High-Resolution Stratospheric Tracer Fields Reconstructed with Lagrangian Techniques: A Comparative Analysis of Predictive Skill

R. DRAGANI, G. REDAELLI AND G. VISCONTI
Department of Physics, University of L'Aquila, L'Aquila, Italy

A. MARIOTTI
ENEA National Agency, Rome, Italy

V. RUDAKOV
Central Aerological Observatory, Moscow, Russia

A. R. MACKENZIE
Environmental Science Department, Lancaster University, Lancaster, United Kingdom

L. STEFANUTTI
IROE, Florence, Italy

(Manuscript received 27 September 2000, in final form 16 November 2001)

ABSTRACT
Numerical experiments and statistical analyses are conducted to determine the skill of different Lagrangian techniques for the construction of tracer distributions. High-resolution potential vorticity (PV) maps are calculated from simulations of the 1996/97 arctic winter stratospheric dynamics using two different numerical schemes—reverse domain filling trajectories (RDF) and contour advection with surgery (CAS)—and data from three meteorological agencies (NCEP, the Met Office, and ECMWF). The PV values are then converted into ozone (O$_3$) concentrations and statistically compared to in situ O$_3$ data measured by the electro chemical ozone cell (ECOC) instrument during the Airborne Polar Experiment (APE) using cross correlation, rns differences, and the Kolmogorov–Smirnov (KS) test.

Results indicate that while Lagrangian techniques are successful in increasing the presence of lower-scale tracer structures with respect to the plain meteorological analyses, they significantly improve the statistical agreement between the simulated and the measured tracer profiles only when there is clear evidence of filaments in the measured data. This better fit is most clearly seen by using the KS test, rather than cross correlation. It is argued that this difference in the performance of Lagrangian techniques can be partly related to the treatment of mixing processes in the framework of the Lagrangian schemes. Statistical analyses also show that the temporal rather than the spatial resolution of the input meteorological fields, used to advect tracers, enhances the predictive skill of the Lagrangian products. The best overall performance is obtained with the Lagrangian product (not gridded) based on high-resolution reverse trajectories calculated along a flight track, in particular when the simulation is initialized with ECMWF data. Other products, such as CAS initialized with ECMWF and 3D-gridded RDF initialized with the Met Office data, show fairly good performances, thus with lower statistical confidence.

1. Introduction
Wintertime stratospheric dynamics at high latitudes is dominated by the stratospheric polar vortex. It has been argued (e.g., McIntyre and Palmer 1984; Newman et al. 1996; Schoeberl and Sparling 1995) that large-scale polar vortex dynamics is coupled with fine-scale filamentary structures, a kind of microerosion of the vortex itself. Similar fine-scale features may be also found in aircraft measurements of chemical tracers in the lower stratosphere (e.g., Plumb et al. 1994; Schoeberl and Newman 1995; Mariotti et al. 2000). To improve the interpretation of these observations, different Lagrangian techniques have been developed in the last few years that reconstruct tracer fields to compare with these observations (Fairlie et al. 1997; Schoeberl and
Newman 1995; Waugh et al. 1994). Lagrangian numerical schemes represent valid tools to explore the evolution of the stratospheric polar vortex and its filamentation. In fact, these tools can be used, starting from low-resolution meteorological analyses, to obtain higher-resolution tracer fields, at horizontal scales that cannot be normally resolved by chemical transport models. These techniques display the small-scale information already contained within the time series of the meteorological analyses (Mariotti et al. 1997) based on the assumption that the small-scale flow is controlled by the large-scale quasi-horizontal advection (Waugh and Plumb 1994).

In this paper we explore the potential of reverse domain filling trajectory (RDF) and contour advection with surgery (CAS) schemes in simulating the structure of filaments near the polar vortex edge, as reported by experimental data. For this purpose, reconstructed potential vorticity (PV) fields are compared with aircraft ozone observations made during the Airborne Polar Experiment (APE). The use of PV as a tracer of stratospheric motion is based on the fact that this quantity is conserved inside each air parcel for adiabatic frictionless motion, a condition that we can assume is met in the wintertime middle stratosphere, over timescales of 1–2 weeks (see e.g., Schoeberl and Sparling 1995). Moreover, since the scale of filaments is usually much smaller than the scales of the horizontal (isentropic) atmospheric motion, it is reasonable to assume that the PV anomalies associated with the filaments are not able to alter the local wind fields and that PV filaments tend to be passively advected by large-scale quasi-horizontal flow (Schoeberl and Newman 1995).

In this work, six different kinds of PV “products” are considered: the output fields obtained by the RDF technique with four different final spatial resolutions, the CAS output field, and also the plain (original) PV meteorological analyses. The statistical tests applied to this selection of techniques allow us to estimate the influence of the spatial resolution on the quality of the RDF reconstructed fields, to compare the results of RDF and CAS techniques, and to quantify the improvement in the results of these Lagrangian methods in comparison to the plain meteorological analyses. The influence of different meteorological fields used to advect the tracers is also tested using winds and temperature data from three different meteorological agencies [the National Centers for Environmental Prediction (NCEP), the Met Office, and the European Centre for Medium-Range Weather Forecasts (ECMWF)].

