A Cautionary Note on the Use of Meteorological Analysis Fields for Quantifying Atmospheric Mixing

A. STOHL AND O. R. COOPER

Cooperative Institute for Research in Environmental Sciences, University of Colorado/NOAA Aeronomy Laboratory, Boulder, Colorado

P. JAMES

Technical University of Munich, Munich, Germany

(Manuscript received 29 September 2003, in final form 14 January 2004)

ABSTRACT

Offline atmospheric transport models are normally driven with meteorological analyses. However, subsequent analysis fields are dynamically not consistent with each other, because they are produced in independent data assimilation cycles that lack strong dynamical constraints between each other. In this paper, it is shown that when these data are used with Lagrangian transport models, spurious mixing results from the dynamic inconsistencies. As a consequence, quantities such as potential vorticity or specific humidity that tend to be conserved along trajectories are found to be significantly less well conserved when analysis data are used than when forecast data are used for the trajectory calculations. This leads, for instance, to enhanced stratosphere–troposphere exchange. It is also shown that the dispersion of initially neighboring particles occurs more rapidly with the analysis than with the forecast data. It is therefore concluded that small-scale tracer structures develop too quickly in Lagrangian models, due to the inconsistencies between the driving wind fields.

1. Introduction

It is well known that with offline Eulerian transport models severe numerical problems can occur if there are inconsistencies in the driving meteorological data between the winds and the surface pressure tendencies. They can lead to the violation of mass conservation, artificial numerical diffusion and/or the generation of spurious tracer gradients (e.g., Jöckel 2001). In contrast, there is little awareness that similar problems also exist with Lagrangian models. In fact, Lagrangian models are often used to address problems for which Eulerian models are deemed not adequate because of their diffusive character and/or lack of resolution. Many of the Lagrangian model studies concentrate on mixing processes in the atmosphere, for which it is essential that any artificial mixing due to numerical problems is significantly less strong than actual atmospheric mixing. Lagrangian advection per se is nondiffusive and it may thus seem that, in principle, the accuracy of Lagrangian models in terms of mixing is limited only by the number of particles used (e.g., Norton 1994). However, as we shall show in this paper, the seemingly ideal mixing properties of Lagrangian models are spoiled when dynamically inconsistent wind fields are used.

Quantification of mixing between different air masses is a very important yet unsolved problem in the atmospheric sciences. The distribution of water vapor in the troposphere is sensitive to mixing (Pierrehumbert 1998), for instance, as is the strength of tracer gradients across frontal surfaces (Cohen and Kreitzberg 1997). Reliable quantification of atmospheric mixing is particularly important for atmospheric chemistry, because many chemical transformations depend nonlinearly on the concentrations of the reactants. This means that the chemical composition of neighboring air parcels with different composition, averaged over both parcels, depends on whether mixing between them occurs before or after chemical reactions take place (Edouard et al. 1996). The rate of mixing, thus, codetermines the chemical composition of the atmosphere. The rate of ozone loss in the stratospheric polar vortex, for instance, depends on mixing processes (Tuck et al. 2003). An example of an important mixing process is the exchange of air between the stratosphere and the troposphere (Stohl et al. 2003). Because of the very different chemical composition of these two atmospheric layers, the chemical reactions are sensitive to how fast stratospheric and tropospheric air masses mix with each other (Esler et al. 2001).

Corresponding author address: A. Stohl, Cooperative Institute for Research in Environmental Sciences, University of Colorado/NOAA Aeronomy Laboratory, R/AL4, room 2A101, 325 Broadway, Boulder, CO 80305.
E-mail: astohl@al.noaa.gov

Mixing is a process that involves a cascade from the largest down to the smallest atmospheric scales. The first step toward mixing is the generation of small-scale tracer structures by the stirring action of the large-scale winds, the so-called chaotic advection (Ottino 1989; Pierrehumbert and Yang 1993). During this step, an air parcel retains its chemical composition and, in the absence of friction and diabatic processes, its dynamical properties; that is, no mixing occurs on the molecular scale. Nevertheless this step is important, because the creation of small-scale structures increases the area/volume ratio of tracer filaments, thus creating an opportunity for molecular mixing across their surface. ''Real'' mixing (i.e., mixing that brings together chemical reactants at the molecular level) occurs by molecular diffusion on the boundaries of these filaments. Typically, turbulence will be an intermediate step between the large-scale stirring and the molecular-scale mixing, at least in the troposphere. Like stirring, turbulence does not really mix the air, but the scales of turbulence are so small that mixing via molecular diffusion will follow suit. All these processes together lead to a fractal tracer structure in the atmosphere (Tuck et al. 2003), which is currently not well understood.

To quantify atmospheric mixing, all three steps leading to it (stirring, turbulence, molecular diffusion) must be considered. Of these, molecular diffusion is in principle the easiest to determine, as it only depends on the thermodynamical properties of the air. However, it acts at spatial scales far below the resolution of atmospheric models, which produce artificial mixing before molecular diffusion becomes important. Turbulence is the most difficult to quantify, because it can be triggered by several different mechanisms (convection, breaking gravity waves, wind shear, radiation) and it is difficult to measure, because it occurs intermittently in relatively small localized patches that are hard to sample (e.g., Alisse et al. 2000). Furthermore, its small scale prohibits it from being resolved in global meteorological models.

