Statistical Perspectives on Stratospheric Transport

L. C. Sparling
Goddard Space Flight Center, Greenbelt MD and
Joint Center for Earth Systems Technology, University of Maryland Baltimore County,
Baltimore, MD

Short title: SPARLING: STATISTICAL PERSPECTIVES ON STRATOSPHERIC TRANSPORT

Abstract.

Long-lived tropospheric source gases, such as nitrous oxide, enter the stratosphere through the tropical tropopause, are transported throughout the stratosphere by the Brewer-Dobson circulation, and are photochemically destroyed in the upper stratosphere. These chemical constituents, or "tracers" can be used to track mixing and transport by the stratospheric winds. Much of our understanding about the stratospheric circulation is based on large scale gradients and other spatial features in tracer fields constructed from satellite measurements. The point of view presented in this paper is different, but complementary, in that transport is described in terms of tracer probability distribution functions (PDFs). The PDF is computed from the measurements, and is proportional to the area occupied by tracer values in a given range. The flavor of this paper is tutorial, and the ideas are illustrated with several examples of transport-related phenomena, annotated with remarks that summarize the main point or suggest new directions. One example shows how the multimodal shape of the PDF gives information about the different branches of the circulation. Another example shows how the statistics of fluctuations from the most probable tracer value give insight into mixing between different regions of the atmosphere. Also included is an analysis of the time-dependence of the PDF during the onset and decline of the winter circulation, and a study of how "bursts" in the circulation are reflected in transient periods of rapid evolution of the PDF. The dependence of the statistics on location and time are also shown to be important for practical problems related to statistical robustness and satellite sampling. The examples illustrate how physically-based statistical analysis can shed some light on

aspects of stratospheric transport that may not be obvious or quantifiable with other types of analyses. An important motivation for the work presented here is the need for synthesis of the large and growing database of observations of the atmosphere and the vast quantities of output generated by atmospheric models.

1. Introduction

For a comprehensive view of how the atmosphere works and how it responds to anthropogenic forcing over a range of length and time scales, we must look at both observational and model data from many different perspectives, using a variety of physically relevant analysis tools. Common among these are field maps that emphasize the patterns in field contours, the zonal mean structure of the field, and mean vertical profiles (see, e.g. *UARS special validation issue*). There is no doubt about the value of these pictures in developing an understanding about stratospheric transport, particularly the spatial structure of a "tracer", a dynamical or chemical quantity that is conserved on the transport timescales of interest.

Tracer-rich air enters the stratosphere in the tropics where it is lofted to higher altitudes by the upwelling branch of the circulation. At high altitudes, the air becomes depleted of tracer through photochemical loss, and tracer-poor air is then carried poleward and downward in the extratropics by the downwelling branch of the circulation. The tracer field develops a generic zonal mean structure, an example of which is shown in Figure 1 for the chemical tracer N_2O. The figure is based on measurements taken during northern hemisphere (NH) winter with the Cryogenic Limb Array Etalon Spectrometer (CLAES) on the Upper Atmosphere Research Satellite(UARS)[*Roche, et al*, 1996]. Large scale chemical differentiation is apparent from the closely spaced contours in the subtropics and at the edge of the polar vortex. While the large scale chemical differentiation has its origins in the overall circulation together with the distribution of sources and sinks, it is maintained because the timescale for global scale horizontal mixing is slow relative to the timescale for vertical advection. This "chemical labelling" of the tropical, midlatitude and polar air masses allows us to track the transport between these regions over a wide range of scales, and it underlies many of the ideas presented here.

Pictorial representations of spatial variability are important for developing a qualitative understanding of the gross features of constituent transport, but there are many problems that require a more quantitative approach. For example, to what extent can it be said that air masses are chemically distinct and how do we define boundaries between them? What signatures of mixing between air masses can be identified in the data and how can they be quantified? There has been much recent interest in these issues [e.g. *Volk et al.*, 1996; *Hitchman et al.*, 1994] not only because of their importance for an understanding of stratospheric transport in general, but also because they are relevant to a number of important current problems such as midlatitude ozone loss and the effects of aircraft on the chemical balance of the atmosphere [*Stolarski et al.*, 1995].

There are also practical problems related to the collection and analysis of observations of the atmosphere and their comparison to models. Since models cannot be expected to reproduce all of the structural details in a tracer field, how then can comparisons be made in a way that is independent of the rapid and reversible deformations of field contours? How do we compare sets of observations that are not colocated, or that have different spatial resolutions or different spatial and temporal coverage? When is a collection of data statistically robust, and what can reasonably be inferred about global scale transport from a sparsely sampled field? These problems require a quantitative, statistical approach to data and model analysis, and while we cannot address all of these questions here, we will suggest ways of looking at them.

The statistical point of view adopted here is based on an analysis of P(χ), the area-weighted probability distribution function (PDF) of a discrete sampling of a chemical or dynamical field $\chi(\phi, \lambda)$ at a single vertical level. The area weighting takes into account the typically nonuniform spatial sampling of data points. P(χ) is the likelihood of observing a field value in the range χ to $\chi + \Delta\chi$ in a given region of the atmosphere, and is normalized so that $\Sigma_i P(\chi_i) = 1$.

An equivalent, but more geometric, point of view is to consider P(χ) as proportional to the area enclosed by the χ and $\chi + \Delta\chi$ contours in a contour plot of the field. If there is no dissipation and if the field is conserved, then a non-divergent flow will deform the field contours but will leave P(χ) invariant. This idea was used by *Yang* [1995] who investigated the time evolution of potential vorticity and tracer PDFs in an effort to isolate nonconservative processes in a dynamical model. Meridional coordinates based on the area enclosed by chemical tracer or potential vorticity contours on isentropic surfaces have also been used to study non-conservative or dissipative processes in various formulations and applications of the Lagrangian mean idea (see, e. g. *McIntyre*, 1980; *Dunkerton et al*, 1981; *Butchart and Remsberg*, 1986; *Hess*, 1991; *Schoeberl and Lait*, 1991; *Nakamura*, 1996). Dissipative effects can be included, as in the "modified Lagrangian mean" (MLM) approach [*McIntyre*, 1980], which includes dissipation in the form of a strain-enhanced diffusion [*Nakamura*, 1996]. These ideas are based on the fact that while the contours of a quasi-conserved field rapidly change shape, the area enclosed by them changes much more slowly. The effects of reversible and rapidly varying horizontal advection can therefore be removed by transformation to a conserved coordinate, and dissipative processes can be isolated.

The area coordinate $\mathcal{A}(\chi)$ in Lagrangian mean formulations is based on the total area enclosed by contours, while the PDF $\mathrm{P}(\chi)$ is the area between contours separated by a small interval $\Delta\chi$, thus $\mathrm{P}(\chi) = \mathcal{A}(\chi + \Delta\chi) - \mathcal{A}(\chi)$. Equivalently, $\mathcal{A}(\chi)$ is the cumulative PDF, obtained by summing over all bins with values greater than or equal to χ

$$\mathcal{A}(\chi) = \sum_{\chi' \geq \chi} P(\chi'). \tag{1}$$

In one of the earliest applications of the conserved coordinate idea, *Butchart and Remsburg* [1986] investigated the dynamics of the polar vortex by looking at the time evolution of the vortex area. They defined the "equivalent latitude" based on the area within a potential vorticity contour, an idea that has been particularly useful in looking at processes in the winter polar regions. The equivalent latitude ϕ_{eq} can be defined for any field that varies monotonically with latitude in a mean sense and is defined by $\phi_{eq} \equiv sin^{-1}(\mathcal{A}(\chi))$.

The horizontal wind in the stratosphere is non-divergent to a good approximation [*Koshyk et al.*, 1999], so the global PDF of a long-lived tracer on a single vertical level evolves in time mainly from the combined effects of vertical advection and mixing. A global PDF, however, has the disadvantage in that it combines the meridional variability of the northern and southern hemispheres. For most kinds of transport problems, it is more useful to compute the PDFs separately in each hemisphere. The hemispheric PDF will also evolve from cross-equatorial transport, but on long timescales. The distribution in a smaller geographical region such as a latitude band on the other hand, can change significantly from day to day as air with different characteristics is horizontally advected through the region.

The "statistical view" of transport presented here means only a point of view in which the PDF of a tracer or dynamical field is the main object of investigation. The intermittency and large-scale non-homogeneity in atmospheric dynamical processes gives rise to non-Gaussian variability which cannot be described solely by the first and second moments of the distribution, thus the entire distribution needs to be considered. The emphasis here is not specifically on separating conservative from non-conservative processes. We are interested instead in more general properties like the shape of the PDF, for example whether it is multimodal, symmetric, or has long tails, and what this might tell us about the mean and eddy transport.

Rather than focus on a particular science question, the point of this paper is to illustrate the applications of PDFs to transport-related problems through a variety of examples. The examples are based on well-known middle atmosphere phenomena, and are annotated with remarks that emphasize the main idea, point out how the ideas can be applied in other contexts, or suggest new directions. Section 2 describes the tracer data used, the computation of area weighted PDFs, and also includes an interpretation of the structure of the PDF in terms of a χ vs. ϕ scatter plot of tracer observations. Section 3 illustrates several statistical aspects of tracer variability in transition regions between air masses, using the subtropical boundary as an example, and Section 4 discusses the statistical properties of fluctuations about the most probable tracer value. A view of the time evolution of the PDF during an early winter warming and the evolution of the PDF throughout northern winter 92/93 is presented in Section 5. Here the idea of chemical entropy is introduced to describe the partitioning of the winter hemisphere into distinct air masses. Another dynamical example summarizes the tracer variability at the end of winter, demonstrating how polar vortex breakup and trapping in the Aleutian high can be summarized by the structure and time evolution of the PDF. Section 6 deals with statistical robustness, with examples that show how the tracer PDF survives a drastic reduction in data, and also includes a discussion of homogeneity and stationarity. Section 7 considers PDFs of gridded meteorological data, illustrated by a comparison of two potential vorticity analyses. We close with an overview, some caveats and some ideas for future directions.

2. Data, calculations and notation

In this paper, the emphasis is on tracer PDFs computed from satellite observations having global coverage. The general ideas however, can be applied to any set of observations provided that any geographically biased sampling is taken into account when interpreting the results or comparing with other data or models.

For simplicity, most of the ideas will be illustrated with a single data set, the

CLAES V. 7 measurements of the long-lived stratospheric tracer N_2O [*Roche, et al*, 1996]. The lifetime of N_2O is several months below 40km [*Ko et al*, 1991]. Figure 2 shows the nearly uniform areal coverage of measurements accumulated over 4 days, with the sampling pattern for a single day highlighted. The dense spatial sampling is not instantaneous, but this is not crucial for fields that are conserved over a few days. PDFs of a conserved quantity based on several days of data can be thought of as averages over daily realizations. A yaw manuever every 36 days keeps the same side of the spacecraft away from the sun for instrument and solar array purposes. Figure 2 shows the coverage for a north-facing period, with latitude range 35S to 80N; the sampling pattern is reversed when the satellite faces south. The hole in the observation field from 80° to the pole is unfortunate but is only 1.5% of the area of the hemisphere. Because of the limited vertical resolution ($\sim$ 2.5 km), most of the results in this paper use the data on the reported UARS pressure levels, rather than interpolating to potential temperature surfaces. The main results presented here are similiar on pressure and isentropic surfaces.

Given a set of observations χ_j of a tracer field $\chi(\phi, \lambda)$ that has a nearly equal area spatial distribution, the PDF $P(\chi)$ can be approximated as the fraction of observations having mixing ratios χ_j in the range χ to $\chi + \Delta\chi$. When the observations are sufficiently dense, this will be a good approximation to the fractional area enclosed by the contours χ and $\chi + \Delta\chi$ of the tracer field. Satellite measurements typically do not have an equal area distribution, however. The northern hemisphere CLAES sampling pattern shown in Figure 2 for example, is zonally uniform, but oversamples near the 35S and

80N turnaround points. When the number of observations is sufficiently large, $P(\chi)$ can be computed by randomly deleting a latitude-dependent number of observations so that the number remaining in the latitude band ϕ to $\phi + \Delta\phi$ is proportional to $\cos(\phi)\Delta\phi$. $P(\chi)$ is then just proportional to the number of remaining observations with mixing ratios in the range χ and $\chi + \Delta\chi$. PDFs generated by deleting different subsets can then be averaged. Alternatively, each observation χ_j can be weighted with a latitude dependendent weight $w_j < 1$ constructed so that the sum of the weights of the measurements in the latitude band ϕ to $\phi + \Delta\phi$ is proportional to $\cos(\phi)\Delta\phi$. $P(\chi)$ is then computed by adding the properly normalized weights of the observations χ_j for which $\chi \leq \chi_j < \chi + \Delta\chi$. For the results presented here, we have found that the PDFs generated using these two methods are nearly indistinguishable.

In general, the PDF is dependent on both location and time, and the statistics are then said to be "nonhomogeneous" and "nonstationary", respectively. When the space and time dependence is being emphasized, it will be explicitly indicated by the notation, e.g. $P(\chi;\lambda)$ or $P(\chi;t)$. The PDF will always be normalized by the area of the particular geographical region under consideration.

Another kind of PDF is useful when we want to describe the statistics of a "support", that is a region in space over which the tracer field has a specified range of values. In this case we use the notation of conditional probability. For example, $P(\phi\,|\chi)$ is the PDF of the latitudes of measurements having mixing ratios in a small neighborhood of the value χ.

