12-h forecasts (section 2b), instead of initialization data, valid at 0000 and 1200 UTC, a new set of forecast trajectories (one month, one site only) was
compared to two different sets of reference trajectories. The forecast trajectories were generated using NGM data that were on the same horizontal grid as that used to create the one-year set of forecast trajectories, but were on pressure levels rather than sigma levels. One reference meteorology dataset was composed of a series of initialization and 6-h forecasts extracted from the 48-h forecasts (similar to Maryon and Heasman 1988), the other was the NGM archive. The mean relative error over the 24- to 48-h travel time was about 0.23 using the NGM archive and about 25% less, or 0.17, using the analysis extracted from the forecast, a result comparable to those shown in Fig. 3. The difference resulting from these two reference datasets is small in comparison to the variability in error at a given site (discussed later in this section). In addition, both of these reference datasets are an approximation of the true atmosphere. Hence, the trajectory errors with reference to the NGM archive can be assumed to have a small baseline error.

The net mean angular separation between all the forecast and reference trajectory pairs is $3^\circ$–$4^\circ$ for travel times up to about $+18$ h, decreasing to near zero by about $+36$ h. This means the forecast trajectory is biased to the left of the reference trajectory early in the forecast, but there is no bias later in the forecast.

Since the meteorology data are on a regional scale (spacing about 185 km) and with the high relative error for small travel times, the error at a later time such as $+36$ h or the time-mean error ($+24$ to $+48$ h average) of forecast-reference pairs will be shown in subsequent analyses.

The relative errors $E_R$ plotted in Fig. 4 are the means over the trajectory duration for all the trajectories over the entire year, but the distribution about the mean is also important for application to operational forecasts. Box plots for the $E_A$ and $E_R$ distributions at $+36$ h are shown in Fig. 5. The horizontal lines of the boxplot show the 10th, 25th, 50th, 75th, and 90th percentiles. The distributions are fairly broad and are similar because the mean travel distance at $+36$ h is about 1000 km. The median of the $E_R$ distribution is 0.26. Half of the relative errors are between 0.15 and 0.44 (the 25th and 75th percentiles), with 10% of the errors less than 0.09 and 10% greater than 0.68. To apply this distribution to a forecast trajectory, two circles could be drawn on a map (not shown) of the trajectory centered on the $+36$-h trajectory endpoint with radii equal to 26% and 68% of the travel distance to that endpoint. Half of the $+36$-h reference trajectory endpoints of this study would be within the inner circle and 90% would be within the outer circle.

To determine what may be contributing to the range of error, the error in terms of origin altitude, site, and time of day are now considered. The RTEs for the three altitudes are fairly similar, but slightly greater for higher altitudes: 0.31 for 500 m, 0.32 for 1000 m, and 0.34 for 1500 m. The site A RTE, 0.45, is higher than for the other sites, which range from 0.26 for site C to 0.35 for sites D and F (Table 3). The number of trajectories used to compute the RTE for site A is fewer than for the other sites because many forecast trajectories reach the top of the grid before $+24$ h. They reach the grid top because of the high terrain at the origin and possibly because of more low pressure systems causing upward transport than at site D, which is also in high terrain but farther to the south. Differences in RTEs for trajectories beginning at 0000 and 1200 UTC are small. For instance, the mean RTE for site C, 500 m trajectories beginning at 0000 UTC is 0.24, but 0.26 for 1200 UTC trajectories. Thus, the differences in error are greater among the various sites than for the three levels, suggesting that regional weather patterns rather than differences in the wind fields at various altitudes has the greatest influence on the error.

The results by origin site and altitude are summarized in Table 4, which shows $E_A$ and $E_R$ at $+36$ h. For the various origin site-altitude combinations, $E_A$ ranges from 235 to 449 km and $E_R$ from 0.26 to 0.54, both about a factor of 2. For a given origin site, $E_A$ and $E_R$ are generally about 15% greater for trajectories begin-