Compared to turbulence, stirring might be thought of as being understood reasonably well. However, fundamental problems remain (Tuck et al. 2003), and this paper discusses one of them. Because Eulerian models can handle the creation of tracer filaments by chaotic advection only down to the resolution of their gridboxes, Lagrangian techniques have been developed that presumably are accurate to virtually any desirable scale. Some of them, like contour dynamics (Dritschel 1989), contour advection (e.g., Waugh and Plumb 1994) and reverse domain filling (Sutton 1994), have become very popular in stratospheric research (although some studies also question some of their aspects, e.g., Dragani et al. 2002). The prospect of capably resolving mixing on small scales has also spawned the development of Lagrangian chemistry transport models (McKenna et al. 2002). Other types of models capable of simulating fine-scale structures are Lagrangian particle dispersion models (e.g., Stohl et al. 1998), which simulate both stirring and turbulence. The success of any of these techniques, like their Eulerian counterparts, is largely reliant on the wind fields used. As we shall show below, using analyzed wind fields—the ones used most often—causes artificial mixing.

2. The problem

Most global tracer transport models are used with analyses from meteorological centers, for example, the European Centre for Medium-Range Weather Forecasts (ECMWF) or the National Centers for Environmental Prediction (NCEP). Analyses are produced in a process called data assimilation, which blends meteorological observations with a short-range forecast, yielding meteorological fields that are as close as possible to the observations and at the same time fulfill the dynamic model constraints. Different data assimilation techniques exist, the most advanced of which use variational algorithms (Courtier and Talagrand 1987). The variant used at ECMWF is four-dimensional (White 2002), where during a certain period of time (currently 12 h) the model forecast is adjusted iteratively to minimize a so-called cost function, a measure of the deviation between the forecast and the observations. During that period, the meteorological fields evolve in a dynamically consistent manner. Analyses are the result of the assimilation cycle and can later be used offline.

The time sequence of meteorological fields produced by a model forecast is also dynamically fully consistent. Even though after some time a forecast might differ substantially from the real state of the atmosphere because no more observations are assimilated, it still represents a physically sensible realization of how the atmosphere might have evolved. On the other hand, analysis fields taken from different assimilation cycles are dynamically not fully consistent with each other, because there is no dynamic constraint for the transition from one assimilation cycle to the next, except that a forecast from the previous cycle is passed on.

For some aspects of atmospheric transport, dynamic inconsistencies between subsequent analyses may be of relatively little concern, but not so for mixing. Dynamic flow features like the tropopause, which normally act as barriers to mixing, may be slightly displaced in an analysis relative to what a forecast from the previous analysis would have predicted. The reason for such a displacement is unphysical and comes about through the insertion of observational data. In a transport model used with these analyses, an air parcel may find itself on the other side of the tropopause without actually crossing it in a physically meaningful way, possibly leading to excessive mixing between the troposphere and the stratosphere.

As a side aspect of their study on trajectory accuracy, Stohl and Seibert (1998) compared the conservation of potential vorticity (PV) and specific humidity (q) along trajectories extending either from an analysis to the next available forecast field, or vice versa. Although their trajectories were only 3 h long, they found that both PV and q were significantly less well conserved when the calculations were done from a forecast to the next analysis field than when they were done from an analysis to the first forecast field. In this paper, we again contrast the conservation of dynamical tracers along trajectories calculated using analyses and forecasts, but using longer time periods and with the aim to explore atmospheric mixing. Furthermore, we directly compare the stirring properties of analyses and forecasts.

3. Method

Three-dimensional trajectories were calculated with the FLEXTRA model (Stohl et al. 1995; Stohl and Seibert 1998; Stohl et al. 2001) using ECMWF (White 2002) global meteorological fields (resolution $1° \times 1°$, 60 vertical levels). Two experiments were made, the first one using twelve 10-day forecasts from 0000 UTC at the first day of every month in the year 2002, and the second one using twelve corresponding 10-day series of analyses for the same time periods. Forecast fields were available from ECMWF every 3 h until 72 h into the forecast and every 6 h thereafter. Analyses were available only every 6 h, but were supplemented with 3-h forecasts during the first 72 h in order to have the same time resolution as the forecasts. This was done in order to achieve the highest possible time resolution, which is critical for trajectory accuracy (Stohl et al. 1995) and tracer conservation (Stohl and Seibert 1998). The insertion of the forecasts does not remove the inconsistencies in the analysis sequence, but condenses them into the 3-h periods from the 3-h forecasts to the next analyses. To explore whether this affected our results, we made two more experiments using only 6-hourly data.