Example A: Tracer PDFs during southern winter and northern summer

This example illustrates the multimodal structure of the tracer PDFs under two dynamically different conditions. Figure 3 is a mapping between the structural features in the PDF and the spatial structure in a χ_j vs ϕ_j scatter plot of the observations. Here χ_j and ϕ_j are the mixing ratio and latitude of the jth N_2O observation, at altitude z = 30 km (15 mb). The scatter plot is on the left and the PDF is on the right, turned on its side in order to make clear that the value of each bin is a horizontal integration through the scatter plot. The single-level PDF presented in Figure 3a is based on a total of about 34,000 individual observations, accumulated over the SH winter period 920817-920916. The double plot shows a correspondence between two points of view; horizontal regions in the scatter plot map into narrow peaks in the PDF, and steeply sloping regions map into minima. Figure 3a shows three distinct peaks associated with the polar vortex, tropics and midlatitude "surf zone", where horizontal mixing is strong during winter. The sharpness of the distribution about each of the peaks is a measure of chemical homogeneity. A narrow peak indicates efficient horizontal mixing of air imported from adjacent regions. Strong minima define the vortex and subtropical boundaries, a clear indication that the three air masses are distinct with respect to the tracer χ. Figure 3b is similiar to Figure 3a, but is based on data from the northern summer period 920719-920810. In the summer hemisphere, there is no vortex peak, and a broad minimum separates the tropical and extratropical air masses. The multimodal structure of PDFs based on GFDL SKYHI model potential vorticity and N_2O fields has also been discussed by *Yang*, [1995].

3. Statistical aspects of transition regions

The following examples illustrate ways in which tracer variability near transition regions in the atmosphere can be characterized. The subtropical boundary will be used as an example, but the ideas can be applied to other transition regions such as the tropopause or the polar vortex edge.

The subtropics is the boundary between the upwelling and downwelling branches of the Brewer-Dobson circulation. As discussed in connection with Figure 1, the mean circulation partitions the atmosphere into tropical and extratropical air masses. To characterize the altitude dependence and seasonality of the stratospheric circulation, and to make progress in the development of simplified models of stratospheric tracer transport [*Plumb*, 1996], observations must be used to define the subtropical boundary more precisely. In addition, differences among models and model/data comparisons are often made clearer when PDFs are used together with contour plots in the analysis of spatial structure [*Douglass et al*, 1999].

Example B: Defining the subtropical boundary

There are different ways in which observations can be used to define air mass boundaries. *Grant et al.*, [1996] for example, estimated the location of the subtropical boundary from the confinement of volcanic aerosols released during tropical volcanic eruptions. Here, we show how the location of the boundary can be estimated from the subtropical minimum in the PDF of a long-lived tracer.

The subtropical minimum in P(χ) occurs at a mixing ratio χ^* that we will call the "tracer boundary" of the subtropics. Returning to the summer hemisphere example of Figure 3b, we see that $\chi^* \sim 140$ ppbv. Mapping across from the PDF to the scatter plot along the horizontal lines and then down to the latitude axis locates the support of the tracer boundary in the range 30N-36N. The term "support" here refers to the latitude range over which the tracer field takes on values in a neighborhood of χ^*. More generally, a support is a region in space over which a dynamical or chemical field takes a certain value, or where a certain condition is satisfied. For the winter hemisphere in Figure 3a, χ^* is also near 140 ppbv, but the support is spread over a wider range of latitudes, and air having chemical characteristics of the subtropics can be found well into the midlatitudes. Midlatitude tracer values near χ^* could either occur from a recent influx of subtropical air or from air originating deeper into the tropics that has partially mixed with midlatitude air. In order to identify the location of the boundary, we consider the PDF of the support of the tracer boundary χ^*, defined as the conditional PDF P($\phi|\chi*$) of the latitudes of observations that have mixing ratios near χ^*. P($\phi|\chi*$) is calculated separately in the northern and southern hemispheres; it is a coincidence that in this example χ^* has the same value in both hemispheres. Figure 4 shows that P($\phi|\chi*$) has peaks at 33N and 16S, with a long tail into the southern midlatitudes. We identify the most probable values 33N and 16S as the northern and southern geographical subtropical boundaries. The width of the distribution near the peak is a measure of the size of the transition region. Note that it is the most probable value of P($\phi|\chi*$) that identifies the boundary, not the mean. The mean is poleward of the most probable latitude in the winter hemisphere because eddy transport brings air with mixing ratios near χ^* to higher latitudes.

Remark: The PDF differs from a map in that it only gives information about how much area is occupied by a certain range of field values, but does not tell us how this area is arranged in space. This is an advantage for problems in which morphological details of the field are irrelevant, or when field contours are rapidly fluctuating and spatial details are not predictable. The previous example shows how statistical information about the spatial characteristics of the field can be recovered by looking at the statistics of the support.

Example C: Measures of chemical differentiation

When can it be said that two air masses are chemically distinct?

Figure 5 shows the tropical (t) and midlatitude (m) PDFs separately for the northern winter period 921225-930105 at altitude z=32 km(10 mb). In this example, "tropical" is the region from the equator to the subtropical boundary at 17N, and "midlatitude" is the region between the subtropical boundary and 55N. The extent to which two air masses are chemically distinct with respect to the tracer χ can be defined in terms of the extent to which their tracer PDFs are resolved. One measure of the degree to which two overlapping distributions are resolved is the separation of the peaks, relative to their standard deviations, and we define a measure of chemical distinctness δ by

$$\delta \equiv |\tilde{\chi}_t - \tilde{\chi}_m|/(\sigma_t + \sigma_m), \tag{2}$$

where $\tilde{\chi}_t, \tilde{\chi}_m$ and σ_t, σ_m are the mode and standard deviations, respectively, of the tropical and midlatitude mixing ratio PDFs. Figure 5 illustrates the calculation of δ. The "inner" standard deviations about the mode are indicated in the figure and are computed only from observations in the range $\tilde{\chi}_m \leq \chi \leq \tilde{\chi}_t$ so that δ will be sensitive mainly to the variability in the subtropics. When $\delta > 1$, the air masses are chemically distinct with respect to the tracer χ. In the example shown here, δ=1.3.

Another measure of chemical similiarity is the area of overlap of the PDFs, indicated by the shaded region in Figure 5. When A=0, the air masses are chemically disjoint, and when A=1, the two PDFs are coincident and the air masses cannot be distinguished chemically. Because the two air masses considered here are contiguous, the overlap area can also be thought of as a measure of the area of the transition region between them. For the example in Figure 5, A=.18, thus by this measure the transition region occupies about 18% of the total area from the equator to 55N.

To give a feel for the magnitudes of A and δ, consider two overlapping Gaussians, both with standard deviation σ=1 and with peaks at $\tilde{\chi}_1$ and $\tilde{\chi}_2$. When $|\tilde{\chi}_1 - \tilde{\chi}_2|$=1, A=0.34 and δ=1, and when $|\tilde{\chi}_1 - \tilde{\chi}_2|$=1.5, A=0.15 and δ=1.5.

The dependence of δ and A on altitude is shown in Figure 6, based on measurements accumulated over the north-facing period 921222-930105. Figure 6 shows that during this time the tropics and midlatitudes are least chemically distinct in the range 30-35 km, something that we can also see qualitatively from the latitudinal gradients in the zonal mean tracer contours in Figure 1. The overlap area A varies in the opposite sense as we might expect, and has a stronger dependence on altitude.

Remark: The overlap of the PDFs in Figure 5 occurs to a large extent because the tropical and midlatitude regions are contiguous, and have similiar values of mixing ratio at their common boundary. We have looked at the PDFs in contiguous air masses in order to characterize the area of the transition region, but in other applications, one might want to consider other choices. The PDF well inside the midlatitudes, from 35-50N for example, might have little overlap with the PDF in the deep tropics, except during times when tracer flux from the tropics to the midlatitudes is strong.

Remark: A dynamical interpretation of the measures A and δ , in the spirit of *Mahlman* [1986] and *Holton* [1986], would view the numerator in Eq. 2 as a "slope steepening" term because the mean circulation increases the separation of the peaks and therefore the subtropical gradient, and the denominator as a "slope flattening" term because eddy mixing across the subtropics increases the variance and decreases the subtropical gradient.

Example D: Zonal mean vs. zonal mode

Contour plots of tracer fields from both models and measurements frequently show tongues of tropical air being drawn into the midlatitudes during winter periods when the circumpolar flow is perturbed [*Leovy et al.*, 1985; *Randel et al.*, 1993; *Waugh et al.*, 1993]. This is the origin of the high tracer values at high latitudes in Figure 7a, which is a χ vs. ϕ scatter plot of tracer measurements at altitude z=32 km(10 mb). The measurements were accumulated over the northern winter period 921222-930105. During this time, a field map (not shown) indicates that the imported tropical air has become wrapped around the polar vortex. Overplotted are the zonal mean $\overline{\chi}(\phi)$ and the zonal mode, $\tilde{\chi}(\phi)$, defined as the mean and most probable values, respectively, of the mixing ratio in a 2° latitude band centered on ϕ. In other words, $\tilde{\chi}(\phi)$ is the peak value of the PDF $P(\chi;\phi)$ and $\overline{\chi}(\phi)$ is its first moment. The zonal mean and zonal mode coincide over most of the summer hemisphere but are very different in the northern winter hemisphere, particularly in transition regions between air masses where $\tilde{\chi}(\phi)$ changes more abruptly. The zonal mode is also less sensitive to eddy transport. The influx of tropical values during this time increases the zonal mean relative to the zonal mode, especially in the winter surf zone, region (c), where $\tilde{\chi}(\phi)$ is nearly constant. It should be noted that a few of the midlatitude measurements here have mixing ratios that are well above the tropical values. There seem to be problems with these particular measurements and they have not been included in the calculation of the zonal mean; they have no effect on the zonal mode. Further discussion of the statistics in the different regions a,b,c can be found later in Example F.

Remark: Modal vertical profiles in individual latitude bands are the tracer values that are observed most often at each altitude. The PDF is often bimodal in transition regions and the mean mixing ratio could actually be close to the minimum in the PDF and therefore have a relatively low likelihood of actually being observed. Modal vertical profiles can also be computed from the altitude dependence of the peaks in the global

PDF, without worrying about which latitude range to average over.

Example E: Defining air mass boundaries from the gradient in the zonal mode

Thus far we have seen two ways of defining boundaries between air masses using PDFs: chemically, by the minimum in $P(\chi)$ at the tracer boundary χ^*, and geographically as the most probable latitude of tracer values in a neighborhood of χ^*. Geographical boundaries can also be defined as the latitude where the most probable tracer value is changing most rapidly, that is by the extrema in the gradient $d\tilde{\chi}(\phi)/d\phi$.

The latitudinal gradient of the zonal mode is plotted in Figure 7b, along with the gradient in the zonal mean for comparison. The subtropical minimum in $d\overline{\chi}(\phi)/d\phi$ is very broad, and a clear boundary is not easily identified. In contrast, there are sharp peaks in $d\tilde{\chi}(\phi)/d\phi$ at the subtropical boundaries near 15S and 15N. The other peak occurs near 60N and marks a transition between the surf zone and polar vortex air masses. However, we note that identifying a fixed latitude as the physical vortex boundary is not meaningful when the vortex is distorted or far off the pole, and should be regarded as a boundary in a statistical, rather than geometrical sense, similiar to the equivalent latitude.

Also plotted for comparison is the standard deviation about the zonal mean $\sigma(\phi) = [\overline{(\chi - \overline{\chi(\phi)})^2}]^{1/2}$, averages again taken over all data within a 2° latitude band centered on ϕ. The zonal variability is largest in transition regions; near the northern subtropical edge for example $\sigma \sim 40$ ppbv, or about 40%. A similiar idea was used by *Remsberg and Bhatt*, [1996], who considered the zonal variance of nitric acid as a diagnostic for mixing across the subtropics.

Two methods have been described for identifying the geographical location of the subtropical boundaries: from the latitude support of tracer values near χ^* and from the gradient in the zonal mode. We have used both measures to compute the subtropical boundaries over a range of altitudes for the data shown in Figure 7. Figure 8 shows that the two methods give similiar results, except near z$\sim$27 km in the southern hemisphere. If we may assume that the subtropical boundaries as defined here from tracer gradients are close to the boundaries between the upwelling and downwelling branches of the circulation, then Figure 8 is a picture of the tropical pipe [*Plumb*, 1996]. An interesting feature is the narrowing and northward jog of the pipe near z=32 km, where δ is a minimum and A is a maximum (see Figure 6).

4. Contributions to tracer variability from transport on different timescales

The large scale latitudinal structure of the tracer field is shaped by the time-integrated effects of mixing and transport. In this section we are interested in the shorter timescale variability that is superimposed on the large scale structure. The statistics of fluctuations from the most probable tracer value at each latitude is a way to separate the effects of mean and eddy transport. As we show below, the fluctuations from the zonal mode show the signatures of recent eddy transport as well as a more "aged" background variability.

Example F: Fluctuations about the zonal mode

Given an observation χ_j, at latitude ϕ_j, we define a fluctuation from the zonal mode by $\epsilon_j \equiv \chi_j - \tilde{\chi}(\phi = \phi_j)$. This removes the large scale latitudinal variation of the mixing ratio that is shaped by the longer timescale transport. The PDF P(ϵ) characterizes the statistics of the fuzzy envelope of displacements of the observations from the zonal mode curve, as in Figure 7a. We have defined the fluctuations relative to the mode rather than the mean because, as Figure 7a shows, the zonal mean includes the effects of recent eddy transport and is not a proper reference state for our purposes here. The zonal mode generally varies more slowly than the mean because it is by definition the tracer value that occupies the most area. (Exceptions do occur, for example during stratospheric warmings as shown in a later example.)