A somewhat similar analysis was conducted by Fairlie et al. (1997) to quantify the level of predictive skill of analyzed and 5-day backward RDF PV in reconstructing nitrous oxide tracer structure in the Southern Hemisphere during the Airborne Southern Hemisphere Ozone Expedition–Measurements for Assessing the Effects of Stratospheric Aircraft (ASHOE–MAESA) campaign. In their study, they found no evidence for statistically significant improvement in predictions based on RDF reconstruction with respect to the analyzed one. Our results confirm some of the Fairlie et al. (1997) results, showing that, from a rigorous statistical point of view, high-resolution reconstructions show only a marginal improvement with respect to the plain meteorological PV. However, the analysis discussed in our paper differs from Fairlie et al. (1997) not only for the different environment studied (Northern Hemisphere winter, instead of the Southern Hemisphere), the different chemical species considered (ozone instead of nitrous oxide), and the use of different Lagrangian techniques, besides the single RDF types, but basically for the different approach. The basic aim of this paper is in fact, to investigate the impact of all the a priori assumptions (such as the type of technique, the source of the meteorological input, the resolution of the input/output grids, etc.) in enhancing or reducing the agreement of the reconstructed data with the tracer measures, rather than to bring new evidence/estimate of the absolute statistical degree of significance of the various results. Such an analysis would eventually lead us to select a combination of those guesses that identify Lagrangian products showing the best overall performances, and thus the best candidate to be routinely used.

The next section contains a description of meteorological and experimental data used in this work, while the description of the Lagrangian techniques used can be found in section 3. Section 4 presents the statistical analyses and the discussion of the results. Conclusions are presented in the last section.

2. Meteorological and experimental data

Meteorological analyses (winds, temperature and PV) from three different sources are used: NCEP analyses (Newman et al. 1989), Met Office Upper Atmosphere Research Satellite (UARS) analyses (Swinbank and O’Neill 1994), and ECMWF operational analyses (ECMWF 1995). These data are made available on a global pat-pong grid on pressure levels. The NCEP data have a spatial resolution of 5° in longitude and 2° in latitude, and are available once a day at 1200 UTC over 18 pressure levels from 1000 mb up to 0.4 mb. Met Office analyses have a spatial resolution of 3.75° in longitude and 2.4° in latitude over the same pressure levels of NCEP data, and are also given once a day at 1200 UTC. ECMWF analyses are available on a regular 1.125° by 1.125° grid over pressure levels from 1000 mb up to 30 mb, four times a day at 0000, 0600, 1200, and 1800 UTC.

The O₃ data used in this work are obtained from in situ measurements by the ECOC electro-chemical ozone sonde on board the M55-Geophysica aircraft during APE (Kyro et al. 2000). The APE mission was based at the arctic polar circle during the winter of 1996/97 (Stefanutti et al. 1999) and studied the local chemical and physical properties of the lower stratosphere. Dur-
ing the APE mission, six scientific flights were performed during the period from 23 December to 14 January. In particular, flights in the first part of the campaign, up to the end of 1996, were purposely planned mainly to study stratospheric transport and mixing processes inside and along the edge of the polar vortex. During this phase, global RDF trajectory calculations were performed in forecast mode and used as a tool to plan the flight tracks, predicting the position of the polar vortex edge and the occurrence of filamentary structures (Stefanutti et al. 1999; Redaelli et al. 1997). The ECOC instrument provided in situ O$_3$ measurements along the flight track with a time response of 30 s, an estimated accuracy of 6%, and a precision of 3 ppbv. ECOC O$_3$ data used in this study are smoothed using a 2-min average in order to remove the part of the signal associated with structures on spatial scales that cannot be resolved by either the analyzed or the Lagrangian products. Considering the aircraft mean speed (≈730 km h$^{-1}$), this averaging time corresponds to an equivalent degraded horizontal resolution of 25 km. During ascents and descents, which were typically performed with vertical velocity of 5 m s$^{-1}$, the resolution of the averaged data is 600 m. In this study, O$_3$ is used as a chemical inert tracer. While the photochemical lifetime of ozone in the lower stratosphere is long enough to warrant such hypothesis under most conditions, the potential polar stratospheric cloud (PSC) processing in the wintertime polar vortex region should be considered. However, analyses performed for the APE–Polar Ozone, Leewaves, Chemistry, and Transport (POLECAT) campaign using ECOC and ozoneonde data, together with chemical transport model, showed no evidence of chemical ozone depletion at the cruising altitudes of the aircraft (Kyro et al. 2000). The absence of chemical depletion is also in agreement with observations showing no PSCs at aircraft cruising altitudes during the campaign (Stefanutti et al. 1999).

Figure 1 shows the aircraft routes and ECMWF PV meteorological analyses, for all the scientific flights considered in this study. In Fig. 2, we present the measured O$_3$ profiles and the potential temperature ($\theta$) along each of the flights. The concentration of O$_3$ fluctuations with various frequencies and magnitudes can be seen during all flights. In particular, rapid and intense O$_3$ variations were observed while flying at almost constant $\theta$ along sections of the flights performed on 23, 29, and 31 December. Under the hypothesis that these variations are not related to $\theta$ excursions, and that O$_3$ is not undergoing rapid localized depletion, it is reasonable to suppose that quasi-horizontal (isentropic) advection may be responsible for large-scale mixing, bringing close together air masses with different tracer content; in this scenario the O$_3$ fluctuations are an indication that the plane is cross sampling vortex filaments carrying distinct O$_3$ concentrations.

However, it should be also taken in account that part of the observed variability in O$_3$ data along each flight track could be attributed to the vertical gradient in the O$_3$ and the plane is departure from a constant potential temperature level. It means that if the aircraft is sampling ozone in a region where its vertical gradient is very steep, even small changes in $\theta$ could, in principle, lead to rather significant changes in O$_3$ at the altitude of the aircraft, resulting in a kind of “false” filament in the flight data. The impact of small excursion in potential temperature ($\Delta \theta$) on fluctuations in ozone ($\Delta$O$_3$) has been estimated, using aircraft O$_3$ and $\theta$ data for all the 6 flights, obtaining values of $\Delta$O$_3$/$\Delta \theta$ in the range of 0.01–0.04 ppmv K$^{-1}$ over the $\theta$ surfaces of interest. Such values can be used to define a kind of threshold to separate $\theta$-based ozone variations from fluctuations due to isentropic advection.