Nine-day forward trajectories were started at 0300 UTC on the first day of every month. Starting locations were on a $1° \times 1°$ global grid every 50 hPa from 50 to 700 hPa, yielding a total of more than 10 million trajectories per experiment. Such a large number of trajectories was required because comparisons between forecasts and analyses can be made only on a statistical basis, not for individual trajectories, as the forecasts diverge from the analyses and thus represent slightly different meteorological situations. Potential vorticity, q and equivalent potential temperature Θ_e were calculated at every trajectory position and were written to output files every 24 h. These are quasi-conserved quantities, which can be used to tag air parcels and check the accuracy of trajectory calculations (Austin and Tuck 1985; Stohl and Seibert 1998). Potential vorticity is not conserved in the presence of friction or diabatic processes; q changes upon condensation, evaporation, and mixing (and slowly due to methane oxidation); and Θ_e is modified by diabatic heating and changes in q. All these processes are fastest in the planetary boundary layer, but above it they are often rather slow, particularly in the stratosphere. Apart from the physical processes, errors in the trajectory calculations (e.g., due to the interpolation of the winds; see Stohl 1998) will also lead to artificial nonconservation of the tracers along the trajectories, introducing additional noise. In the following, we assume that on average the above effects are the same for the two ensembles of forecast and analysis trajectories, and that any statistical differences between the two trajectory sets can be attributed to the dynamical inconsistencies of the latter. As a statistical parameter we use

$$\Delta T(t) = \sum_{i=1}^{N} |T_i(0) - T_i(t)|, \tag{1}$$

where T can be any of the tracers, t is time (from 0 to 9 days), and N is the total number of trajectories used for the comparison; ΔT was evaluated using forecast T values along the forecast trajectories and analysis T values along the analysis trajectories.

Now, how do errors in the conservation of the tracers relate to mixing? Consider the case of PV. If PV increases from a value smaller than 2 PVU (potential vorticity units, 1 PVU $= 1 \times 10^{-6}$ m^2 K kg^{-1} s^{-1}) to a value greater than 2 PVU this indicates transport from the troposphere to the stratosphere if the tropopause is defined as the 2-PVU surface (in the Northern Hemisphere). If the nonconservation of PV is greater for the analysis than for the forecast trajectories, stratosphere–troposphere exchange (STE) will be artificially enhanced in the former. We check this explicitly by calculating the number of the trajectories starting polewards of 30° that cross the 2-PVU surface. Similar to PV, the change with time of the other tracers also indicates the tendency of trajectories to cross airstream boundaries. Because of the wind shear across airstream boundaries, this likely also leads to enhanced stirring and subsequent mixing (further down the scale cascade) between contrasting air masses.

The separation over time of neighboring trajectories can be used to directly determine the intensity of stirring. The stronger the stirring, the stronger is the strain on particles that are initially located close to each other, and the faster these particles will separate. Lyapunov exponents, for instance, have often been used for quantifying how fast particles separate (e.g., Pierrehumbert and Yang 1993). Another parameter often used for related problems and which is also calculated for this study is the diffusion coefficient (e.g., Maryon and Buckland 1995; Pudykiewicz and Koziol 1998)

$$K(t) = \frac{1}{2}\frac{d\overline{r^2(t)}}{dt}, \tag{2}$$

where $\overline{r^2(t)}$ is the mean square separation of particles at time t. Mean square particle separations were calculated for every trajectory using the four initially neighboring trajectories; K, implying Gaussian turbu-