Figure 9 shows P(ϵ) for the three regions indicated in Figure 7a. The PDFs sometimes have long tails, but the typical variability, that is, the small fluctuations around the most probable value (the "core") is Gaussian with a standard deviation that is indicated on each of the plots. In the southern low latitudes, region (a), P(ϵ) is Gaussian over the entire range of ϵ to a good approximation. In fact, the PDFs in individual latitude bands (not shown) throughout the tropics and southern hemisphere are all nearly Gaussian with a variance that depends weakly on latitude.

In the northern (winter) hemisphere on the other hand, the fluctuations from the zonal mode are not Gaussian and are more strongly latitude dependent. In region (b), P(ϵ) has long tails on both sides and a Gaussian core is not so clear, relative to the other two cases. Thus, there is no clear background in the NH subtropics, in contrast to the SH subtropics. P(ϵ) is not symmetric about ϵ=0; positive fluctuations are stronger and are observed somewhat more frequently. This is consistent with an asymmetry in the transport across the subtropics, in which outflow from the tropics is more likely than inflow into the tropics.

We have defined the midlatitude surf zone, region (c), as the latitude range over which $\tilde{\chi}(\phi)$ is nearly constant, and here P(ϵ) is a highly asymmetric one-sided Gaussian. Negative fluctuations are Gaussian, but the distribution has a long positive tail, the signature of high mixing ratios imported from lower latitudes. While the zonal mode is nearly constant across the midlatitudes, the fluctuations about the zonal mode are latitude dependent. For example, P(ϵ) in a narrow latitude band centered at 50N (see Figure 7a) has a tail that is due solely to subtropical air wrapped around the vortex. On the other hand, in the interior of the surf zone near 40N, well away from transition regions, P(ϵ) is nearly Gaussian.

Gaussian variability can arise for many different reasons and the Gaussianity by itself, while evidence that the Central Limit Theorem is somehow at work, offers little other insight about its origin. A Gaussian core in P(ϵ) could be due to the fact that the satellite measurements are spatial integrations over smaller scale inhomogeneities. It could also arise from random errors in the measurements, due to both the instrument and the satellite retrieval. Another source of Gaussian variability is atmospheric mixing, a consequence of the Central Limit Theorem applied to the collective effects of mixing over long times. If the source of the variability could somehow be turned off, the Gaussian would continue to narrow, approaching a delta function in the long-time limit of complete chemical homogeneity (and perfect measurements).

Ultimately, the reason for the asymmetry in P(ϵ) in the surf zone (Figure 9c) is that low N_2O air enters the region mainly by vertical descent from the upper level surf zone, while high N_2O air enters sideways from the tropics. The isentropic transport from the tropics to the surf zone is episodic and rapidly brings in air that is much more chemically distinct than air that is continually descending from above. These different pathways also have different mixing histories. The descending air is mixing vertically by diabatic dispersion, the vertical dispersion that arises from variability in diabatic heating along parcel paths. This process is most important in the winter surf zone where the gradient in diabatic heating is large [*Sparling et al.*, 1997], and we might expect it, together with the vigorous horizontal mixing in the winter surf zone, to give rise to a typical or background variability that is Gaussian. We see only half of the Gaussian in 9(c), because the right positive half is masked by the strong tail. The tail represents "anomalous" variability: air recently imported from the tropics that has not fully mixed down to become part of the background. The mixdown can be slow at times, particularly when the imported tropical air becomes wrapped around the polar vortex or trapped within the Aleutian anticyclone.

The lined area between P(ϵ) and the Gaussian shown in Figure 9c is an estimate of the fraction of the surf zone, averaged over the time period, containing air that has some subtropical character. In this example, the "atypical" air occupies 26% of the surf zone. A portion of this is just poleward of the subtropical boundary and another is at higher latitudes, wrapped around the vortex.

Remark: To summarize, the effects of transport and mixing on long, intermediate and short timescales can be seen in the tracer statistics. The large scale structure described by the zonal mode evolves on the longest timescale. Superimposed on this is a Gaussian core of typical variability that arises from mixing and transport over intermediate time scales. Superimposed on the Gaussian background are non-Gaussian tails from recent eddy transport that have not yet mixed in to become part of the background. (This separation of "mean" and "eddy" timescales is not always so clear however. This is demonstrated later in example G, a period during which global scale changes in the tracer field were occuring very rapidly.) The general ideas here can be used in any problem where one needs to distinguish between typical and anomalous variability or where a background context for a perturbation must be defined. [see, for example *Thompson et al.* 1999]. It is important to emphasize, however, that the tails do not *by themselves* say anything about irreversible transport, since they can also arise from wavelike disturbances that move air back and forth across fixed geographical boundaries. In the above example however, the imported air has become wrapped around the vortex, and a return to the subtropical reservoir is unlikely.

Measurement precision

In principle, the width of the Gaussian core should give information about the mixing rate, but in practice this is possible only when the measurement precision is smaller than the width of the Gaussian core. *Roche et al.* [1996] have estimated the random errors in the CLAES v.7 N_2O data to be 5.7 ppbv at the 10 mb level, during the summer. This is significantly smaller than the standard deviation of 17 ppbv in the summer hemisphere shown in Figure 9a. In this case the Gaussian most likely arises from ongoing mixing across the southern latitudes.

Since the weakest dynamical variability is in the high latitude summer southern hemisphere, we expect that the tracer distribution there would be a good test of the instrument precision. Figure 9d shows $P(\epsilon)$ during the SH summer period 930113-930127 at 10 mb and at high latitudes 50S-80S. The observed variability is indeed weak, and is Gaussian with a width that is much smaller than what was found in the previous cases. The standard deviation of 5.2 ppbv is in fact somewhat smaller than the reported uncertainty of 5.7 ppbv at this level, and is likely at the limit of the instrument precision. Estimates of the overall measurement precision, which includes both instrumental and retrieval uncertainties, are more difficult to make in the highly variable winter hemisphere [S. Massie, personal communication].

5. Dynamics: time evolution of the PDF

The examples in this section illustrate the time evolution of the tracer PDF under non-conservative dynamical processes acting on both seasonal and weekly timescales.

Example G: An early winter warming

A warming during the first 3 weeks of December, 1992 resulted in an episode of strong diabatic cooling and increased descent which rapidly brought down low mixing ratios into the midlatitudes. This example is mainly concerned with the evolution of the PDF during this time. A brief discussion of the tracer PDFs during fall and late winter is also included to provide some context for the highly nonstationary conditions during December.

The data used here is from the October, December and February UARS north-facing periods; there is no middle or high latitude CLAES data during November and most of January. Since we are interested in diabatic effects, we study the evolution of the PDFs on the 950K surface, but it should be noted that the time evolution of the PDF is not that different on the 6.8 mb level. The CLAES v.7 temperatures were used to interpolate mixing ratios to the 950K surface; θ=950K is the mean potential temperature over the northern 6.8 mb surface during this time.

Figure 10 shows how the NH tracer distribution evolved over the course of the 92/93 winter. Each curve is the PDF of 5 consecutive days of data, centered on the date indicated in the figure. For clarity, the PDFs are displayed as smooth curves rather than individual bins. The peaks at low, intermediate and high mixing ratios correspond to the vortex, midlatitude and tropical air masses, respectively.

During October (top panel) the polar vortex spins up and the early winter strong polar descent [Rosenfield et al, 1994] brings down very low mixing ratios. This, together with the isolation of the vortex, gives rise to the strong peak near 10 ppbv. The smaller peak near 40 ppbv corresponds to the emerging surf zone. Changes in the PDF are small and fairly gradual throughout October; the early and late October distributions are similiar.

Maps of the PV (potential vorticity) field in late November indicate that the vortex moved off the pole and the perturbed flow pulled tropical air into the midlatitudes. The tracer distribution on 921204 (middle panel) shows a fourth peak at subtropical values near 90 ppbv which could be the result of mixing between the tropics and midlatitudes. The PDF changes drastically during the following two weeks. The fraction of midlatitude air with mixing ratio below 50 ppbv increases throughout December; the lower mixing ratios could not have come from the relatively small pool of vortex air, but are instead due to vertical advection of low mixing ratios from above. The surf zone peak also sharpens, an indication that the region is becoming more homogeneous, presumably as a result of an increase in horizontal mixing from Rossby wave breaking. By early January (lower panel), the surf zone is well-established as a distinct, well-mixed air mass. This rapid evolution in the tracer distribution is followed by a period of more gradual change, and the PDF looks much the same when UARS again faces north in mid-February (bottom panel). The tropical peak has become broader and has extended to higher mixing ratios, but the midlatitude peak and the subtropical minimum have changed little. In the surf zone and subtropics then, the time-integrated effects of horizontal mixing, influx from the tropics and diabatic descent are either balanced, or are small during this time.

Another way to study the development of the surf zone is to look at the way in which the zonal mode curve $\bar{\chi}(\phi)$ changes over the winter. Figure 11 shows the zonal mode for the same time periods considered in Figure 10, except that each curve is computed from measurements accumulated over two weeks. The flat midlatitude portion of the curve is the surf zone. It extends as winter progresses, and begins to shrink by mid-February. The zonal mode curve tilts upward at high latitudes during times when large amounts of tropical air are transported poleward.

Figure 12 shows how the rapid change in the tracer distribution during December is correlated with the rising temperatures and increased diabatic cooling. Here we have used NCEP (National Center for Environmental Prediction) temperatures as described in *Finger et al.*, [1993] and references therein. Diabatic cooling rates were computed from the NCEP temperatures (courtesy Joan Rosenfield, *Rosenfield, et al.* [1994]). The three curves in the figure show the time evolution of 1) A_χ, the fraction of the midlatitude (35–55N) area occupied by mixing ratios less than 50 ppbv, 2) A_Q, the fractional area for which the diabatic cooling rate is less than -5K/day, and 3) A_T, the fractional area occupied by temperatures above 230°K. (Heating rates are in degrees potential temperature per day.) The cutoff values for heating rate and temperature were chosen on the basis of the shape of the heating rate and temperature PDFs (not shown), which developed large tails during this time. NCEP temperatures have been used instead of the CLAES temperature retrievals to independently verify the relation between the diabatic descent and decreasing mixing ratios. The figure shows that the area occupied by low tracer mixing ratios increases at roughly the same rate as the area of warm temperatures and strong cooling. The correlation between the changes in the tracer and temperature fields could also have been demonstrated by simply considering the time evolution of the midlatitude-averaged quantities, but this would not give any information about the scale of the event. The area calculation shows for example that the area occupied by mixing ratios below 50 ppbv grows from 50% to 80% of the midlatitude area.

Remark: One diagnostic for model transport in CTMs is the extent to which the model transport reproduces the observed strong subtropical tracer gradients [*Douglass et al.*, 1999]. Example G shows that the relatively brief December warming event was important in creating a subtropical gradient that changed little over the remainder of the winter. The misrepresentation of such events in models can therefore lead to weak gradients in modeled tracer fields.

Measures of variability: the entropy

For some problems, it is convenient to have a simple measure of the tracer variability that is derived from the full PDF. For example, we might want to summarize the evolution of the PDF from day to day by a single function of time. The area occupied by a mixing ratio above or below some cutoff, as in Figure 12, is one example and is useful when changes in the tails of the distribution is of interest. A more common choice is to consider the first or second moments (the mean and variance) of the distribution. In general, a unimodal short-tailed distribution is well-characterized by the mean and variance, and they completely specify a Gaussian distribution. Higher moments become important when the PDF has long tails or is not symmetric about the most probable value, but they are sensitive to undersampling. Another problem with moments is that they describe the distribution of probability around some reference point such as the mean or mode, but the choice of this reference point is not so clear when the PDF is multimodal.

When looking at transport, we are often interested mainly in the structure of the PDF, and not so much on the value of the quantity being binned. One way to summarize the structure in the PDF is to consider the entropy of the distribution. The tracer PDFs are computed by binning the observations into M bins of width $\Delta\chi$. If p_i is the fraction of observations in each bin, the Shannon entropy [*Shannon and Weaver*, 1949] is

$$S = -\sum_{i=1}^{M} p_i \log p_i \tag{3}$$

which is a measure of the information content of the field. This is not a thermodynamic quantity, although it is interesting that the thermodynamic entropy can be expressed in the same way.

The entropy is a maximum when the M tracer bins are equally probable so that $p_i = 1/M$. In this case the tracer field has maximum variability and the entropy is equal to the maximum entropy S_{max}=log M. By analogy with thermodynamics, we can also define the effective "number of states" Ω=e^S, with Ω_{max}=M when the variability is a maximum. The number of possible states is the maximum range of the variability divided by the bin size $\Delta\chi$, where $\Delta\chi$ is on the order of the absolute precision of the measurements. The measurement uncertainty, then, reduces the information content of an ensemble of measurements. The minimum entropy state is the state of chemical homogeneity, to within the bin size $\Delta\chi$. If the region is perfectly homogeneous, then p_i=1 for one bin and is zero for the rest, and S=log 1 =0. It may seem odd that a homogeneous air mass has minimum entropy, since we are used to thinking of a well-mixed state as an equilibrium state, a state of maximum entropy. The difference is that we are talking about the entropy in "chemical space". If we partition physical space into cells, homogeneity occurs when each cell has the same amount of tracer. The configurational entropy in physical space is then a maximum [*Sparling and Schoeberl*, 1995], because a field without structure conveys no information.