<table>
<thead>
<tr>
<th>Site</th>
<th>RTE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.45</td>
</tr>
<tr>
<td>B</td>
<td>0.30</td>
</tr>
<tr>
<td>C</td>
<td>0.26</td>
</tr>
<tr>
<td>D</td>
<td>0.35</td>
</tr>
<tr>
<td>E</td>
<td>0.32</td>
</tr>
<tr>
<td>F</td>
<td>0.35</td>
</tr>
</tbody>
</table>
Table 4. Annual absolute ($E_a$) and relative ($E_r$) error for travel time +36 h by origin site and altitude.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Altitude</th>
<th>Site</th>
<th>Site</th>
<th>Site</th>
<th>Site</th>
<th>Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_a$ (km)</td>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
</tr>
<tr>
<td>500 m</td>
<td>383</td>
<td>296</td>
<td>253</td>
<td>235</td>
<td>274</td>
<td>280</td>
</tr>
<tr>
<td>1000 m</td>
<td>434</td>
<td>328</td>
<td>269</td>
<td>243</td>
<td>292</td>
<td>271</td>
</tr>
<tr>
<td>1500 m</td>
<td>449</td>
<td>336</td>
<td>288</td>
<td>263</td>
<td>312</td>
<td>276</td>
</tr>
<tr>
<td>$E_r$</td>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
</tr>
<tr>
<td>500 m</td>
<td>0.46</td>
<td>0.30</td>
<td>0.26</td>
<td>0.36</td>
<td>0.31</td>
<td>0.33</td>
</tr>
<tr>
<td>1000 m</td>
<td>0.54</td>
<td>0.32</td>
<td>0.28</td>
<td>0.40</td>
<td>0.33</td>
<td>0.37</td>
</tr>
<tr>
<td>1500 m</td>
<td>0.52</td>
<td>0.32</td>
<td>0.29</td>
<td>0.44</td>
<td>0.38</td>
<td>0.42</td>
</tr>
</tbody>
</table>

At 1500 m than 500 m, a trend noted earlier. Also, again, site A has the greatest error and site C (Ontario) has the smallest relative error. From the distributions of $E_r$ at +36 h by site for the 500-m trajectories for the year, the errors vary by about a factor of 3 between the 25th and 75th percentiles (i.e., from about 0.15 to 0.45 for sites B, C, E, and F and from about 0.2 to 0.6 for sites A and D, Fig. 6). Here $E_r$ is about 0.1 at the 10th percentile and about 0.6 at the 90th percentile, with slightly higher values for sites A and D.

b. Seasonal

The errors were averaged by season to investigate the relationship, if any, between season and error. Figure 7 shows seasonal mean absolute error and trajectory travel distance for times +24, +36, and +48 h. Travel distances are low in the summer and high in the winter, as expected, because of typical differences in wind speed through the year. Since the absolute error is also higher in the winter, and lower in the summer, the relative error $E_r$ (not shown) is generally constant through the year. Box plots of seasonal $E_r$ at +36 h show nearly identical distributions over the four seasons (not shown).

5. Trajectory cluster analysis

To determine if trajectory error is associated with synoptic situation, a trajectory cluster analysis program was developed, seasonal clusters were identified for each trajectory origin, and mean errors for each cluster were calculated. Cluster analysis is a common statistical technique and is described in textbooks such as Gordon (1981) and Romesburg (1984). Clustering is a process of grouping similar objects together whereby differences among individual elements in a cluster are minimized and differences among clusters are maximized. When applied to trajectories, the goal is a collection of distinct clusters representing different classes of synoptic regimes over the duration of the trajectory; with each cluster being composed of similar trajectories. Note that a traditional synoptic classification is valid at a given time, but that trajectory cluster analysis includes the temporal evolution of the wind fields. Clustering has been applied to air parcel trajectories to investigate air chemistry (e.g., Moody and Galloway 1988; Moody and Samson 1989; and Dorling et al. 1992) and to develop long-range transport climatology (e.g., Harris and Kahl 1990; Harris 1992).

a. Cluster method

The cluster program uses Ward’s method (see Romesburg 1984), as did Moody and Galloway (1988), and is described as follows. The cluster variables are the trajectory endpoints, which represent wind speed and direction. Given $N$ trajectories, each initially defined as a separate cluster with zero spatial variance, where the cluster spatial variance is the sum of the squared distances between the endpoints of the cluster’s component trajectories and their (cluster) mean. Successive steps through the clustering process combine the two clusters that result in the minimum increase in total spatial variance (TSV), where TSV is the sum of all the cluster spatial variances. The com-
b. Cluster results

Clustering was done by season on the forecast trajectories originating at each site (6) and level (3). The forecast trajectories were clustered, rather than the reference trajectories, because the emphasis in operations is on the forecast. Because many forecast trajectories reached the grid top before 48-h and clustering requires all trajectories to have the same duration, only trajectories with at least 36-h duration were used. This resulted in clustering of about 40%, 55%, and 70% of all possible trajectories originating at altitudes 1500, 1000, and 500 m, respectively. No clustering for site A was done for summer at 500 m, winter and summer at 1000 m, and all seasons at 1500 m because fewer than 30 trajectories were available for each of these cases. Typically, about six or seven clusters were identified for each origin.