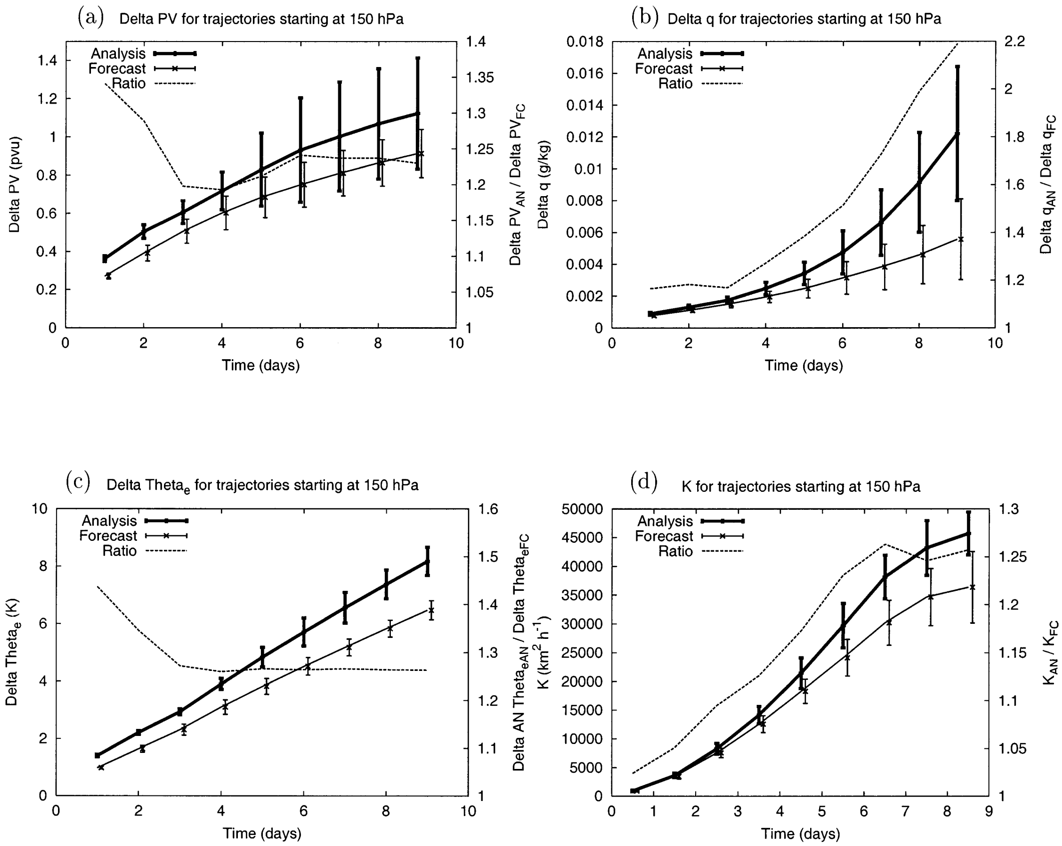

FIG. 1. (a) ΔPV, (b) Δq, (c) $\Delta\Theta_e$, and (d) K averaged over all trajectories starting from the 150-hPa level for the analysis trajectories (thick solid lines) and the forecast trajectories (thin solid lines). The dashed lines show the ratio between the analysis and the forecast values. Error bars give the standard deviations for the 12 individual analyses and forecast ensembles. Error bars for the forecast trajectories are offset by 0.1 day for clarity.

lence, is not a particularly good parameter to quantify mixing at large scales, where dispersion is not Gaussian. Therefore, the K values calculated here should be considered as statistical indicators of mixing, but not as physical parameters quantifying it.

4. Results

Figure 1a shows ΔPV calculated using Eq. (1) for all trajectories starting at the 150-hPa level, an altitude where diabatic effects on PV should be minimal and, thus, differences between the two experiments should be most clearly revealed. This is indeed the case: PV is significantly less well conserved along the analysis trajectories than along the forecast trajectories. For both experiments ΔPV increases with time, due to diabatic effects and trajectory errors. However, ΔPV is consistently higher by 20%–35% for the analysis trajectories than for the forecast trajectories. As the experimental setup was the same for the two experiments, except for the different types of meteorological data used, we trace this back to the dynamic inconsistencies between the analysis fields. The month-to-month variability of ΔPV was also higher for the analysis trajectories than for the forecast trajectories, indicating that PV conservation by the analysis trajectories was especially poor for particular months. We note, though, that ΔPV for the analysis trajectories exceeded ΔPV for the forecast trajectories for every individual month; that is, the analysis fields produced consistently more mixing than the forecast fields.

Looking at the average heating rates along the trajectories starting at 150 hPa and distinguishing between cases of heating and cooling, we find that the average 24-h heating rate was 0.8 K day^{-1} using the forecasts, but 1.4 K day^{-1} using the analyses, while cooling rates were 1.5 and 2.2 K day^{-1} for forecasts and analyses, respectively. Both heating and cooling rates were substantially larger using the analyses, which can be interpreted as the result of occasional artificial temperature "shocks" caused by data assimilation (Austin and Tuck 1985).

How representative is the 150-hPa level shown in Fig. 1a? For all individual starting levels, ΔPV was larger for the analysis than for the forecast trajectories. However, the relative differences between the two datasets decrease toward lower levels, and at 600 hPa they are only about 5%. This is to be expected because diabatic

TABLE 1. Number of STT and TST trajectories for the forecast (FC) and analysis (AN) datasets, and relative differences (%) between the two. The total number of trajectories starting poleward of 30° was 7.04 million.

Time	STT_{FC}	STT_{AN}	$\frac{\Delta STT}{STT_{FC}}$	TST_{FC}	TST_{AN}	$\frac{\Delta TST}{TST_{FC}}$
1 day	78 567	92 581	18%	85 042	94 025	11%
3 days	196 218	215 737	10%	154 056	163 655	6%
9 days	382 608	420 138	10%	180 572	191 199	6%

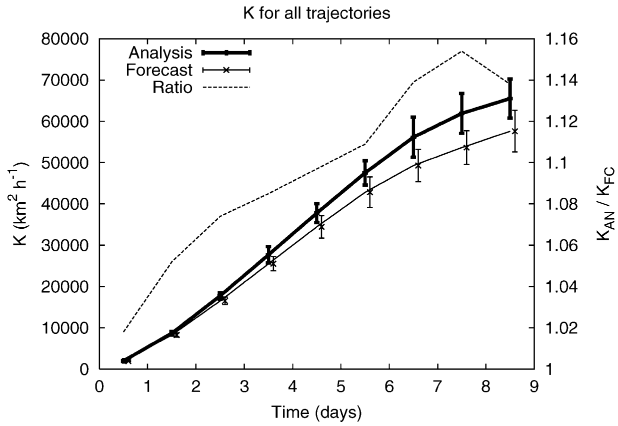

FIG. 2. Same as Fig. 1d, but for all trajectories.

processes, especially in the boundary layer, make PV a poor tracer in the lower troposphere, and the extra noise reduces the contrast between the two experiments. Averaged over all 14 starting levels from 50 to 700 hPa, relative ΔPV differences are 20% after 1 day, and 16% after 9 days.