The time evolution of Ω summarizes the evolving structure in the PDF. Figure 13 plots Ω versus time during the warming of the previous example. In this example, the bin size $\Delta\chi = 5$ ppbv, and the mixing ratios were binned into 50 bins, so S_{max}=log 50. Overall, the entropy decreases with time, the diabatic circulation and horizontal mixing acting together to produce a hemisphere that is more highly "ordered" into distinct air masses. The decrease is not monotone, but is punctuated by episodes of increased entropy that are due in part to mixing between the air masses.

Example H: Polar vortex breakup and trapping in the Aleutian high

The decay of the polar vortex at the end of winter typically begins first at high altitudes and progresses downwards [*Randel*, 1992]. As the polar jet wind speed decreases, the transport barrier separating the vortex from the surrounding surf zone weakens and the air masses begin to mix together. Interesting details about the mixing of polar and midlatitude air masses after the final warming can be found in *Hess* [1991].

The PDFs show this mixing process as a filling in of the vortex edge minimum. This can be seen in Figure 14, which compares the high latitude tracer PDFs for the second half of February and the first half of March 1993 at a range of vertical levels. The polar region is of interest here, so the PDFs are based only on observations in the high latitude band 50-80N. The polar cap 80-90N is excluded, so for example the area under the vortex peak in the PDF is somewhat smaller than the actual total area of the polar vortex.

A comparison of the late February (shaded) and early March(line) PDFs shows little change in the chemical separation of the vortex and surf zone below the 10 mb level, because the vortex at the lower levels is still strong and dynamically isolated. At higher levels on the other hand, the March vortex and surf zone peaks have merged to an extent that increases with altitude. At 38 km and above, the polar jet is weakening and the minimum between the vortex and midlatitude peaks is filling in, an indication that the air masses are mixing together. These changes are not due to chemistry, because they are occuring on a time scale that is much faster than the photochemical lifetime of N_2O at these high latitudes [*Ko et al.*, 1991]. The evolution from a bimodal to a unimodal tracer distribution means that there has been an increase in the amount of air having chemical properties intermediate between those of the two air masses. If advection and chemistry can be ruled out, this can only happen from mixing between the air masses.

The high N_2O tails in the PDFs are due to the influx of subtropical air into the polar region which had its origins in a mid-February wave 1 perturbation that moved the vortex off the pole. Low PV air was drawn up to high latitudes, the Aleutian anticyclone intensified and high N_2O air became trapped within it. Poleward transport of tongues of ozone under similiar conditions was first seen in LIMS observations by *Leovy, et al.* [1985]. Trapping of tropical aerosol in the Aleutian High has also been described by *Harvey et al.* [1999]. The trapping appears most pronounced near the 10 mb level where the high N_2O values appear as a third distinct air mass, and we see a clear minimum in the PDF marking a tracer boundary between the midlatitude and the Aleutian air masses. High values of χ do not appear as a third distinct air mass at other levels during this time, either because the imported tropical air has now mixed with the surrounding air, or because the tropical air was not strongly trapped in the first place. The minima in the PDF at the 10 mb level, at 26 and 90 ppbv, are the tracer boundaries between the three air masses, and we define the polar vortex by $\chi \leq 26$ ppbv, the surf zone by $26 < \chi \leq 90$ ppbv and the anticyclone by $90 < \chi \leq 150$ ppbv. The support of each of these ranges of mixing ratio during the first half of March can be seen in Figure 15a, which clearly shows the compact support of the high N_2O values poleward of 50N, and their apparent subtropical origin. Trajectories computed with the UKMO winds show that most of the trapping actually occured in mid-February when the anticyclone was intensifying over 180W. After this time, the anticyclone moved eastward, transporting its cache of high N_2O to 90W.

The spatial structure in the high latitude N_2O field is consistent with the PV field during this time. The bottom plot is a "reverse domain fill" (RDF) plot [*Sutton et al.*, 1994] showing the high resolution PV field on 930310. Parcels were initialized on 930228 with the value of the PV field at their initial locations and advected for 10 days. The PV and wind fields are from the UKMO assimilation system [*Swinbank and O'Neill*, 1994]. The point of this figure is to show the flow structure; the actual PV values are not important for our purposes here. The anticyclone over Canada contains low subtropical PV, the dark comma-shaped region is the distorted polar vortex containing high PV, and the region between them has intermediate PV values.

Remark: If the observing period had been longer, it would have been possible to track the time course of the mixing of polar and tropical air at the 10 mb level during the vortex breakup by following the time evolution of some measure of the mixing, for example the entropy. Unfortunately, the CLAES monitoring of this "two-ring circus" in the north polar region comes to a halt in mid-March when UARS turns south to face a quiescent summer hemisphere.

6. Statistical robustness

Because atmospheric fields have spatial and temporal structure over a wide range of scales, the question arises as to whether a sparse or limited spatial or temporal sampling of the field is representative of the full range of variability. A satellite platform can provide data with nearly synoptic, global coverage, and by subsampling the data it is possible to assess the statistical robustness of atmospheric measurements taken on other platforms or with instruments that have more restricted or biased sampling.

Example I: Robustness under subsampling

A collection of data is called "statistically robust" if the statistics have converged, that is, if the PDF does not change when more data is added, and is not sensitive to an unbiased reduction of data. The shaded PDF in Figure 16 is the PDF of all observations from the equator to 50N, accumulated for one month starting on 920925. There are about 700 measurements per day in this latitude range. The observations were subsampled by randomly selecting a subset of 350 (12/day) observations out of the total of 21,000, thus the subsample contains less than 2% of the data. The PDF of this subset is the "subsampled" PDF, also plotted in Figure 16. The overall shape of the PDF, including the location of the extrema, survives this rather severe reduction of data surprisingly well. One reason for this is that the meteorological conditions are not changing too rapidly during this time (see Fig. 10a). Another reason is that while the subsampling is sparse, it is unbiased, i.e. more or less uniform in space and time.

Another example of sparse subsampling, but one that is not uniform in area, is the sampling by the solar occultation instrument HALOE (HALogen Occultation Experiment) on UARS [*Russell et al.*, 1993]. Within each 24 hour period, 15 sunrise/sunset observations are made along each of two latitude circles; the latitudes change each day and sweep the hemisphere in about one month. During the time period studied here, there are about 350 observations from the equator to 50N, with no observations in the northern hemisphere during the middle third of the time period.

The dotted curve in Figure 16 is the "HALOE-sampled" PDF. This is the PDF of a subset of 350 CLAES observations that are closest to the times and locations of the 350 HALOE measurement locations during this one month period. The comparison shows that the extrema of the PDF survive the HALOE subsampling, however the total area under each peak does not.

The statistical robustness of a temporally or geographically restricted set of observations depends on the scale of variability relative to the size of the spatial or temporal sampling window. For example, global-scale changes in the tracer field were taking place rapidly during the warming event in Example G. The PDF under such highly non-stationary conditions would not likely be as robust as the example considered here, and would be especially susceptible to sparse HALOE-type sampling. Problems related to synoptic variability and insufficient sampling have also been discussed by *Bithell et al.*, [1994].

Remark: A variety of methods are used to reconstruct fields from a sparse set of observations. These include trajectory mapping, [*Pierce et al.*, 1994; *Morris et al.*, 1995], Fourier mapping [*Salby*, 1982], and PV reconstruction [*Schoeberl and Lait*, 1991]. One way to quantitatively test these different methods would be to compare the PDFs of an original test field with a field that has been reconstructed from an unbiased subsampling of the original field. A conservative algorithm will fill in missing information in the tracer field without changing the total amount of tracer in each mixing ratio bin. In other words, the PDFs of the subsampled points and the reconstructed fields should be the same.

Example J: Homogeneity and Stationarity

The extent to which a temporally or geographically restricted sampling of a tracer field is representative of the full range of variability depends on the homogeneity and stationarity of the tracer field. This depends in turn on the homogeneity and stationarity of the underlying dynamics. The time dependence of a global or hemispheric tracer PDF on a single vertical level is due to vertical advection, chemistry and mixing. The tracer statistics are nearly stationary during times when these different sources of variability are either slow or balanced. Figure 17 shows the 10mb PDFs in the southern hemisphere during 930113-930208. The curves in Figure 17a are the PDFs $P(\chi;t)$ for successive 4 day intervals during that time, i.e. days 1-4, 5-8, etc. There is some variability in $P(\chi;t)$ from each 4 day period to the next, but the overall structure of the PDF varies little during this time. This is in contrast to the rapid time evolution during the NH early winter warming (Figure 10b).

While we do not expect the tracer statistics to be homogeneous in latitude, they are often homogeneous in longitude, particularly in the SH. Figure 17b shows $P(\chi;\lambda)$, the tracer PDFs in disjoint 30° longitude slices. Each curve includes all the measurements in this time period. The zonal symmetry of the field is reflected in the weak dependence of $P(\chi;\lambda)$ on longitude. The tracer distribution in each longitude slice is thus representative of the global statistics.

This is not the case in the NH during October 1992. Large scale zonal inhomogeneity in the tracer field is illustrated by the 10mb PDFs shown in Figure 18. The different curves are again the PDFs in 30° longitude slices. The peak at high mixing ratios is larger than in the previous figure because the region of upwelling is shifted into the NH during this time. In the region 0-30E (grey curve), the distribution has a strong vortex peak near 20 ppbv because the vortex is off the pole over the Greenwich meridian. The black curve (150-180E) shows that the surf zone, the peak near 60 ppbv, is building on the other side of the globe. Sampling that is zonally restricted clearly does not encompass the full range of variability of the field. This example is interesting because while inhomogeneous, the statistics during this time were fairly stationary. Note that the time period here is the same as that considered previously in Figure 16. Because the statistics are stationary, a daily sparse sampling, accumulated over the month, is representative of a more spatially dense set of observations.

Remark: As we have seen previously, another way to look at spatial structure is to consider the statistics of the support. For example, persistent zonal asymmetries in poleward transport can be characterized statistically by $P(\lambda \mid \chi^*)$. The connection with the underlying dynamics could then be made by combining this with statistical studies of the winds, as in [*Pierce and Fairlie*, 1993] and [*Pawson et al.*, 1996]

7. PDFs of gridded meteorological fields

While our emphasis thus far has been on PDFs of chemical tracer data, one can look at PDFs of gridded model or meteorological data in the same way. Gridded data is often represented by a field map such as a contour plot of the field on a single vertical level. The next example shows the correspondence between PDFs and contour plots in order to make clear the connection between the two points of view.

Example K: Contour plots and PDFs

The computation of area-weighted PDFs of gridded data is easily implemented by first interpolating the global gridded field to an equal area distribution of N points. The fractional area between two tracer contours is then computed by simply counting the number of points in the region bounded by the contours and dividing by N.

Figure 19a shows a contour plot of the UKMO 500K potential vorticity analysis on Dec 1, 1995. During this time, the decaying southern polar vortex and the growing northern polar vortex are shifted off the poles toward each other over the Greenwich meridian. The PDF of the same field is shown in Figure 19b. P(q) is the fraction of the global area enclosed by the PV contours q and q+Δq. The bin size Δq=6 pvu and the bin values have been chosen in order to match the contours in the contour plot, thus the value of each bin is the fractional area between the contours in Figure 19a. The PV bins occupying the largest area in each hemisphere are shaded; these bins correspond to the northern and southern surf zones and the value P(q) is the area of the dotted regions in Figure 19a, between the $|q|$ =24 and $|q|$=30 pvu contours. The surf zones together comprise about one-third of the global area.

Minima in P(q) correspond to regions where the gradient in potential vorticity is large and contours are closely spaced, while maxima correspond to more homogeneous regions where the contours are far apart. The polar vortices appear as small maxima at the ends of the distribution, and the minima near q=±40 pvu are the vortex boundaries.

The minima are coincident with the location of the vortex edge as computed from the algorithm of *Nash et al.* [1996], indicated in the figure by the thin vertical lines. There are hints of additional minima near the subtropical values ±20 pvu, but these are not clearly defined with only one day of data, even when a smaller PV bin size is used. The next figure shows how the multimodal structure becomes better defined when the PDFs are based on several days of data.

Example L: Comparing two gridded fields

Figure 20 shows the potential vorticity PDF for the UKMO analysis as in Figure 19b, except here the data has been accumulated over the 4 day period Nov. 27 to Nov 30, 1995 in order to better define the multimodal structure. The PDF for the GSFC (Goddard Space Flight Center) PV analysis is also shown for comparison. Details about the UKMO and GSFC stratospheric analyses can be found in *Coy and Swinbank*, [1997]. P(q) has the same overall structure in both analyses, both have minima at the equator and vortex edge and strong peaks in the homogeneous midlatitudes. In the UKMO distribution, minima also appear at the edge of the tropics near ±10°, while the GSFC analysis has only a ledge there.

The polar vortex edge appears as a rather broad minimum in both analyses. The dotted vertical lines indicate the location of the UKMO vortex edge, computed from the zonal mean latitudinal PV gradient [*Nash et al*, 1996]. The range of uncertainty, defined from the inflection points of the zonal mean PV, is indicated by dashed lines. The vortex edge defined in this way agrees with the location of the minimum in P(q).