Hereafter, we consider only these last kind of features as a signature of “possible filaments,” and use them to test the skill of Lagrangian techniques in reproducing tracer fields at scales smaller than the resolution of the input tracer or meteorological fields used in the simulations. Labels in Figs. 1 and 2 indicate different flight sections selected for our analysis, as described in section 4.

3. The Lagrangian techniques

a. Reverse domain filling trajectory

RDF trajectory techniques are based on the simultaneous calculation of atmospheric trajectories of many parcels, each of them of infinitesimal dimension and unable of being deformed during its motion by shear or strain (Schoeberl and Sparling 1995; Morris et al. 1995). To avoid having inhomogeneous parcel distributions at the time of interest (in the case of this work at the time of the flight), “reverse” trajectories are calculated with particles advected backward in time, starting from a regular grid at the moment of interest (Sutton et al. 1994). The resolution of this grid of parcels determines the resolution of the reconstructed tracer fields. The trajectories used for the RDF analysis are calculated with a global trajectory model that uses winds obtained from the assimilation schemes of meteorological agencies. The code used is an isentropic model (Redaelli 1997) that conserves potential temperature explicitly and makes use of a fourth-order Runge–Kutta time integration scheme to advect particles with a time step of about 20 min. Horizontal components of wind velocity are calculated at particle locations by cubic-spline interpolation from values at surrounding grid points. Potential temperature conservation is accomplished by interpolating wind fields on the selected isentropic surface with the potential temperature being calculated from temperature on pressure levels.

The hypothesis that a parcel’s potential temperature is conserved is an approximation that imposes a temporal limit for the length of the simulations, estimated at about 10–15 days (Schoeberl and Sparling 1995). Furthermore, it is important to note that the trajectory
calculation does not include diffusive terms, and this could lead to an incorrect simulation of local dispersion properties when the integration time becomes too long. In this study, 5-day backward trajectories are calculated, following Sutton et al. (1994), Schoeberl and Newman (1995), and Redaelli (1997), where it has been observed that 5-day integrations were long enough to capture the most important fine-scale features. Potential vorticity is assumed to be conserved along the trajectory [very reasonable on a 5-day period (Schoeberl and Sparling 1995)] and is used to label each parcel and follow its path during the simulated motion.

Four different categories of RDF simulation are considered. Three of them are obtained by defining the PV tracer on a fixed bidimensional grid at the time of interest and differ only in resolution and area of the horizontal grid. The calculations are performed on multiple isentropic surfaces selected with a vertical resolution of 5 K in the range of potential temperature measured during each flight. For comparison with the observed aircraft data for every flight, the gridded fields are first interpolated to the horizontal location where parcels have been sampled and then interpolated vertically to their potential temperature surface. The fourth kind of RDF calculation is obtained by interpolating the PV directly to the three-dimensional (3D) positions in which...
parcels have been sampled during each flight. Then, for every interpolated PV value, the backward trajectory is calculated. In this way, resulting tracer values are already over a one-dimensional track instead of over a three-dimensional grid. For this calculation, a constant time sampling is chosen along the examined flight track, so that horizontal and vertical resolution of the final reconstructed tracer may vary slightly.

Hereafter, RDF$_0$ indicates the RDF calculations on an global domain with range of latitudes starting from 54$^\circ$N up to 90$^\circ$N, with a latitudinal resolution of 1$^\circ$ and a longitudinal resolution that varies with the latitudes, being 2.5$^\circ$ at $\sim$$65^\circ$N; RDF$_t$ indicates that calculations are performed on a less extended regular grid, chosen to completely cover only the flight track area, with fixed spatial resolutions of 1$^\circ$ in latitude and 2$^\circ$ in longitude; RDF$_{xl}$ is RDF performed on the same grid of RDF$_t$ but at higher spatial resolution, 0.5$^\circ$ in latitude and 1$^\circ$ in longitude.

RDF$_{st}$ refers to the nongridded RDF product with an average resolution of $\sim$25 km along the flight track, corresponding to about 0.2$^\circ$ either in longitude or latitude. RDF calculations in these four different categories are performed using analyses from NCEP, Met Office, and ECMWF sources. Tables 1 and 2 provide a summary of the products and the acronyms used in this paper.

b. Contour advection with surgery

CAS, an algorithm derived from contour dynamics [Dritschel and Saravanan (1994) and reference therein], is a numerical method that applies exactly to the two-dimensional motion of any inviscid fluid that has a generalized vorticity invariant (PV in this case; Dritschel and Legras 1993; Waugh and Plumb 1994). It is based on the piecewise uniform representation of PV, which allows the evolution of the 2D motion of the fluid to be depicted by that of the PV isolines (contours) that divide regions of uniform PV. The CAS technique was originally developed independently by Norton (Norton 1994) and Waugh and Plumb (Waugh and Plumb 1994).