STE estimates are known to be highly sensitive to errors in the meteorological data (e.g., Gettelman and Sobel 2000). In order to estimate what consequences the dynamical inconsistencies have for STE estimates, we calculated the number of trajectories starting poleward of 30° that crossed the 2-PVU surface, a commonly used definition for the position of the tropopause. We distinguished between stratosphere-to-troposphere transport (STT) and troposphere-to-stratosphere transport (TST), according to the definitions of Stohl et al. (2003). Table 1 shows the number of STT and TST events for time scales of 1, 3, and 9 days, and for the two experiments. Both STT and TST events occurred between 6% and 18% more frequently with the analysis than with the forecast trajectories. Notice that there is an asymmetry between STT and TST, with the frequency of STT events being overestimated more strongly than the number of TST events. Thus, Lagrangian STE estimates based on meteorological analysis data, which have become popular recently (e.g., Stohl 2001; Wernli and Bourqui 2002; James et al. 2003; Sprenger and Wernli 2003), will be biased high. Note, though, that Eulerian STE estimates using these data will suffer from similar, if not more severe, numerical problems.

In order to check whether the use of the 3-h forecasts that were embedded into the series of 6-hourly analysis fields caused any problems, we repeated the above analysis using only 6-hourly data. Values of ΔPV were about 10% higher both for the analyses and the forecasts, due to larger errors from temporal interpolation, but the relative differences between the two experiments were very similar to those reported above. As we found the same to be true for all other investigations, we do not further discuss these experiments.

Figure 1b shows Δq for trajectories starting at the 150-hPa level. As for PV, the analysis trajectories conserve q much less well than the forecast trajectories. Furthermore, the relative differences increase with travel time, from about 20% after 1 day to 220% after 9 days. Such a behavior would indeed be expected for a good tracer, indicating that the series of inconsistencies generated upon every step from analysis to analysis add up. Values of Δq are larger for the analysis than for the forecast trajectories at all altitudes. Relative differences decrease with increasing pressure, but even at 700 hPa, systematic differences of 3%–11% remain.

Figure 1c shows $\Delta\Theta_e$ for trajectories starting at 150 hPa. A similar picture emerges as with PV: the analysis trajectories show less conservation and the relative differences decrease from 44% to 26% as travel time increases from 1 to 9 days. The relatively small error bars show that there is relatively little variation in mean $\Delta\Theta_e$ between the 12 months, with the analysis trajectories being consistently higher. In contrast to PV and q, relative differences between forecasts and analyses have a minimum of 0% to 10% in the middle troposphere (450–500 hPa) for $\Delta\Theta_e$.

The results obtained for the conservation of PV, q, and Θ_e all point toward enhanced mixing when using the analysis fields. But can this indeed be related to artificially enhanced stirring? Figure 1d shows the diffusion coefficient K for the 150-hPa starting level. Values of K increase with travel time and are of a similar magnitude as those reported in previous studies (e.g., Maryon and Buckland 1995; Pudykiewicz and Koziol 1998). Analogous to the tracer studies, K values are larger for the analysis trajectories than for the forecast trajectories, with relative differences increasing from about 2% to 26% from days 1 to 9. The relative differences are largest at the two uppermost levels, where they exceed 35%, and decrease to 8.5% at 700 hPa on day 9. Figure 2 shows K values averaged over all trajectories. As for the single 150-hPa level (Fig. 1d), K increases faster with time for the analysis than for the forecast trajectories. Especially at the higher levels, the differences are substantial, indicating a much stronger strain on neighboring particles when analyses instead of forecasts are used. Qualitatively the same general result is obtained when using Lyapunov exponents, which are not reported here. The fact that differences increase with time is especially worrisome.

Table 2. Values of ΔPV, Δq, and $\Delta\Theta_e$ for forward trajectories extending from analyses to 3-h forecasts (an2fc), backward trajectories from the 3-h forecasts to the analyses (fcback2an), and forward trajectories extending from the 3-h forecasts to the next analyses (fc2an). All trajectories were starting at 150 hPa. The error ranges gives the standard deviations obtained for the monthly experiments.