The fractional area of the polar vortex is the area under the vortex peak, integrated from the dotted line to the end of the distribution, and is about 3 % of the global area for each vortex. The PDF shows that the area of the edge region, that is, the area between the dashed lines, is comparable to the area of the vortex itself. This example shows how two fields can be quantitatively compared through the entire range of field values. We can see very clearly the quantitative agreement near the poles, the more homogeneous UKMO midlatitudes, and stronger subtropical and equatorial PV gradients in the UKMO analysis. A comparison with Figure 19b shows how the UKMO subtropical minima in P(q) near ±20 pvu become better defined by accumulating data over several days. This is in contrast to a time averaged map where spatial features can become blurred under time averaging.

Remark: For global PDFs of fields that are nearly conserved over a week or so, day-to-day differences in the PDF are small and combining several days of data refines the structure in the PDF without introducing new features. New features can appear on the other hand, when the PDF changes rapidly with time. For example, suppose that we had global measurements of a certain field that were clustered around some value χ_1 on day 1 and clustered around χ_2 on day 2. If χ_1 and χ_2 are widely separated, then combining these two days produces a bimodal PDF, i.e. a prominent new feature has been introduced that is not characteristic of the field on any one day. If χ_1 and χ_2 are closer together, the combined PDF will remain unimodal but will have a larger variance. In either case, the PDF loses its interpretation as a measure defined on the spatial structure of the field at a particular time. Of course, the PDF retains its more general meaning as the likelihood that the field has a given tracer value sometime during that two day time interval.

8. Overview

Much of what has been learned about stratospheric transport over the past couple of decades has been inferred from the morphology of chemical fields that have been constructed from observations of long-lived tracers. The statistical view of large scale tracer variability that has been presented here is a complementary view in which a gridded field or a global set of observations is thought of not only as a field map, but also as a probability distribution. The two points of view go hand in hand; the PDF summarizes and quantifies the field variability, but its proper interpretation requires a global view of the spatial structure in the field.

While it certainly cannot be said that binning data is an innovative idea, sometimes even a modest shift in point of view can lead to new insights. The examples were chosen to illustrate different aspects of the transport that are not obvious or quantifiable by looking only at the spatial structure of the field. Examples A-E and L illustrate how the structure in the tracer PDFs can be used to characterize different transport regimes and the boundaries between them. One obvious advantage of PDFs, pointed out in Examples K and L, occurs when data is accumulated over a timescale on which the field is conserved. This blurs spatial structure in a map, which changes under conservative rearrangements of the field. But it sharpens and refines structure in the PDF which is nearly invariant under the horizontal advection. This allows quantitative studies of the evolution of the stratospheric circulation. For example, the seasonal migration of transition regions can be calculated from tracer measurements and directly related to the evolution of similiarly defined boundaries in dynamical fields [J. L. Neu et al, Variability of the tropical pipe boundaries, *in preparation*]. Changes in the tracer field due to non-conservative transport processes, illustrated by Examples G and H, can be investigated by looking at the time evolution of the PDF or some measure of its structure, such as the entropy. The vortex breakdown example shows how mixing between two distinct air masses takes place by the filling in of the minimum in the tracer PDF. The increase in entropy associated with such mixing is a physically appealing way to look at irreversibility. Example F illustrates how the statistics of fluctuations about the zonal mode can be used to separate the effects of transport on different timescales, and also suggests a way of defining typical versus anomalous variability.

Examples I and J deal with the properties of homogeneity and stationarity, and the implications for statistical robustness. A dataset that has full temporal and spatial coverage, such as the satellite measurements used here, provides a means for investigating the homogeneity and stationarity of the dynamical processes that underlie the homogeneity and stationarity of the tracer field. These properties also have a direct bearing on the important issue of statistical robustness. Whether a set of measurements represents the full range of variability of a geophysical field depends on the temporal and spatial variability of the field. Issues related to statistical robustness arise in a number of practical contexts, for example in the design of data-gathering strategies, when it is necessary to know the number and spatial distribution of observations that will be sufficient to answer the science question at hand. These considerations are also important in the analysis of data from aircraft field missions or satellite validation campaigns where the sampling of the atmosphere is of necessity spatially and temporally restricted.

Tracer distributions are often multimodal or asymmetric, and in these cases the mean of the distribution can be very different from the most probable value. When the PDF has a long tail, moments can be dominated by the rare but intense events in the tails which are often undersampled. The shape of the PDF is particularly important when it is necessary to distinguish between the background or natural variability, and a superimposed transient or anthropogenic perturbation. The average value over a region affected by some perturbation is clearly not the background, because the average includes the perturbation itself. The background distribution is better characterized by the variability in a neighborhood of the most probable value. In addition, the mean tracer value in transition regions where the PDF is bimodal may not be representative of the typical tracer value and in fact can be close to the *least* probable value. This is a well-known problem in the northern polar region when the vortex is off the pole, and zonal means are averages over a bimodal surf zone + polar vortex tracer distribution. These examples are reminders that, for non-Gaussian variability, "average" and "typical" do not always mean the same thing. We note that while the zonal mode is useful in a diagnostic framework, the development of a prognostic equation for the mode - an extremum of the PDF - would likely be a difficult mathematical problem.

Inferences about transport that are based on the structure of the PDF must take into account artifacts due to both random and systematic errors in the measurements. Under dynamically active conditions, the range of variability in satellite measurements is an amalgam of real atmospheric variability and a processing uncertainty that is itself dependent on dynamical conditions, for example through the temperatures used in the retrieval. The possibility of obtaining information about mixing from the Gaussian core in a tracer PDF is certainly worth pursuing, but with the understanding that Gaussian variability can also arise from measurement uncertainties.

For simplicity, and for ease of comparisons between models and data, it is convienent to compute the PDFs within fixed geographical regions. This might not be the best choice in applications where irreversible mixing is of interest. Wave transport can lead to reversible fluxes of chemically distinct air back and forth across a fixed boundary, giving rise to tails in the PDF. To remove the effects of rapid and reversible transport, air mass boundaries can be defined with respect to a dynamical coordinate such as PV. For example, the tracer distribution within the northern polar vortex can be defined as the conditional PDF $P(\chi|q \geq q^*)$, where q* is the potential vorticity at the vortex edge, defined by the minimum in P(q). This approach may be less useful in the subtropics however, and given the differences one finds in PV analyses there (Example L), it might be best to avoid introducing a model-dependent uncertainty into the observations.

To learn more about what tracer PDFs have to say about stratospheric mixing and transport requires investigations of the PDF from data collected over a range of dynamical conditions. Studies of the structure of tracer PDFs generated from numerical mixing experiments where the forcing and dissipation can be specified [e.g. *Pierrehumbert*, 1994] are also necessary for the interpretation of measurement PDFs. Further progress also requires a theoretical framework for understanding and interpreting the structure and time evolution of the PDF. Some recent progress has been made for layerwise flows and simple models of advection/diffusion [*N. Nakamura*, pers. comm; *Majda and Kramer*, 1998; *Pierrehumbert*, 1999] which can be applied to mixing processes on isentropic surfaces. The evolution of the tracer distribution on the timescale of the stratospheric circulation is a fundamentally different problem, however, which must take into account vertical advection. It must also include a correct representation of the "sources", i.e. the intermittent injection of tracer from both the tropics and winter polar region into the midlatitudes. In addition, the frequently made simplifying assumption of statistical homogeneity clearly does not apply to stratospheric tracer fields because of the large scale meridional structure and the presence of large scale vortices and jets. It may be necessary to adopt a simplified representation of the large scale transport, such as *Plumb's* [1996] tropical pipe model.

The examples presented here are a small sample of ways in which PDFs can be used to quantify and synthesize the large amount of observational and model data that is now becoming available. To limit the scope of this paper, the discussion was restricted to stratospheric transport inferred from one-point distributions of satellite observations. Other statistical tools, such as conditional or two-point statistics, will also be useful for other kinds of problems that require data from different measurement platforms.

Acknowledgments. I would like to acknowledge the NASA/AEAP program and the

EOS Interdisciplinary Program for support of this work. Thanks to A.Roche, purveyor of the CLAES observations used here, for discussions on instrument uncertainty, and to Anne Douglass, Stacey Hollandsworth, Steven Pawson, and Richard Swinbank for very helpful comments on the manuscript. This work would have proceeded even more slowly without the remarkable GSFC Code 916 computer/data acquisition system, and the talents of David Gent, Pat Guimaraes, Leslie Lait, Eroc Nash and Cid Praderas. Special thanks to Julio Bacmeister, Larry Coy and Mark Schoeberl for criticisms of this work and for many discussions that have helped me to formulate the ideas here.

References

Bithell, M., L. J. Gray, J. E. Harries, J. M. Russell III, A. F. Tuck, Synoptic interpretation of measurements from HALOE, *J. Atmos. Sci.*, 51, 2942–2956, 1994.

Butchart, N. and E. E. Remsberg, The area of the stratospheric polar vortex as a diagnostic for tracer transport on an insentropc surface, *J. Atmos. Sci.*, 43, 1319-1339, 1986.

Coy, L., and R. Swinbank, Characteristic of stratospheric winds and temperatures produced by data assimilation, *J. Geophys. Res., 102*, 25763–25781, 1997.

Dunkerton, T. J., C.-P. F. Hsu and M. E. McIntyre, Some Eulerian and Lagrangian diagnostics for a model stratospheric warming, *J. Atmos. Sci., 38*, 819–843, 1981.

Evaluation of the UARS data, JGR special issue, *J. Geophys. Res., 101*, No. D6, 1996.

Finger, F. G., M. E. Gelman, J. D. Wild, M. L. Chanin, A. Hauchecorne, and A. J. Miller, Evaluation of NMC upper-stratospheric temperature analyses using rocketsonde and lidar data, *Bull. Am. Meteorol. Soc., 74*, 789–799, 1993.

Grant, W. B., E. V. Browell, C. S. Long, L. L. Stowe, R. G. Grainger and Alyn Lambert, Use of volcanic aerosols to study the tropical stratospheric reservoir, *J. Geophys. Res., 101*, 3973–3988, 1996.

Harvey, V. L., M. H. Hitchman, R. B. Pierce and T. D. Fairlie, Tropical aerosol in the Aleutian High, *J. Geophys. Res., 104*, 6281–6290, 1999.

Hess, P. G., Mixing processes following the final stratospheric warming, *J. Atmos. Sci., 48*, 1625–1641, 1991.

Hitchman, M. H., M. McKay and C. R. Trepte, A climatology of stratospheric aerosol, *J. Geophys. Res., 99*, 20,689–20,700, 1994.

Holton, J. R., A dynamically based transport parameterization for one-dimensional photochemical models of the stratosphere, *J. Geophys. Res., 91*, 2681–2686, 1986.

Ko, M.K.W., N.K. Sze and D.K.Weisenstein, Use of satellite data to constrain the model-calculated atmospheric lifetime for N_2O: implications for other trace gases, *J. Geophys. Res., 96*, 7547–7552, 1991.

Koshyk, J.N., B.A. Boville, K. Hamilton, E. Manzini and K. Shibata, The kinetic energy spectrum of horizontal motions in middle-atmosphere models, *J. Geophys. Res., in press*, 1999.

Mahlman, J. D., Mechanistic interpretation of stratospheric tracer transport, *Advances in Geophysics*, Vol. *28A*, *Climate Dynamics*, B. Saltzman, Ed., Academic Press, 301–320, 1985.

Majda, A. J., and P. R. Kramer, Simplified models for turbulent diffusion: theory, numerical modelling, and physical phenomena, *Phys. Repts. 314*, 238–574, 1999.

Morris, G. A., M. R. Schoeberl, L. C. Sparling, P. A. Newman, L. R. Lait, L. Elson, J. Waters, R. A. Suttie, A. Roche, J. Kumer, and J. M. Russell, Trajectory Mapping and Applications to Data from the Upper Atmosphere Research Satellite, *J. Geophys. Res., 100*, 16491-16505, 1995.

Leovy, C.B., C.-R. Sun, M. H. Hitchman, E. E. Remsburg, I.J.M Russell, L.L.Gordley, J. C. Gille, and L. V. Lyjak, Transport of ozone in the middle stratosphere: Evidence for planetary wave breaking, *J. Atmos. Sci., 42*, 230–244, 1985.

McIntyre, M. E., Towards a Lagrangian-mean description of stratospheric circulations and chemical transport. *Phil. Trans. Roy. Soc. London, A296*, 129–148, 1980.

Nakamura, N., Two-dimensional mixing, edge formation, and permeability diagnosed in an area coordinate, *J. Atmos. Sci., 53*, 1524–1537, 1996.

Nash, E. R., P. A. Newman, J. E. Rosenfield and M. R. Schoeberl, An objective determination of the polar vortex using Ertel's potential vorticity, *J. Geophys. Res., 101*, 9471-9478, 1996.

Pawson, S. and Kubitz, Climatology of planetary waves in the northern stratosphere, *J. Geophys. Res., 101*, 16,987–16,996, 1996.

Pierrehumbert, R. T., Tracer microstructure in the large-eddy dominated regime, *Chaos Soliton Fract*, 4, 1091-1110, 1994.

Pierrehumbert, R. T., Lattice models of advection-diffusion, *submitted.*

Pierce, R. B. and T. D. A. Fairlie, Observational evidence of preferred flow regimes in the northern-hemisphere winter stratosphere, *J. Atmos. Sci.*, 50, 1936–1949, 1993.

Pierce, R.B., W.L. Grose, and J.M. Russell, Evolution of southern hemisphere air masses observed by HALOE, *Geophys. Res. Lett., 21(3)*, 213–216, 1994.