In this framework, the stratospheric flow is approximated by parcels that are constrained to conserve PV and move isentropically. The evolution of the flow is
4. Results and discussion

a. Analysis technique

In order to assess the improvement obtained by using Lagrangian methods to resolve observed small-scale features, PV fields from meteorological analyses (hereafter ANL; see Table 1 for specific acronym) are also considered, along with the RDF and CAS simulated ones, when comparing to experimental data. On the whole, for each scientific flight, we consider four high-resolution reconstructed PV outputs (three 3D regularly gridded RDF and one 3D irregularly gridded CAS), one 1D profile along the flight track, and PV from the analyses; that is six fields altogether for each of the three meteorological data sources: ECMWF, NCEP and Met Office (so actually eighteen PV profiles are considered for each comparison with observed O₃ data).

Using a least squares linear fit between each of the PV profiles calculated over the quasi-isentropic sections of flights and the corresponding O₃ measured data, PV is translated into O₃ reconstructed values. ECOC O₃ data are restricted to the 420–480 K potential temperature range, which corresponds approximately to altitudes in the range 15–21 km. For the comparison the flights are divided into sections during which the plane followed a quasi-isentropic track within this altitude range (specifically as reported in Table 3: one section during flights on 23 December and 9 January, two sections during flights on 29 December and 7 January, and three during flights on 31 December and 14 January, a total of twelve quasi-isentropic flight sections. Locations of the flight sections are shown in Figs. 1 and 2). A tolerance of ±5 K is considered when selecting quasi-isentropic sections of the flights. Such value of θ corresponds also to the vertical resolution of the not-gridded products.

Table 2. Spatial resolution used for each product. Here, Δx and Δy indicate the spatial resolution in kilometers between two grid points in west-east and north-south directions, respectively, for the gridded product. "Res" indicates the average resolution in kilometers, along the flight track in the case of RDFₑₑ, and along a PV contour in the case of CAS scheme.

<table>
<thead>
<tr>
<th>Product</th>
<th>Δ long</th>
<th>Δ lat</th>
<th>Δx</th>
<th>Δy</th>
<th>Res</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANLₑₑ</td>
<td>5°</td>
<td>2°</td>
<td>235</td>
<td>222</td>
<td>nat</td>
</tr>
<tr>
<td>ANLₑₑ</td>
<td>1.125°</td>
<td>1.125°</td>
<td>53</td>
<td>125</td>
<td>nat</td>
</tr>
<tr>
<td>ANLₑₑ</td>
<td>3.75°</td>
<td>2.4°</td>
<td>176</td>
<td>267</td>
<td>nat</td>
</tr>
<tr>
<td>ANLₑₑ</td>
<td>2.5°</td>
<td>2.5°</td>
<td>130</td>
<td>130</td>
<td>nat</td>
</tr>
<tr>
<td>RDFₑₑ</td>
<td>2.8°</td>
<td>1°</td>
<td>135</td>
<td>111</td>
<td>nat</td>
</tr>
<tr>
<td>RDFₑₑ</td>
<td>2°</td>
<td>1°</td>
<td>94</td>
<td>111</td>
<td>nat</td>
</tr>
<tr>
<td>RDFₑₑ</td>
<td>1°</td>
<td>0.5°</td>
<td>47</td>
<td>56</td>
<td>nat</td>
</tr>
<tr>
<td>RDFₑₑ</td>
<td>0.2°</td>
<td>0.2°</td>
<td>11</td>
<td>22</td>
<td>25</td>
</tr>
<tr>
<td>CAS</td>
<td>nat</td>
<td>nat</td>
<td>nat</td>
<td>nat</td>
<td>nat</td>
</tr>
</tbody>
</table>

* Values are calculated at the mean latitude of 65°N.
In the remainder of the paper, flights are identified by Roman numbering while subsequent sections along the same flight are labeled in alphabetical order (see Table 3).

To compare the PV fields with the observed \(O_3\) data, PV values are interpolated to the latitude, longitude and potential temperature of the 2-min-averaged sampled data. Then a least squares linear fit of the six PV profiles to the \(O_3\) observations is calculated for each isentropic part of the APE flights. This relationship is used to translate PV values into \(O_3\). Figures 3, 4, and 5 give an example showing the basic steps of the method used to reconstruct \(O_3\) profiles from PV. Figure 3 presents a comparison between PV determined from ECMWF analyses and results from RDF and CAS calculations, also with ECMWF data, at the time of two given flights. PV fields from RDF and CAS present large-scale features similar to the low-resolution meteorological PV analyses, plus finer-scale filamentary features, absent in the analyses. Figure 4 presents PV versus ECOC \(O_3\) scatterplots for each of the six PV profiles, from analyses, and from the various RDF and CAS simulations all calculated using ECMWF data. For each scatterplot a linear least squares fit is calculated to provide a relation that allows a conversion from PV values to \(O_3\) concentrations. The positive slope of the linear least squares fits implies that PV and \(O_3\) are directly correlated. Figure 5 shows the predicted \(O_3\) profiles from the six PV products compared to the measured \(O_3\) profile. From this figure, Lagrangian products, and especially the RDF ones, seem to fairly match the variations of the 2-min-averaged \(O_3\) and to capture the supposed filaments as sampled by the aircraft, even if they are not able to reproduce well the amplitude of the observed variations. At the same time, the predicted profiles obtained from plain ANL do not show a good agreement with the observed profile when intense and rapid tracer variations are sampled. However, while we can conclude by visual inspection that Lagrangian techniques seem to generally improve the fit, it is more difficult to define from a statistical point of view the real improvement of the predicted profiles with respect to the observed one and so their agreement.