Experiment	Dynamical property	ΔPV (PVU)	Δq (g kg^{-1})	$\Delta\Theta_e$ (K)
an2fc	Consistent	0.12 ± 0.01	$0.23 \pm 0.02 \times 10^{-3}$	0.33 ± 0.01
fcback2an	Consistent	0.12 ± 0.01	$0.23 \pm 0.02 \times 10^{-3}$	0.34 ± 0.01
fc2an	Inconsistent	0.20 ± 0.01	$0.32 \pm 0.03 \times 10^{-3}$	0.59 ± 0.02

So far we have not yet addressed the critical question of how the mixing occurring with the analyses and forecasts compares to reality. Is the enhanced mixing found for the analyses overestimating real mixing? Using physical reasoning, we have argued in section 2 that we expect artificial mixing from the dynamical inconsistencies of the analysis fields, which, among other effects, cause airstream boundaries (e.g., the tropopause) to have unphysical position changes superimposed over their real physical motion. Due to these position changes, an air parcel may find itself at the other side of an air stream boundary without actually physically crossing it. Since there is normally wind shear across airstream boundaries, this will also lead to enhanced dispersion. But is the enhanced dispersion we have found in our study really due to this unphysical effect, or could it be overwhelmed by other effects that also lead to enhanced mixing when using the analyses? Processes that are missing in the forecast model could be introduced indirectly into the analyses through the assimilation of observation data. If these processes are physical and lead to enhanced mixing that is stronger than the extra mixing due to dynamical inconsistencies, the strength of mixing when using the analyses may indeed be closer to reality than the strength of mixing when using the forecasts. However, great care is taken in the data assimilation to filter out all processes that are unphysical from the model's point of view, because such processes would cause severe numerical problems with the forecast. Therefore, it is not clear how extra mixing could result from processes that are inserted indirectly by the observations and which are reflected in the analysis, but are not operative in the model forecast.

More can be learned by investigating when exactly the extra mixing occurs. For this, we calculated three sets of 3-h trajectories: dynamically inconsistent forward trajectories from a 3-h forecast to the next analysis (experiment ‘‘fc2an’’), dynamically consistent backward trajectories from the same forecast to the previous analysis (‘‘fcback2an’’), and dynamically consistent forward trajectories from this previous analysis to the forecast (‘‘an2fc’’). Apart from the shorter duration of the trajectories, the setup for these experiments was the same as for those discussed previously.

Diffusion coefficient K was almost identical for the three experiments, because the 3-h time period is too short to yield meaningful values, given the relatively large initial separation of trajectories on the starting grid, but tracer conservation was very different. Table 2 reports the results for the 150-hPa starting level (results for other levels were similar). Because of the reversibility of trajectory calculations, the results for the experiments an2fc and fcback2an are almost identical. These experiments differ only by where the trajectories were started (for the forward calculations from the regular grid at the analysis time; for the backward calculations from the regular grid at the forecast time). All three quantities, PV, q, and Θ_e, are conserved significantly less well for the inconsistent experiment fc2an, in agreement with previous results (Stohl and Seibert 1998). This confirms that the larger conservation errors in the analysis sequence are indeed caused by dynamical inconsistencies. Furthermore, this rules out the possibility that they are due to differences in the wind field characteristics. A reviewer of this paper has suggested that the analyses may contain more finescale structures in the wind fields due to the assimilation of observation data and that this may cause the different mixing for forecasts and analyses. On the basis of the above results, however, this can be ruled out, because all of the above experiments were based on one analysis and one forecast field. If only the field characteristics at a particular time were responsible, the three quantities should have been conserved equally well in the three experiments.

5. Discussion and conclusions

In this paper we have shown that three tracers (PV, q, and Θ_e) are consistently less well conserved along trajectories calculated using meteorological analysis data than along trajectories calculated using forecast data. Furthermore, we found that diffusion coefficients and Lyapunov exponents are also substantially larger with the analysis data than with the forecast data. As the model setup (trajectory model, time and space resolution of the input data, etc.) was identical for the two experiments, we trace this back to the dynamic inconsistencies that exist between subsequent analysis fields. Three-hour trajectories calculated in a dynamically consistent way (from an analysis to a forecast) and in an inconsistent way (from a forecast to a subsequent analysis) showed large differences in tracer conservation. This confirms that, indeed, dynamical inconsistencies are responsible for the relatively poorer tracer conservation when using the analysis fields.

STE was also enhanced using the analysis fields relative to using the forecast fields, a direct consequence of PV being less well preserved in the analysis experiment. This adds to other difficulties with estimating STE using meteorological analysis data (Gettelman and Sobel 2000). For the same reason, we also expect enhanced mixing across other barriers to mixing such as the stratospheric polar vortex (e.g., Schoeberl and Newman 1995) when Lagrangian techniques are used in combination with analysis fields compared to when they are used with forecast fields.

While the differences were largest at higher atmospheric levels, this does not necessarily imply that the inconsistencies are less important at lower levels. The tracers we have used are all not well conserved in the presence of diabatic effects which introduces a high level of noise in the conservation statistics, reducing the differences between the two experiments in the lower troposphere. However, the time scales of mixing in the lower troposphere are much shorter than in the upper troposphere or lower stratosphere, because of convective activity, boundary layer turbulence, etc., effects which we have not considered here. This may reduce the significance of the enhanced stirring, because tracer filaments are destroyed by these processes before they can cascade down to the smallest scales. In the stratosphere, in contrast, the time scale of mixing is very long, because turbulence occurs only intermittently and there is no convection. This means that chaotic advection can act for a longer time than the 9 days we have considered here before filaments are destroyed. A relatively large part of the small-scale tracer structures obtained with Lagrangian models may therefore be artificial, not only because of erroneous spatial locations of tracer filaments (e.g., Dragani et al. 2002), but also statistically in terms of the richness of the small-scale structure. This in turn may lead to systematic biases in Lagrangian chemistry models. Another consequence of this is that turbulent diffusion parameterizations for Lagrangian models (e.g., Legras et al. 2003) should be different for analysis than for forecast fields.