Plumb, R. A., A "tropical pipe" model of stratospheric transport, *J. Geophys. Res., 101*, 3957–3972, 1996.

Randel, W., Global Atmospheric circulation statistics, 1000-1 mb, NCAR technical note, NCAR/TN-366+STR, 1992.

Randel W. J., J. C. Gille, A. E. Roche, J. B. Kumer, J. L. Mergenthaler, J. W. Waters, E. F. Fishbein, W. A. Lahoz, Stratospheric transport from the tropics to middle latitudes by plantetary-wave mixing, *Nature, 365*, 533–535, 1993.

Remsberg, E.E. and P. P. Bhatt, Zonal variance of nitric acid vapor as an indicator of meridional mixing in the subtropical lower stratosphere, *J. Geophys. Res., 101*, 29,523–29,530, 1996.

Roche, A. E., et al, Validation of CH4 and N2O measurements by the cryogenic limb array etalon spectrometer instrument on the Upper Atmosphere Research Satellite, *J. Geophys. Res., 101*, 9679–9710, 1996.

Rosenfield, J. E., P. A. Newman and M. R. Schoeberl, Computations of diabatic descent in the stratospheric polar vortex, *J. Geophys. Res., 99*, 16,677–16,689, 1994.

Russell, J. M., L.L.Gordley, J. H. Park, S. R. Drayson, W. D. Hesketh, R. J. Cicerone, A. F. Tuck, J. E. Frederick, J. E. Harries and P. J. Crutzen, The Halogen Occultation Experiment, *J. Geophys. Res., 98*, 10,777–10,797, 1993.

Salby, M., Sampling theory for asynoptic satellite observations. Part II: Fast Fourier synoptic mapping, *J. Atmos. Sci., 39*, 2601–2614, 1982.

Schoeberl, M. R. and L. R. Lait, Conservative coordinate transformations for atmospheric measurements. *Proc. Internat. School Phys. "Enrico Fermi"*, Course CXV, ed. J. C. Gille and G. Visconti, North-Holland, 419–431, 1992.

Stolarski, R. S., et al., 1995 Scientific assessment of the atmospheric effects of stratospheric aircraft, NASA Ref. Publ. 1381, 64 pp., 1995.

Sutton, R. T., H. Maclean, R. Swinbank, A. O'Neill and F. W Taylor, High-resolution stratospheric tracer fields estimated from satellite observations using lagrangian trajectory calculations, *J. Atmos. Sci., 51*, 2995–3005, 1994.

Swinbank, R. and A. O'Neill, A stratosphere-troposphere data assimilation system, *Mon. Wea. Rev., 122*, 686-702, 1994.

Thompson, A. M., L. C. Sparling, Y. Kondo, B. E. Anderson, G. L. Gregory and G. W. Sachse, Perspectives on NO, NOy and fine aerosol sources and variability during SONEX, *Geophys. Res. Lett*, in press, 1999.

Volk, C. M., et al., Quantifying transport between the tropical and mid-latitude lower stratosphere, *Science, 272*, 1763-1768, 1996.

Waugh, D. W., Subtropical stratospheric mixing linked to disturbance of the polar vortices, *Nature, 365*, 535-537, 1993.

Yang, H., Three-dimensional transport of the Ertel potential vorticity and N_2O in the GFDL SKYHI model, *J. Atmos. Sci., 51*, 3427-3454, 1995.

L. C. Sparling, Atmospheric Chemistry and Dynamics Branch, Code 916, NASA/Goddard Space Flight Center, MD 20771, sparling@dynarama.gsfc.nasa.gov, (http://hyperion.gsfc.nasa.gov/).

Received ______________________

Figure 1. Zonal mean contour plot of CLAES N_2O, from measurements accumulated over 930101-930118.

Figure 2. CLAES spatial sampling during a UARS north facing period. The large dots show the sampling pattern for a single day; the small dots show the nearly uniform coverage when measurements are accumulated over 4 days.

Figure 3. Left: Scatter plot of CLAES N_2O mixing ratio vs latitude at 30 km (15 mb); Right: equal area PDF $P(\chi)$ vs χ, turned sideways. (a) SH winter, 920817-920916; (b) NH summer 920719-920810. The center of the tropics was at 5N during this time. See example A in text.

Figure 4. $P(\phi \mid \chi^*)$, the PDF of the latitude support of the NH and SH tracer boundaries, for the same altitude and time period as in Figure 3. See example B in text.

Figure 5. Calculation of chemical distinctness δ and overlap area A of the midlatitude and tropical tracer PDFs, for late December 1992, at altitude 32km (10mb). Subscript m refers to midlatitudes (ϕ^* to 55N), subscript t refers to tropics (equator to ϕ^*), and ϕ^*=17N is the subtropical boundary. See example C in text.

Figure 6. Altitude dependence of the chemical distinctness δ and overlap area A of the tropical (equator to ϕ^*) and midlatitude (ϕ^* to 55N) tracer PDFs during 921222-930105. ϕ^* is the subtropical boundary at each altitude (shown in Fig. 8). See example C in text.

Figure 7. Top: χ vs. ϕ scatter plot of measurements at z=32 km (10mb), during 921222-930105, with the zonal mean $\overline{\chi}(\phi)$ and zonal mode $\tilde{\chi}(\phi)$ overplotted. Bottom: $d\overline{\chi}(\phi)/d\phi$ (thin) $d\tilde{\chi}(\phi)/d\phi$ (thick) and σ (dashed), the standard deviation about the zonal mean, vs. latitude . See examples D and E in text.

Figure 8. The tropical pipe during 921222-930105, computed from the most probable latitude support (dashed), and the extrema in $d\tilde{\chi}(\phi)/d\phi$ (solid). See examples B and E in text.

Figure 9. $P(\epsilon)$, the distribution of fluctuations from the zonal mode $\tilde{\chi}(\phi)$ on the 10mb level, for the observations in Figure 7a, and for each of the regions a-d indicated there. Measurements in region (d) are during 930113-930127. The standard deviation of the Gaussian core is indicated on each plot. See example F in text.

Figure 10. Time evolution of the NH (0 to 80N) tracer PDF on the 950K potential temperature surface. Top: early and late October 1992. Middle: December 1992, during a warming. Bottom: early January and mid February 1993. Each curve is the PDF of measurements accumulated over 4 consecutive days. See example G in text.

Figure 11. Simultaneous increase in the area occupied by high temperatures, strong diabatic cooling and low mixing ratios on the 950K surface during the December 92 warming (see Fig. 10b). A_χ is the area over which $\chi \leq 50$ ppbv, A_T is the area over which the temperatures are $\geq 230K$, and A_Q is the area over which the diabatic cooling rate is less than -5K/day. Areas are expressed as fractions of the midlatitude (35N–55N) area. See example G in text.

Figure 12. Time evolution of the zonal mode on the NH (0-80N) 950K surface during December 1992 (Fig. 10b). See example G in text.

Figure 13. Time evolution of the "number of states" $\Omega \equiv e^S$ on the 950K surface, from 0-80N, during December 1992 (Fig. 10b). $\Omega(t)$ is computed from data in running 3 day intervals. The mixing ratios are binned into 50 bins, thus Ω_{max}=50.

Figure 14. Altitude dependence of tracer PDFs near the pole (50N-80N) during the 1993 NH vortex breakdown. Measurements were accumulated over two consecutive 2 week time periods starting on 930213(shaded) and 930301 (line). See example H in text.

Figure 15. Top: The vortex ($\chi < 26$ ppbv; black dots), surf zone($26 \leq \chi < 90$ ppbv; small dots) and anticyclone/subtropical($90 \leq \chi < 150$ ppbv; grey dots) air masses on the 10mb surface, 930307-930314. The chemically defined boundaries between the air masses are the minima in the 10mb PDF in Figure 14. The thick line is at 50N. Bottom: PV RDF on 850K surface, 930310; parcels initialized 10 days earlier. The anticyclone is over Canada, the polar vortex is the comma-shaped region of high PV. (PV figure courtesy Paul Newman, GSFC). Map projection is orthographic. See example H in text.

Figure 16. Shaded: PDF of all 21,000 measurements on the 10 mb level, from 0–50N during 920925–921025. Black line: PDF of an unbiased subsample of 350 measurements, chosen randomly in space and time. Dotted line: PDF of 350 measurements closest in space and time to the HALOE measurements during this time period. See example I in text.

Figure 17. PDFs during SH summer, 930113-930206, 0-80S on the 10mb surface. Top: each curve is the PDF of all SH data, in disjoint 4 day intervals. Bottom: each curve is the PDF of the time-integrated data in disjoint 30° longitude slices.

Figure 18. PDFs in disjoint 30° longitude slices as in Fig. 17(bottom); data accumulated over the NH fall period 920925-921025, from 0-80N, on the 10 mb surface. Two curves are highlighted: $0 \leq \lambda < 30$E (grey), 150E$\leq \lambda <$180E (black).

Figure 19. (a) Equal area map projection of the UKMO potential vorticity analysis on the 500K surface, Dec 1, 1995. The contour interval $\Delta q=6$ pvu ($1\text{pvu}=10^{-6}\text{m}^2\text{s}^{-1}\text{K}\,\text{kg}^{-1}$) and 90E is at the bottom of the graph. An equal area distribution of points covers the area enclosed by the $|q|=24$ and 30 pvu contours. (b) Equal area PDF of the PV field in (a). P(q) is the fractional area between PV contours q and $q+\Delta q$, with bin size $\Delta q=6$ pvu. The PV bins with the largest area are in the northern and southern surf zones and are shaded; they correspond to the area of the dotted regions in (a). The thin vertical lines mark the vortex edge, as computed from the algorithm of *Nash et al.* [1996]. See example K in text.

Figure 20. Comparison of GSFC (shaded) and UKMO (line) Potential vorticity PDFs on the 500K surface for the 4 day period starting on Nov 27, 1995. The UKMO 4-day PDF shows how the multimodal structure becomes clearer when more data is used (compare with the one-day PDF in Fig. 19b). See example L in text.

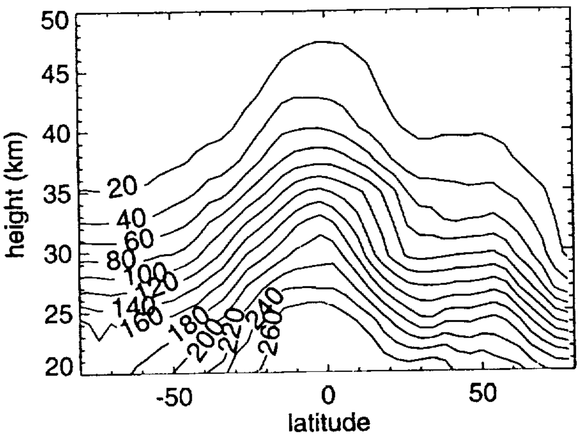

Figure 1. Zonal mean contour plot of CLAES N_2O, from measurements accumulated over 930101-930118.

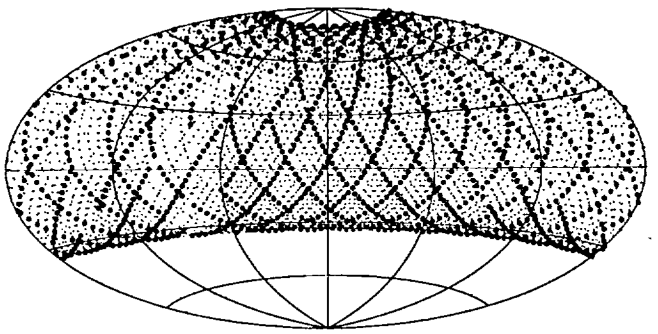

Figure 2. CLAES spatial sampling during a UARS north facing period. The large dots show the sampling pattern for a single day; the small dots show the nearly uniform coverage when measurements are accumulated over 4 days.