As an example, Fig. 8 shows the 36-h forecast trajectories in each of the resulting five clusters for trajectories originating during autumn from site E (east-central Oklahoma) at 500 m. The five clusters each look distinct with the set of trajectories being categorized generally by direction and travel distance. The cluster 1 trajectories originate behind a cold front and go to the south with travel distances of about 1300 km. Cluster 2 contains trajectories going about 1300 km to the northeast because of a high over the eastern United States and an approaching frontal system. The trajectories in cluster 3 tend to go about 850 km toward the north, beginning ahead of a north-south-oriented trough. Cluster 4 trajectories tend to go toward the east and northeast because of a high located over the southeastern United States or a cold front near the origin; travel distances are about 1000 km. The trajectories in cluster 5 do not go in any one direction, but generally stay fairly close to the origin (i.e., within about 450 km), and originate under a high pressure system or near a stationary front.

The results from the site E—500-m cluster runs for the four seasons are shown in Fig. 9 and Table 5. The cluster-mean trajectories are shown in Fig. 9 (Fig. 9d corresponding to Fig. 8) and the cluster-mean error, standard deviation, and the number of trajectories in each cluster are given in Table 5. A cluster-mean trajectory is simply the mean of the set of trajectories forming a cluster. The cluster-mean error (standard deviation) is the mean (standard deviation) of the RTEs of the individual trajectories in a cluster. Cluster numbers (shown at the end of each cluster-mean trajectory) are arbitrarily defined to correspond to cluster-mean error with cluster number 1 having the smallest error and greater numbers having higher error.

Ten clusters were identified for winter, nine each for spring and summer, and five for autumn. The winter cluster-mean trajectories (Fig. 9a) generally are oriented in all directions, with relatively shorter and longer travel distances being apparent. Cluster 10, with a mean trajectory going to the east, contains the greatest number of trajectories and the greatest error. The spring cluster-mean trajectories (Fig. 9b) are similar to the winter case with trajectories of varying travel distances generally oriented in all directions. Cluster 7, with a mean trajectory going to the north, has the most trajectories. The summer cluster-mean trajectories (Fig. 9c) show the predominance of transport to the north and northeast; no cluster-mean trajectories go to the south. The autumn cluster-mean trajectories show little transport to the west (Fig. 9d) even though some trajectories within cluster 5 go to the west (Fig. 8). The cluster 5 mean trajectory is oriented to the south-southwest and is much shorter than the typical cluster 5 trajectory travel distances because the trajectories in the cluster are oriented in many directions. The other cluster-mean trajectories in Fig. 9d are more similar to the component trajectories because of the more unidirectional pattern of the trajectories.

Overall, the differences among the cluster-mean trajectories for a given seasonal cluster run show the effectiveness of the cluster analysis in classifying trajectories. The cluster analysis also identified clusters with relatively high and low mean errors. While most of the clusters had errors of 0.20–0.40, the minimum seasonal errors were in the range 0.12–0.19 (Table 5) and the maxima were about two to four times greater, 0.42 to 0.52. The cluster-standard deviations of the error (Ta-
ble 5) are about 60% of the cluster-mean error, suggesting a difference, but probably not at a high level of significance, between the maximum and minimum errors. The distribution of errors were similar throughout the year.

The minimum and maximum cluster-mean errors and their cluster-mean trajectories for the six origin sites during the four seasons (500-m trajectories) are shown in Fig. 10 to focus on the differences between the extremes over all the origins rather than to investigate differences in mean-cluster errors for a given site. The maximum cluster-mean errors, generally in the 0.4-0.5 range, are about twice the minimum cluster-mean errors. As noted earlier, the errors for site A tend to be higher than for the other sites. No consistent differences among the seasons are apparent, as might be expected given the differences in climate among the sites.