It is important to emphasize that our results do not imply that model calculations based on forecast fields are to be preferred over calculations based on analysis fields. Many transport processes have time scales of 10 days or more, and forecasts cannot normally be said to be even close to actual meteorological conditions after such a long time. Therefore, there is typically no other choice than using analysis data for studying atmospheric transport processes. But this note shows that while analyses remain the preferable data source, transport calculations based on these data also have the undesirable property of artificially enhanced mixing. For some studies, this may be of relatively little concern, but for others it is important that modelers are alert to the possibility that their results may tend to overestimate atmospheric mixing, at least as depicted by the underlying meteorological model. As not all processes leading to mixing are captured by a limited-resolution model, the artificial mixing caused by dynamic inconsistencies may partly compensate for unresolved real mixing. However, the characteristics of unresolved mixing processes may be entirely different from resolved or artificial mixing, which perhaps explains why the fractal distribution of tracers in the atmosphere is currently not well understood (Tuck et al. 2003).

While this paper cautions the interpretation of results from Lagrangian models, it does not question the superiority of Lagrangian models over Eulerian models for quantifying mixing. Similar effects occur in Eulerian models and the resulting numerical problems may in fact be even more severe if not dealt with carefully (e.g., Bregman et al. 2003). However, it was shown that Lagrangian models are not free from artificial mixing either, although this expresses itself in the artificially rapid creation of small-scale structures rather than their diffusive destruction, as more likely would be the case in Eulerian models.

Is there a solution to these problems in sight? Unfortunately, not. The only solution these authors can envisage is the continuous application of four-dimensional variational data assimilation over the same time periods as the Lagrangian models are applied. Ideally, the Lagrangian calculations would be done online with the data assimilation, because the interpolation between grid points in space and time causes similar effects as the dynamical inconsistencies between the analyses. It will take a long time, though, before four-dimensional variational data assimilation becomes feasible over time periods much longer than a few days—if it can ever be realized.

Acknowledgments. We thank A. Tuck and three reviewers for their valuable comments on previous versions of this manuscript, and E.-Y. Hsie for support with our computer environment. ECMWF and the German Weather Service are acknowledged for permitting access to the ECMWF archives. Part of this research was supported by the German Federal Ministry for Education and Research within the Atmospheric Research Program 2000 (ATMOFAST project).

REFERENCES

Alisse, J. R., P. H. Haynes, J. Vanneste, and C. Sidi, 2000: Quantification of stratospheric mixing from turbulence microstructure measurements. *Geophys. Res. Lett.,* **27,** 2621–2624.

Austin, J., and A. F. Tuck, 1985: The calculation of stratospheric air parcel trajectories using satellite data. *Quart. J. Roy. Meteor. Soc.,* **111,** 279–307.

Bregman, B., A. Segers, M. Krol, E. Meijer, and P. van Velthoven, 2003: On the use of mass-conserving wind fields in chemistry-transport models. *Atmos. Chem. Phys.,* **3,** 447–457.

Cohen, R. A., and C. W. Kreitzberg, 1997: Airstream boundaries in numerical weather simulations. *Mon. Wea. Rev.,* **125,** 168–183.

Courtier, P., and O. Talagrand, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. II: Numerical results. *Quart. J. Roy. Meteor. Soc.,* **113,** 1329–1347.

Dragani, R., R. Redaelli, G. Visconti, A. Mariotti, V. Rudakov, A. R. MacKenzie, and L. Stefanutti, 2002: High-resolution stratospheric tracer fields reconstructed with Lagrangian techniques: A comparative analysis of predictive skill. *J. Atmos. Sci.,* **59,** 1943–1958.

Dritschel, D. G., 1989: Contour dynamics and contour surgery: Numerical algorithms for extended, high-resolution modelling of vortex dynamics in two-dimensional, inviscid, incompressible flows. *Comput. Phys. Rep.,* **10,** 77–146.

Edouard, S., B. Legras, F. Lefevre, and R. Eymard, 1996: The effect of mixing on ozone depletion in the Arctic. *Nature,* **384,** 444–447.

Esler, J. G., D. G. H. Tan, P. H. Haynes, M. J. Evans, K. S. Law, P. H. Planterin, and J. A. Pyle, 2001: Stratosphere–troposphere exchange: Chemical sensitivity to mixing. *J. Geophys. Res.,* **106,** 4717–4731.

Gettelman, A., and A. H. Sobel, 2000: Direct diagnoses of stratosphere–troposphere exchange. *J. Atmos. Sci.,* **57,** 3–16.

James, P., A. Stohl, C. Forster, S. Eckhardt, P. Seibert, and A. Frank, 2003: A 15-year climatology of stratosphere–troposphere exchange with a Lagrangian particle dispersion model: 2. Mean climate and seasonal variability. *J. Geophys. Res.,* **108,** 8522, doi:10.1029/2002JD002639.