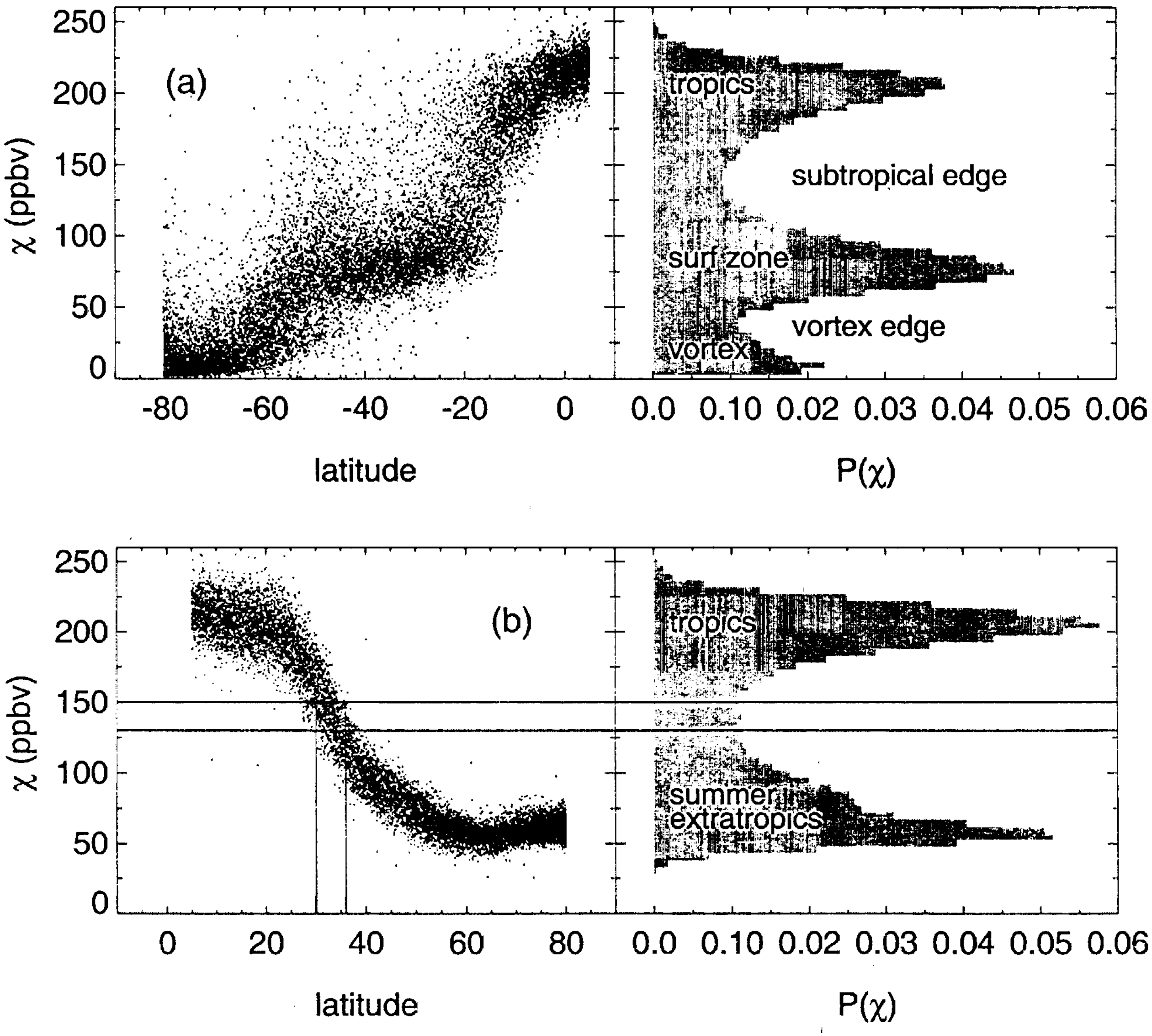

Figure 3. Left: Scatter plot of CLAES N_2O mixing ratio vs latitude at 30 km (15 mb); Right: equal area PDF P(χ) vs χ, turned sideways. (a) SH winter, 920817-920916; (b) NH summer 920719-920810. The center of the tropics was at 5N during this time.

See example A in text.

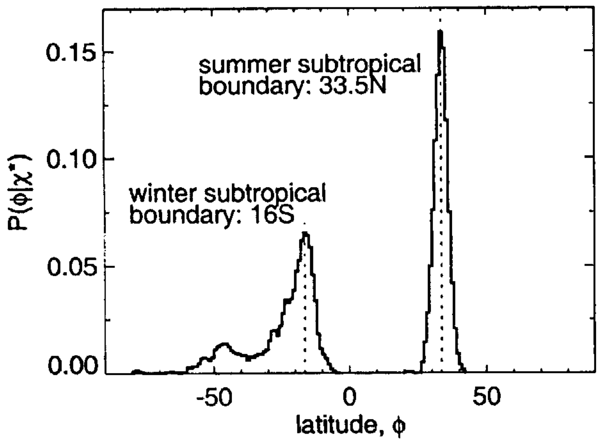

Figure 4. P(ϕ | χ^*), the PDF of the latitude support of the NH and SH tracer boundaries, for the same altitude and time period as in Figure 3b. See example B in text.

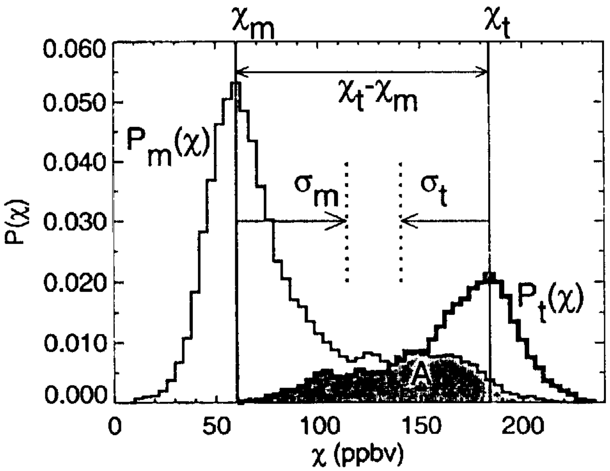

Figure 5. Calculation of chemical distinctness δ and overlap area A of the midlatitude and tropical tracer PDFs, for late December 1992, at altitude 32km (10mb). Subscript m refers to midlatitudes (ϕ^* to 55N), subscript t refers to tropics (equator to ϕ^*), and ϕ^*=17N is the subtropical boundary. See example C in text.

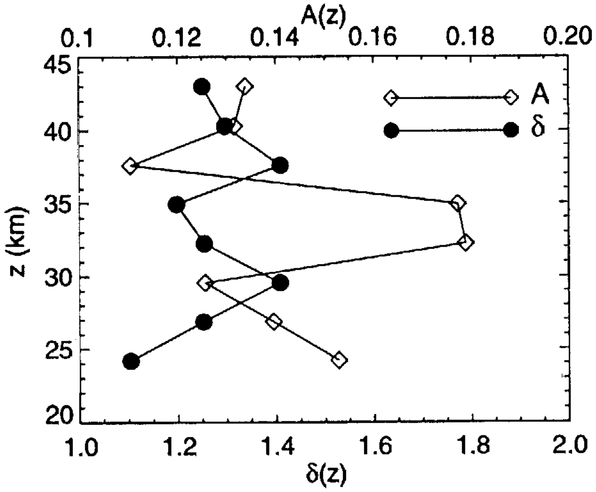

Figure 6. Altitude dependence of the chemical distinctness δ and overlap area A of the tropical (equator to ϕ^*) and midlatitude (ϕ^* to 55N) tracer PDFs during 921222-930105. ϕ^* is the subtropical boundary at each altitude (shown in Fig. 8). See example C in text.

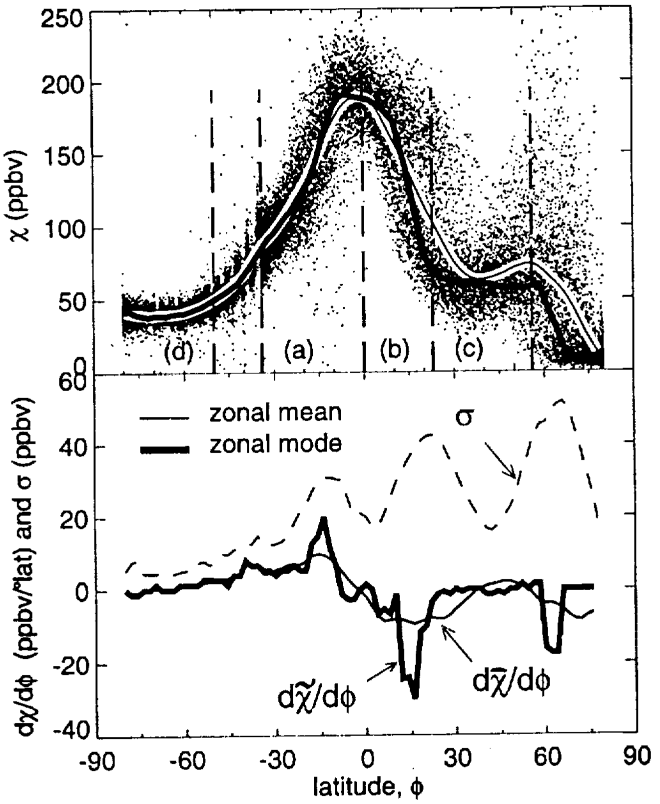

Figure 7. Top: χ vs. ϕ scatter plot of measurements at z=32 km (10mb), during 921222-930105, with the zonal mean $\bar{\chi}(\phi)$ and zonal mode $\tilde{\chi}(\phi)$ overplotted. Bottom: $d\bar{\chi}(\phi)/d\phi$ (thin) $d\tilde{\chi}(\phi)/d\phi$ (thick) and σ (dashed), the standard deviation about the zonal mean, vs. latitude . See examples D and E in text.

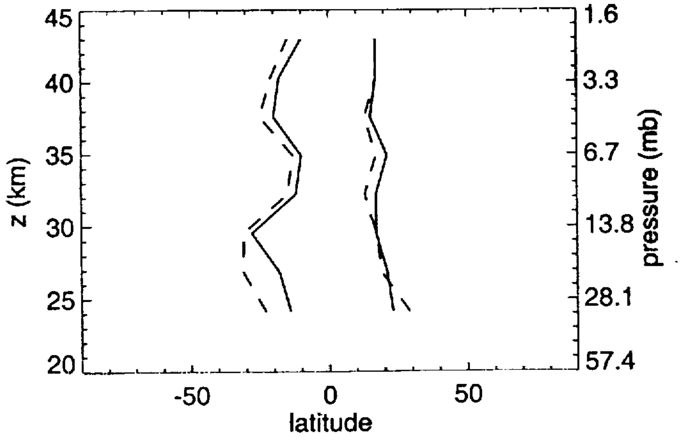

Figure 8. The tropical pipe during 921222-930105, computed from the most probable latitude support (dashed), and the extrema in $d\tilde{\chi}(\phi)/d\phi$ (solid). See examples B and E in text.

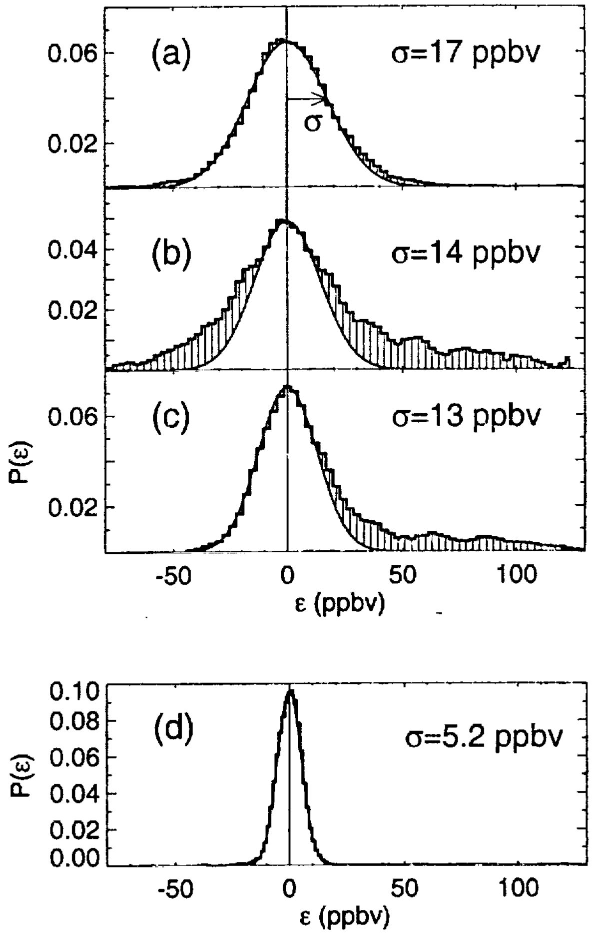

Figure 9. P(ϵ), the distribution of fluctuations from the zonal mode $\tilde{\chi}(\phi)$ on the 10mb level, for the observations in Figure 7a, and for each of the regions a-d indicated there. Measurements in region (d) are during 930113-930127. The standard deviation of the Gaussian core is indicated on each plot. See example F in text.

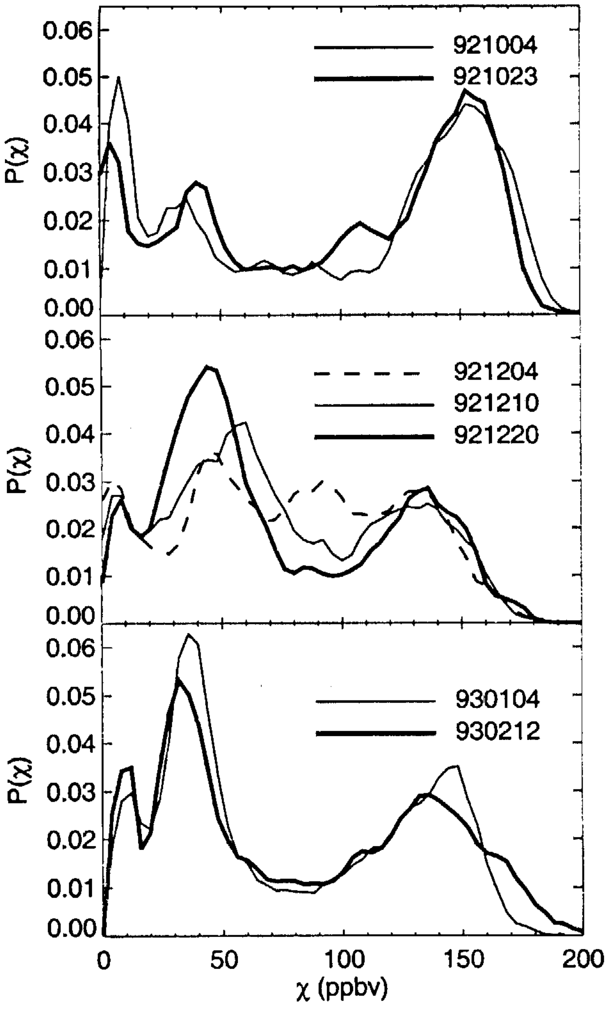

Figure 10. Time evolution of the NH (0 to 80N) tracer PDF on the 950K potential temperature surface. Top: early and late October 1992. Middle: December 1992, during a warming. Bottom: early January and mid February 1993. Each curve is the PDF of measurements accumulated over 4 consecutive days. See example G in text.