**b. Statistical analyses**

As a first statistical analysis we compute the cross correlation coefficients (Fig. 6) and the rms differences (not shown here) between the observed and the reconstructed \(O_3\) profiles for each of the quasi-isentropic sections of the APE flights, using all of the various PV sources (18 all together for each flight section). The cross correlation coefficient \(R\) gives a measure of the agreement between observed and predicted \(O_3\) in comparison to the variance of each dataset. The coefficient \(R\) is a scalar quantity in the interval \([-1.0, 1.0]\), defined as the ratio of the covariances of the sample populations to the product of their standard deviations. A value of \(R = +1\) or \(R = -1\) indicates a perfect fit to a positive or negative linear model, respectively. A value of \(R\) close to \(\pm 1\) indicates a high degree of correlation and a good fit to a linear model. A value of \(R\) close to 0 indicates a poor fit to a linear model (Sveshnikov 1968; Morris et al. 1995).

The rms instead provides a measure of the absolute differences between the observed and the reconstructed \(O_3\) profiles. High cross-correlation values and low rms differences imply good fits. Higher correlations (lower rms values) are found for flights IV, V, and VI, while lower correlation coefficient values (higher rms) are found for the flights I, II, and III. One factor that is presumably making the difference between these two sets of flights is the fact that the Lagrangian simulations show a greater amount of filaments along the flight track during the set of flights that have lower correlation values (higher rms) than for the other set of flights. This means that there is an higher degree of uncertainty in the predictions when the Lagrangian reconstructions show that filaments are being sampled. Fairlie et al. (1997) obtain a similar result and relate higher (lower) rms differences to higher (lower) variability in the observed along-track tracer data and not to a deterioration (improvement) in forecast skill.

Figure 6 also shows that on a given flight section the correlation with the various reconstructed ozone data-sets may yield different results.

Nevertheless, when the correlation coefficients and rms differences are calculated for each set of six PV sources (derived from RDF, CAS and meteorological analyses)—each being calculated using ECMWF, Met Office, and NCEP analyses, and for each, source results are averaged over all 12 quasi-isentropic flight sections—final results show that no single PV data source yields a consistently higher average correlation coefficient (see Tables 4 and 5), and ANL never gives the lowest average cross correlation and is usually among the lowest rms difference. However the variation observed through each column does not appear to be statistically significant given the means and the standard deviations listed, making such an approach inconclusive.

Instead we go back to analyze further why the correlations are so different on the various flight sections, pursuing the idea that the presence of filaments may be the important factor in determining the degree of agreement between the reconstructed ozone value and the observations. Specifically the observed \(O_3\) profiles show possible filaments along the flight track in four cases and none in eight (see Table 3). As discussed in section 4a, as a possible filament finding in a flight dataset, we consider a variation in \(O_3\) that shows no direct relation to a corresponding variation in the \(\theta\) level of the flight, or that is greater than what is expected due to the \(\Delta O_3/\Delta \theta\) value estimated from flight data.

Figure 7 presents the correlation coefficients calcu-
lated for each PV source averaged separately over the two sets of flight sections with and without filaments.

The correlation coefficients over the sections without filaments (left-hand panel of Fig. 7) show values that range approximately between 0.55 and 0.75. The highest values are found for meteorological analyses and RDFST, while reconstructed values from RDFL and RDFSL yield lower mean correlations. This means that in the cases without filaments apparent in the observations, higher mean cross-correlation coefficients are found for the plain meteorological analyses than for the Lagrangian methods, with the unique exception of RDFST. The implication is that the Lagrangian methods are producing a great number of small-scale features that degrade the cross correlation with the observations.

The right-hand panel instead refers to the mean correlation coefficients averaged over the quasi-isentropic sections during which strong and rapid variations in the measured O3 are observed. Most striking is the significant decrease of the mean cross-correlation levels with respect to the left-hand panel, in fact here correlation values range approximately between 0.15 and 0.40. In addition, in this case all the Lagrangian products present higher correlation values than those of the meteorological analyses. More specifically, highest coefficients are found for all RDF reconstructed values, and in particular for RDFST and RDFSL.

Overall these results indicate a high sensitivity of the prediction skill to the local characteristics of the tracer field (i.e., the presence or absence of filaments). This suggests that it may be inappropriate, or at least partially misleading, to quantify the agreement between the measured and the predicted O3 profiles using a statistical index such as cross correlations or rms, which are sensitive to so-called “phase errors” in the predicted O3 distribution, that is, errors in the space–time localization of filamentary structures along a flight track.

For this purpose, following Fairlie et al. (1997), we have extended our statistical analysis applying the Kolmogorov–Smirnov (KS) test to yield a different kind of statistical diagnostic, which takes into account not so much the exact location of the filaments but the overall properties of the tracer distribution along the flight track. More specifically, the KS test uses the maximum difference between the cumulative frequency distributions of the observed and the predicted profiles as a measure
of their agreement (Sveshnikov 1968). Higher KS significance levels signifies better fits. In particular, a KS significance level of less than 0.05 means that the hypothesis that two distributions are identical must be rejected at the 0.05 significance level (less than 5% chance of rejecting a true null hypothesis). A KS significance level of 0.95 indicates a 95% probability that the two distributions are identical (less than 5% chance of accepting a false null hypothesis). To compute the KS statistic and significance levels, we follow Sveshnikov (1968).

Figure 8 shows the KS significance levels for each product calculated over each of the quasi-isentropic sections of the APE flights.

The general tendency is quite similar to that found for cross-correlation coefficients. Highest values characterize the January flights while the lowest ones are obtained for the December flights. Table 6 shows the same averaged results over the total set of flight sections as in Table 4 only for the KS significance levels and their standard deviations. Table 6 shows that KS values vary more depending on the kind of PV source than what had been found for the correlation coefficients reported in Table 4. The ANL runs suggest that NCEP is marginally the best PV source, but the Lagrangian runs suggest that the Met Office is the best all-purpose PV source (left panel in Fig. 9). The best single result is for RDFST using ECMWF, but this may be a function of the frequency of analyses rather than the quality of the PV analyses per se. However, one should be wary of being categorical about PV analyses based on such a short sample.