For many of the origins (see the autumn site E case, Fig. 9d) the cluster-mean trajectory with the smallest travel distance is the cluster with the greatest error and one of the longer mean-cluster trajectories is associated with the minimum error. Mean sea level pressure maps were inspected to relate the clusters with the extreme errors to meteorological synoptic situations. The synoptic situation was noted at the start of the component trajectories. For sites A and D, in the mountains, the relationship between the synoptics and trajectories forming the extreme error clusters was not well defined; hence, only the sites east of the Rockies are discussed. Trajectories beginning both behind and ahead of a cold front were associated with the cluster-minimum error. For both of these situations, generally strong, steady winds are common and produce relatively long, unidirectional trajectories. The postfrontal cases with the cluster-mean trajectory generally going to the south were
observed for winter and spring from sites C (Ontario), E (Oklahoma), and F (North Carolina/South Carolina), for summer from site C, and for autumn from sites C and E (Fig. 9). The prefrontal cases with the cluster-mean trajectory generally going to the northeast were observed for winter and spring from site B (Minnesota), for summer from sites B, E, and F, and for autumn from sites B and F. Trajectories beginning near stationary or slowly moving fronts or troughs, or under weak pressure gradients typical of high pressure systems were associated with the cluster-maximum error. These weather patterns produce relatively shorter trajectories oriented in various directions, and hence a comparatively short cluster-mean trajectory.

The details of the cluster results from the sets of 1000- and 1500-m trajectories, that is, the number of clusters and the cluster-mean trajectories, differed from those for the 500-m trajectories as expected, primarily because trajectories at the three levels are different and because of the missing trajectories at higher levels. Yet the difference between the minimum and maximum cluster-mean errors were again about a factor of 2. For trajectories beginning at all altitudes, 80% of the maximum cluster-mean errors were at least twice the minimum error. The relationship between the higher-level clusters and the synoptics was not investigated.
6. Summary and conclusions

Trajectory models and/or dispersion models driven by forecast meteorology are used to position aircraft for sampling tracer gases or other species during field experiments and to predict regions of high concentration of hazardous substances in response to an emergency. Clearly, in applying forecasts it is important to know their accuracy. The purpose of the present study was to estimate the forecast trajectory error and to relate the error to synoptics. Three-dimensional forecast and analysis (reference) trajectories were computed from six widely spaced sites over North America at three levels within the lower troposphere, many within the boundary layer (500, 1000, and 1500 m above ground), twice a day for a one-year period using NGM data.

Absolute forecast trajectory error refers to the distance between corresponding forecast-reference trajectory endpoints; relative error is the absolute error divided by the forecast trajectory travel distance. For the entire collection of trajectories, mean relative errors are about 0.35 after travel times of +18 h. The forecast is biased a few degrees to the left of the reference for travel times up to about +18 h, but no bias is apparent after +36 h. Differences in annual or seasonal mean relative error among the starting altitudes and among the origin sites were much less than the differences observed from trajectories originating at a given site and altitude. Differences in annual mean relative error between trajectories starting at 0000 and 1200 UTC were also small. The +36-h distributions of relative error for trajectories beginning from 500 m above a given origin site show a large range in error. Half of the errors in most of the distributions are within the range of about 0.15–0.45 (25th–75th percentiles), a factor of 3. Ninety percent of the errors are less than about 0.6, except for site A on the Montana–Wyoming border with 0.9. The apparently higher error at site A may result from the exclusion of many forecast trajectories when they reached the top of the meteorology grid, or from other factors, such as orography or climatology.

Absolute error and travel distance are correlated such that relative error is nearly constant throughout the year. Box plots of relative error by season show nearly identical distributions.

The forecast trajectories were objectively classified through a trajectory cluster analysis, which groups trajectories by direction and travel distance. The trajectory cluster analysis is effectively a classification of wind fields with time, because trajectories are Lagrangian events, and differs from a synoptic classification at a given time, say, the trajectory origin time. Clustering was done by season for each of the origin site-altitude sets of trajectories. Six or seven clusters were typically identified for each origin with each cluster generally containing comparatively long or short trajectories oriented in a given direction. Usually there was one cluster with relatively short trajectories oriented in many directions. Cluster-mean errors were obtained by averaging the time-mean (relative) errors of the cluster’s component trajectories. Differences between the minimum cluster-mean error and the maximum cluster-mean error for all the sets of trajectories clustered was large, that is, about a factor of 2–3. Trajectories in the minimum-error cluster tended to originate within stronger, steady flow either ahead of or behind a cold front. Trajectories in the maximum-error cluster tended to originate in regions of generally slow wind fields such as under a high or near stationary or slowly moving fronts. Thus, the large range in the distribution of relative error from trajectories originating from a given site altitude are likely related to synoptics. Differences in synoptic patterns over North America may contribute to some of the relatively small differences among the six trajectory origin sites.
Fig. 10. Cluster-mean trajectories with the minimum (O) and maximum (+) cluster-mean errors for the 500-m trajectories from the six origin sites (●) for (a) winter, (b) spring, (c) summer, and (d) autumn. Errors (%) are labeled at the end of the trajectories.

REFERENCES