Jöckel, P., R. von Kuhlmann, M. G. Lawrence, B. Steil, C. A. M. Brenninkmeijer, P. J. Crutzen, P. J. Rasch, and B. Eaton, 2001: On a fundamental problem in implementing flux-form advection schemes for tracer transport in 3-dimensional general circulation and chemistry transport models. *Quart. J. Roy. Meteor. Soc.,* **127,** 1035–1052.

Legras, B., B. Joseph, and F. Lefévre, 2003: Vertical diffusivity in the lower stratosphere from Lagrangian back-trajectory reconstructions of ozone profiles. *J. Geophys. Res.,* **108,** 4562, doi:10.1029/2002JD003045.

Maryon, R. H., and A. T. Buckland, 1995: Tropospheric dispersion: The first ten days after a puff release. *Quart. J. Roy. Meteor. Soc.,* **121,** 1799–1833.

McKenna, D. S., P. Konopka, J.-U. Grooβ, G. Günther, R. Müller, R. Spang, D. Offermann, and Y. Orsolini, 2002: A new chemical Lagrangian model of the stratosphere (CLaMS) 1. Formulation of advection and mixing. *J. Geophys. Res.,* **107,** 4309, doi:10.1029/2000JD000114.

Norton, W. W., 1994: Breaking Rossby waves in a model stratosphere diagnosed by a vortex-following coordinate system and a technique for advecting material contours. *J. Atmos. Sci.,* **51,** 654–673.

Ottino, J., 1989: *The Kinematics of Mixing: Stretching, Chaos and Transport.* Cambridge University Press, 364 pp.

Pierrehumbert, R. T., 1998: Lateral mixing as a source of subtropical tropospheric water vapor. *Geophys. Res. Lett.,* **25,** 151–154.

——, and H. Yang, 1993: Global chaotic mixing on isentropic surfaces. *J. Atmos. Sci.,* **50,** 2462–2480.

Pudykiewicz, J. A., and A. S. Koziol, 1998: An application of the theory of kinematics of mixing to the study of tropospheric dispersion. *Atmos. Environ.,* **32,** 4227–4244.

Schoeberl, M. R., and P. A. Newman, 1995: A multiple-level trajectory analysis of vortex filaments. *J. Geophys. Res.,* **100,** 25 801–25 815.

Sprenger, M., and H. Wernli, 2003: A northern hemispheric climatology of cross-tropopause exchange for the ERA15 time period (1979–1993). *J. Geophys. Res.,* **108,** 8521, doi:10.1029/2002JD002636.

Stohl, A., 1998: Computation, accuracy and applications of trajectories—A review and bibliography. *Atmos. Environ.,* **32,** 947–966.

——, 2001: A 1-year Lagrangian climatology of airstreams in the Northern Hemisphere troposphere and lowermost stratosphere. *J. Geophys. Res.,* **106,** 7263–7279.

——, and P. Seibert, 1998: Accuracy of trajectories as determined from the conservation of meteorological tracers. *Quart. J. Roy. Meteor. Soc.,* **124,** 1465–1484.

——, G. Wotawa, P. Seibert, and H. Kromp-Kolb, 1995: Interpolation errors in wind fields as a function of spatial and temporal resolution and their impact on different types of kinematic trajectories. *J. Appl. Meteor.,* **34,** 2149–2165.

——, M. Hittenberger, and G. Wotawa, 1998: Validation of the Lagrangian particle dispersion model FLEXPART against large scale tracer experiment data. *Atmos. Environ.,* **32,** 4245–4264.

——, L. Haimberger, M. P. Scheele, and H. Wernli, 2001: An intercomparison of results from three trajectory models. *Meteor. Appl.,* **8,** 127–135.

——, and Coauthors, 2003: Stratosphere–troposphere exchange: A review, and what we have learned from STACCATO. *J. Geophys. Res.,* **108**, 8516, doi:10.1029/2002JD002490.

Sutton, R., 1994: Lagrangian flow in the middle atmosphere. *Quart. J. Roy. Meteor. Soc.,* **120,** 1299–1331.

Tuck, A. F., S. J. Hovde, R.-S. Gao, and E. C. Richard, 2003: Law of mass action in the Arctic lower stratospheric polar vortex January–March 2000: ClO scaling and the calculation of ozone loss rates in a turbulent fractal medium. *J. Geophys. Res.,* **108,** 4451, doi:10.1029/2002JD002832.

Waugh, D. W., and R. A. Plumb, 1994: Contour advection with surgery: A technique for investigating finescale structure in tracer transport. *J. Atmos. Sci.,* **51,** 530–540.

Wernli, H., and M. Bourqui, 2002: A Lagrangian ''1-year climatology'' of (deep) cross-tropopause exchange in the extratropical Northern Hemisphere. *J. Geophys. Res.,* **107,** 4021, doi:10.1029/2001JD000812.

White, P. W., Ed., 2002: IFS documentation. ECMWF Tech. Rep. [Available online at www.ecmwf.int.]