Similar to what has been done for the correlation coefficient calculations and presented in Fig. 7, we consider KS values calculated for each PV source averaged separately over the two sets of flight sections with and without filaments. The mean KS significance levels are presented for the different PV sources and for these two sets of flights in Fig. 9, respectively, right and left panels.

As was the case for the cross correlations, higher values of the KS statistic are found when averaging over flight segments without filaments (cf. left and right panels of Figs. 7 and 9). However, the KS statistic shows a greater amount of variability than the cross correlations, and this is particularly the case when filaments are present (right-hand panels of Figs. 7 and 9). The KS statistic for the ANL runs are very much smaller...
than those for the Lagrangian runs, when filaments are present. However, again as for the cross correlations (left-hand panel of Fig. 7), the ANL runs produce KS values higher than most of the Lagrangian runs for the flight segments when filaments are not apparent in the observations.

c. Discussion

The higher KS significance level found in the case of filaments for RDF_{ST} in the ECMWF-based calculations (see Fig. 9, right panel), in comparison to what is obtained with the other meteorological analyses, is most likely related to the higher temporal resolution of the ECMWF wind fields. These winds are given every 6 h instead that every 24 h, as for NCEP and Met Office analyses. Instead, the advantage of RDF_{ST} over the other RDF calculations, observed for the same case, could lie in the better spatial resolution of the initial field, that is, placing the parcels more accurately along the flight track. Here, O_3 profiles calculated from gridded output fields may suffer from the effect of spatial interpolation and the related truncation errors.

More generally, when comparing the KS significance levels with the other three RDF gridded fields RDF_{G}, RDF_{L}, and RDF_{SL}, it is apparent that no significant improvement is obtained by choosing as input data meteorological fields with greater spatial resolution. Quality of the KS significance level can be slightly enhanced by increasing the spatial resolution chosen for the final output Lagrangian products, especially in the case when filaments are detected along the track. In conclusion, and confirming what has been shown in previous studies (Mariotti et al. 1997; Waugh and Plumb 1994), the small-scale information missing in the plain meteorological analyses can be reconstructed using a wind time series long enough for advection, while the quality of the reconstruction does not appear to be very much sensitive to the spatial resolution chosen for the input/output fields.

Similarly to what has been said for RDF_{ST}, the CAS calculations based on ECMWF analyses should also benefit from higher temporal resolution wind fields, and this could explain the higher KS significance levels found in comparison with CAS calculations performed with the other analyses. The improvement of the tem-

Fig. 5. The predicted O_3 profiles and the 2-min-averaged ECOC O_3 data for the flight on 961223. Each panel shows the output obtained with a different product, based on ECMWF data.
poral resolution should play an even more important role in the CAS calculations. This is due to the fact that the CAS scheme is based on the piecewise uniform representation of the tracer (PV in this case), with regions of uniform PV separated by contours, each defined by a large number of nodes that are not independent of each other, but are dependent on the local curvature. This representation implies that if the winds at the position of each node are not very accurate, the errors tend to spread to nearby contour sections. The RDF gridded outputs suffer less from this kind of error because each gridded field is obtained by advecting simultaneously, but independently, thousands of parcels.

Our results demonstrate that, statistically, there is sometimes only marginal improvement in the reconstructed ozone fields using Lagrangian techniques rather than the plain meteorological analyses. This is particularly the case when filaments are not apparent in the observations (cf. left-hand panels of Figs. 6 and 8). Among the Lagrangian techniques, CAS does not appear to perform markedly better or worse than the RDF techniques: the cross correlation coefficient for CAS using

![Cross correlation coefficients between the 2-min-averaged ECOC O₃ and the predicted O₃ profiles for each product, over each of the quasi-isentropic sections of the APE flights (see Table 3 for acronyms) and for the three data sources used.](http://journals.ametsoc.org/doi/abs/10.1175/1520-0469(2002)059<1943:HRSTFR>2.0.CO;2)

Table 4. Mean cross-correlation coefficients and their std dev (in parentheses) between the estimated O₃ profiles obtained from the six products and ECOC O₃ data, for each meteorological source. The coefficients are averaged over the 12 quasi-isentropic sections of APE flights, using a simple arithmetic average, with all the sections of data equally weighted.

<table>
<thead>
<tr>
<th>Product</th>
<th>NCEP</th>
<th>Met Office</th>
<th>ECMWF</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANL</td>
<td>0.55 (0.11)</td>
<td>0.57 (0.09)</td>
<td>0.49 (0.11)</td>
</tr>
<tr>
<td>RDF₆₈</td>
<td>0.47 (0.12)</td>
<td>0.50 (0.12)</td>
<td>0.53 (0.07)</td>
</tr>
<tr>
<td>RDF₆₉</td>
<td>0.45 (0.10)</td>
<td>0.57 (0.06)</td>
<td>0.46 (0.07)</td>
</tr>
<tr>
<td>RDF₆₈</td>
<td>0.50 (0.08)</td>
<td>0.55 (0.06)</td>
<td>0.48 (0.06)</td>
</tr>
<tr>
<td>RDF₆₉</td>
<td>0.59 (0.08)</td>
<td>0.61 (0.06)</td>
<td>0.52 (0.09)</td>
</tr>
<tr>
<td>CAS</td>
<td>0.49 (0.08)</td>
<td>0.50 (0.10)</td>
<td>0.52 (0.08)</td>
</tr>
</tbody>
</table>

Table 5. Mean rms differences (10⁻³ ppmv) and their std dev (in parentheses) between the estimated O₃ profiles obtained from the six products and ECOC O₃ data, for each meteorological source. The coefficients are averaged over the 12 quasi-isentropic sections of APE flights, using a simple arithmetic average, with all the sections of data equally weighted.

<table>
<thead>
<tr>
<th>Product</th>
<th>NCEP</th>
<th>Met Office</th>
<th>ECMWF</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANL</td>
<td>26.09 (0.30)</td>
<td>27.08 (0.29)</td>
<td>26.22 (0.31)</td>
</tr>
<tr>
<td>RDF₆₈</td>
<td>26.75 (0.31)</td>
<td>27.31 (0.27)</td>
<td>26.98 (0.30)</td>
</tr>
<tr>
<td>RDF₆₉</td>
<td>25.92 (0.17)</td>
<td>28.60 (0.25)</td>
<td>27.67 (0.28)</td>
</tr>
<tr>
<td>RDF₆₈</td>
<td>26.35 (0.19)</td>
<td>28.49 (0.26)</td>
<td>27.31 (0.27)</td>
</tr>
<tr>
<td>RDF₆₉</td>
<td>25.65 (0.28)</td>
<td>26.90 (0.28)</td>
<td>25.70 (0.31)</td>
</tr>
<tr>
<td>CAS</td>
<td>27.19 (0.28)</td>
<td>25.81 (0.20)</td>
<td>27.83 (0.27)</td>
</tr>
</tbody>
</table>
ECMWF data is relatively high for flight segments with apparent filaments, but relatively low for flight segments without apparent filaments, and the KS significance levels are intermediate among those found for the other techniques for both sets of flight segments.

It has been noted how the general trend of Lagrangian products tend to introduce a great number of small-scale features in the reconstructed fields that degrade the statistical agreement with the observations, and this can be partly related to the effects of mixing processes.

In fact, the reconstructed ozone fields tested above differ in several respects, including their treatment of mixing. The meteorological analyses, by virtue of their grid size, have an implicit (numerical) diffusivity. This implicit diffusivity is two, or more, orders of magnitude greater than current estimates of the horizontal diffusivity in the lower stratosphere (Tan et al. 1998, and references therein). Moreover, the diffusivity in the real atmosphere is sporadic in time and space. The CAS simulations include an explicit diffusivity, set by the delta cut-off value (see section 4b above). The RDF techniques do not include diffusivity.

The relatively poor performance, with respect to the other gridded fields, of the RDF_L, RDF_Sl, and RDF_G techniques, for flight segments where filaments were not apparent suggests that mixing may not be negligible in these cases.

However, particularly as gauged by the KS statistic, CAS, which does include mixing, does little or no better than the RDF techniques for flight segments where filaments were not apparent. A more detailed examination of how to include mixing effects in techniques for the construction of forecast tracer fields when sporadic mixing events take place, must await further study.

Several other sources of errors can contribute to degrade the agreement of the Lagrangian products with the observations. The errors in the meteorological fields—winds and initial PV data used to label each parcel—seem to be an important source of uncertainty (e.g., Morris et al. 1995). Another factor, which is usually neglected for 5-day integrations, is the effect of diabatic cross-isentropic motion (Sutton et al. 1994; Schoeberl and Sparling 1995; Manney et al. 1995; Morris et al. 1998). This approximation is a potential, al-
though minor, source of additional inaccuracies. Numerical and computational errors can, in principle, also contribute to degrade the agreement between the observed and the predicted profiles, but Schoeberl and Sparling (1995) demonstrated that the magnitude of such errors is negligible compared to the other sources of error in these problems.

The fluctuations that could be related to a latitudinal or longitudinal shift of the fine-scale structures on the gridded fields (in other words, to an incorrect time-space localization) could be considered as an added source of error, only partially removed by the KS test.

All of these sources of errors can produce relevant effects on the space-time localization of high-resolution filamentary structures along a selected track.

5. Conclusions

Numerical experiments and statistical analysis are conducted to determine the skill of different Lagrangian techniques in improving the spatial resolution of given global tracer fields, using PV fields obtained from three meteorological agencies. Results from calculations show that the Lagrangian techniques are able to reproduce high-resolution tracer fields, increasing the presence of smaller-scale structures with respect to the meteorological analyses, thus giving at least a valid tool to visualize tracer fields. More difficult to define, from a statistical point of view, is the real improvement of the predicted profiles with respect to the observed one and so their agreement, because of the presence of several sources of error. This analysis shows that a part of the difficulty in demonstrating improvements comes from the fact that Lagrangian techniques seem to create a certain amount of small-scale filaments where none are seen in the data. Comparisons with data from such periods shows better fits using the plain meteorological

### Table 6

Mean KS significance levels and their std dev (in parentheses) between the estimated O₃ profiles obtained from the six products and ECOC O₃ data, for each meteorological source. The coefficients are averaged over the 12 quasi-isentropic sections of APE flights, using a simple arithmetic average, with all the sections of data equally weighted.

<table>
<thead>
<tr>
<th>Product</th>
<th>NCEP</th>
<th>Met Office</th>
<th>ECMWF</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANL</td>
<td>0.23 (0.08)</td>
<td>0.22 (0.06)</td>
<td>0.17 (0.04)</td>
</tr>
<tr>
<td>RDFₘ</td>
<td>0.16 (0.06)</td>
<td>0.24 (0.11)</td>
<td>0.17 (0.06)</td>
</tr>
<tr>
<td>RDFₘ⟲</td>
<td>0.14 (0.06)</td>
<td>0.20 (0.05)</td>
<td>0.10 (0.03)</td>
</tr>
<tr>
<td>RDFₘ⟲⟲</td>
<td>0.16 (0.06)</td>
<td>0.19 (0.08)</td>
<td>0.16 (0.08)</td>
</tr>
<tr>
<td>RDFₘ⟲⟲⟲</td>
<td>0.17 (0.02)</td>
<td>0.24 (0.09)</td>
<td>0.28 (0.13)</td>
</tr>
<tr>
<td>CAS</td>
<td>0.07 (0.01)</td>
<td>0.17 (0.07)</td>
<td>0.18 (0.07)</td>
</tr>
</tbody>
</table>
analysis than by Lagrangian techniques. The degree of parameterization of mixing processes in the Lagrangian techniques (i.e., RDF techniques completely neglect this effect) can be a partial explanation for this behavior. When considering data that contains clear evidence of filaments, the Lagrangian techniques show a better fit than the plain analyses. This better fit is most clearly seen when using the Kolmogorov–Smirnov test, rather than cross correlation.

Of the Lagrangian methods, RDF_{ST} (i.e., the Lagrangian not-gridded product based on high-resolution reverse trajectories calculated along a flight track) gives the best overall performance, as measured statistically by both cross-correlation and the KS significance test. Again, however, the advantages of RDF_{ST} are most clearly seen using the KS test.

The temporal rather than the spatial resolution of the meteorological fields used to advect tracers seems to favor an improvement of the agreement between the reconstructed and the measured tracer data. Such a result is not new, as far the CAS technique is concerned (Waugh and Plumb 1994), but from our analysis it seems to extend also to the other considered Lagrangian techniques.

In addition, the performance of RDF_{ST} is greatly improved by using the ECMWF analyses, which have a higher temporal resolution. This is also the case for CAS. However, it is not clear from our results how this effect cumulates with others, such as the increase in the spatial resolution of the gridded output Lagrangian fields.

In fact, the KS test also shows that RDF_{L}, using Met Office analyses, is about as good as RDF_{ST} and CAS overall.

It should also be considered that other differences besides temporal and spatial resolution may be present among the various meteorological analyses (which can influence the quality of the reconstructed fields) and it is in principle impossible to say what is responsible for the observed differences in the results. A more complete analysis of the relative influence of resolution of the input fields could be assessed by performing new tests using degraded spatial and temporal resolution of each single dataset, in a method similar to Waugh and Plumb.
(1994). Such an approach has not been considered in the framework of this paper, and it is left to future study.

Finally, the sole increase of the spatial resolution of the grided RDF output seems to have only a minor impact, with no real statistical significance, on the quality of the reconstructed fields, at least for the studied cases.

Previous studies have applied Lagrangian techniques to the predictions of constituent profiles along a flight track, based on PV analyses. In particular, Fairlie et al. (1997) conducted a somewhat similar analysis during the ASHOE–MAESA campaign, to quantify the level of predictive skill of analyzed and 5-day backward RDF PV in reconstructing nitrous oxide tracer structure in the Southern Hemisphere. Our simulations generally support some of their conclusions showing that, from a rigorous statistical point of view, high-resolution reconstructions show only a marginal improvement with respect to the plain meteorological PV. On the other hand, while Fairlie et al. (1997) found a general degradation in predictive skill when using RDF techniques, our results suggest instead that Lagrangian products very often improve the predictions when there is clear evidence of filaments in the measured data. Also, as already discussed in the introduction section, our study differs from Fairlie et al. (1997)—aside from the differences in the hemisphere, techniques, and tracer dataset chosen—especially for the aim of this paper which is mostly to select Lagrangian products that show the best overall performance. The current study, even if extended to various Lagrangian techniques and meteorological datasets, has been concerned exclusively with PV–$\mathrm{O}_3$ prediction based on meteorological analyses of the Northern Hemisphere, and its results should be considered only in this framework. A generalization of the results to all other situations is, in principle, limited by all the general circumstances not represented in this study (i.e., particular meteorological situations, hemisphere and period chosen, future evolution of quality and characteristics of the chosen dataset, etc.), and cannot be given with the available data. It may be anticipated that a more complete dataset, containing a greater number of in situ tracer observations, covering different hemispheres and winter periods, could help in this direction, but it is beyond the scope of our paper.

The understanding of a possible improvement in the performance of Lagrangian techniques with a detailed inclusion of mixing processes in the framework of calculation must await further study. However, our results may have some implications for the use of Lagrangian techniques in the analysis of data from airborne missions, such as APE–POLECAT and ASHOE–MAESA. Such techniques are useful, as has been shown by the KS statistic, for understanding the measurements once made. Finally some particular product, such as RDF$_{st}$ initialized with ECMWF data, gives a good overall performance and can also be considered as a visualization product for planning the mission itself, as was recently done for the 1999 Antarctic APE–Geophisica Aircraft in Antarctica (GAIA) campaign (Carli et al. 2000).

Acknowledgments. The authors want to thank NCEP/CPC team for permission to use NCEP data and the British Atmospheric Data Center for provision of Met Office data. Part of this work was funded by the European Communities under contract APE and by the Italian Space Agency (ASI). The authors also want to thank two anonymous reviewers for helpful comments and suggestions.

REFERENCES


Redaelli, G., 1997: Lagrangian techniques for the analysis of strato-