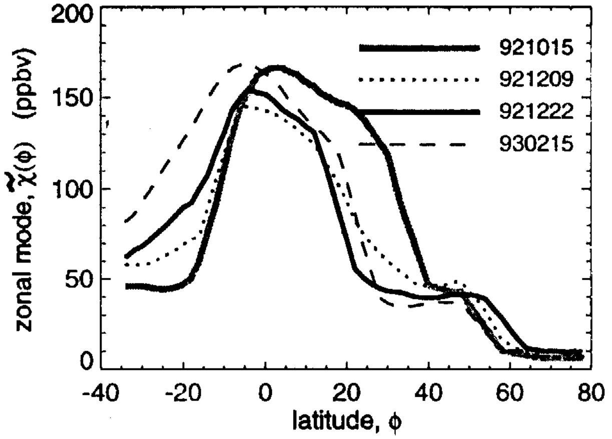

Figure 11. Time evolution of the zonal mode on the NH (0-80N) 950K surface during December 1992 (Fig. 10b). See example G in text.

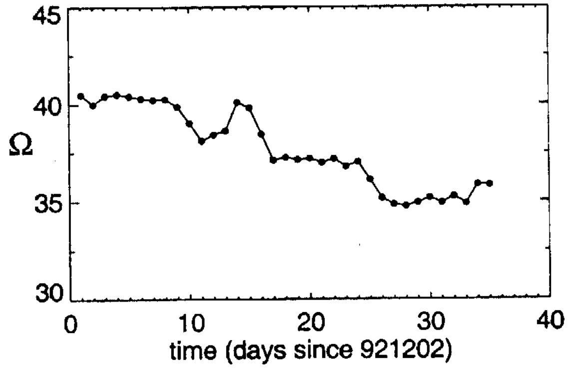

Figure 13. Time evolution of the "number of states" $\Omega \equiv e^S$ on the 950K surface, from 0-80N, during December 1992 (Fig. 10b). $\Omega(t)$ is computed from data in running 3 day intervals. The mixing ratios are binned into 50 bins, thus Ω_{max}=50.

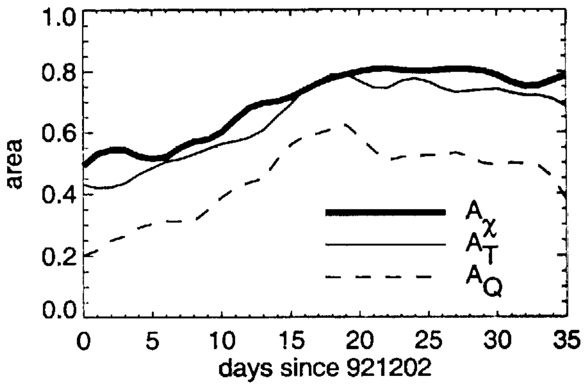

Figure 12. Simultaneous increase in the area occupied by high temperatures, strong diabatic cooling and low mixing ratios on the 950K surface during the December 92 warming (see Fig. 10b). A_χ is the area over which $\chi \leq 50$ ppbv, A_T is the area over which the temperatures are $\geq 230K$, and A_Q is the area over which the diabatic cooling rate is less than -5K/day. Areas are expressed as fractions of the midlatitude (35N–55N) area. See example G in text.

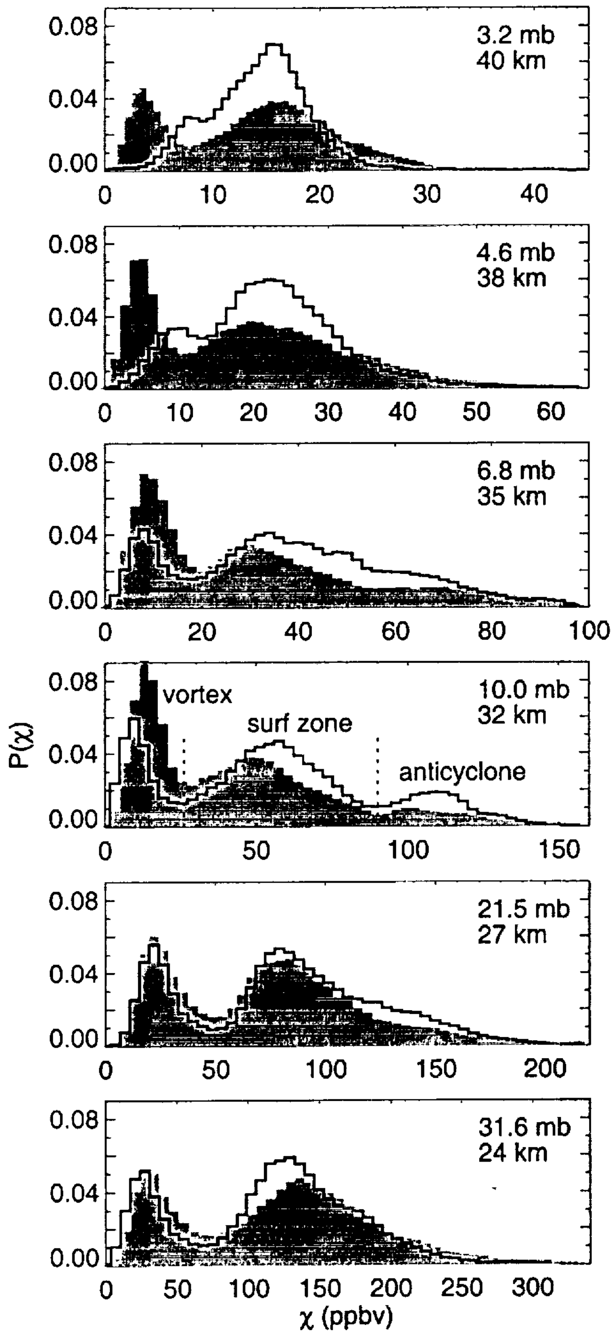

Figure 14. Altitude dependence of tracer PDFs near the pole (50N-80N) during the 1993 NH vortex breakdown. Measurements were accumulated over two consecutive 2 week time periods starting on 930213(shaded) and 930301 (line). See example H in text.

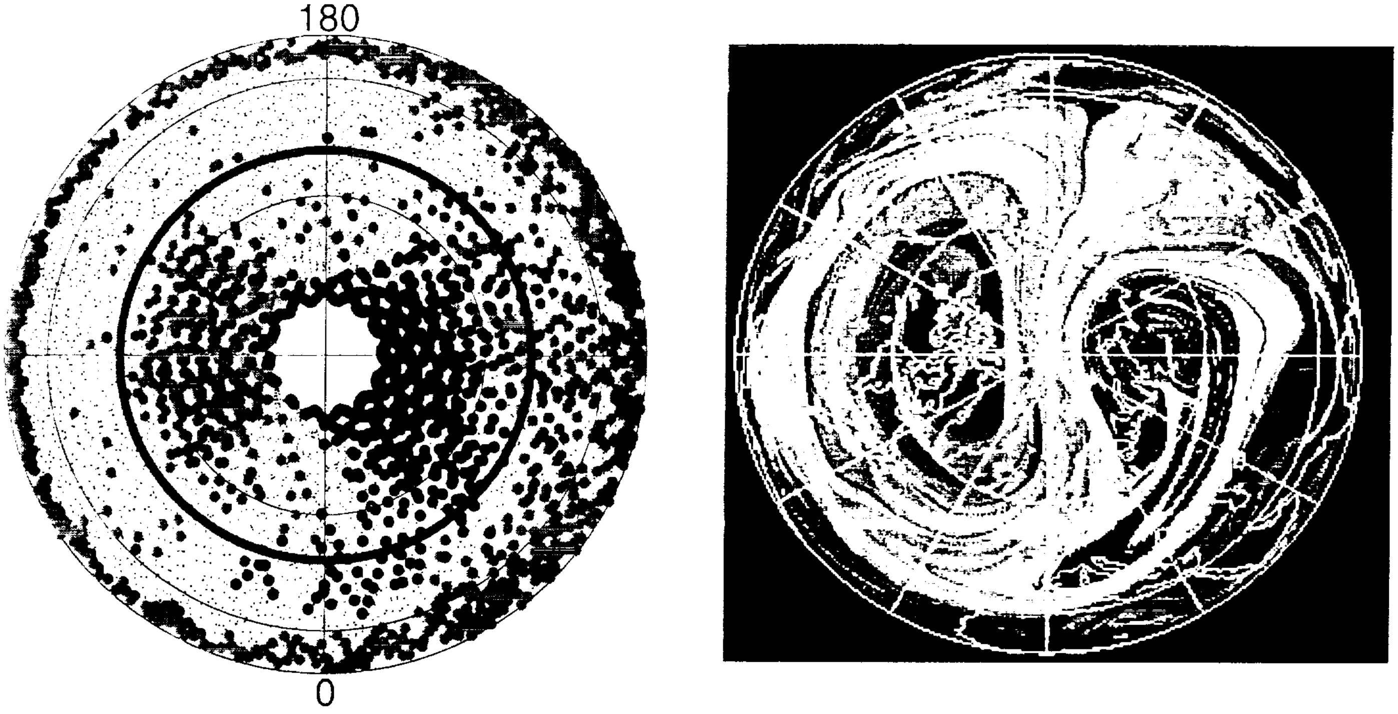

Figure 15. Top: The vortex ($\chi < 26$ ppbv; black dots), surf zone($26 \leq \chi < 90$ ppbv; small dots) and anticyclone/subtropical($90 \leq \chi < 150$ ppbv; grey dots) air masses on the 10mb surface, 930307-930314. The chemically defined boundaries between the air masses are the minima in the 10mb PDF in Figure 14. The thick line is at 50N. Bottom: PV RDF on 850K surface, 930310; parcels initialized 10 days earlier. The anticyclone is over Canada, the polar vortex is the comma-shaped region of high PV. (PV figure courtesy Paul Newman, GSFC). Map projection is orthographic. See example H in text.

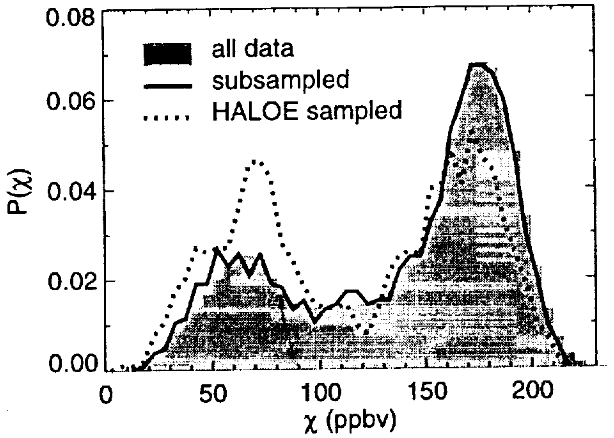

Figure 16. Shaded: PDF of all 21,000 measurements on the 10 mb level, from 0–50N during 920925–921025. Black line: PDF of an unbiased subsample of 350 measurements, chosen randomly in space and time. Dotted line: PDF of 350 measurements closest in space and time to the HALOE measurements during this time period. See example I in text.

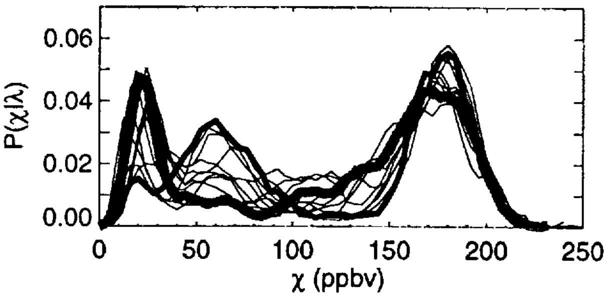

Figure 18. PDFs in disjoint 30° longitude slices as in Fig. 17(bottom); data accumulated over the NH fall period 920925-921025, from 0-80N, on the 10 mb surface. Two curves are highlighted: $0 \leq \lambda < 30E$ (grey), $150E \leq \lambda < 180E$ (black). See example J in text.

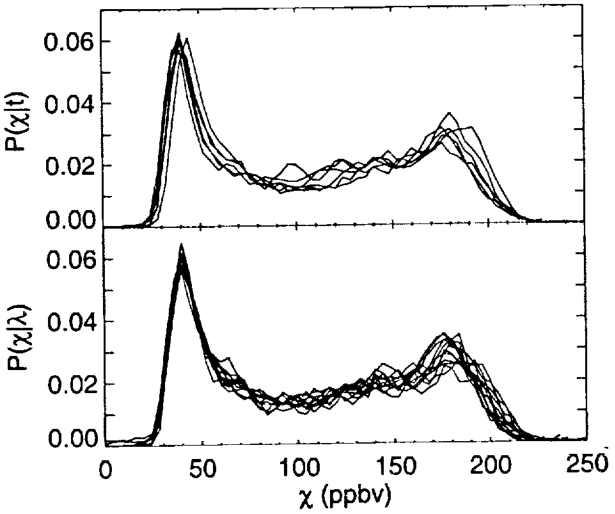

Figure 17. PDFs during SH summer, 930113-930206, 0-80S on the 10mb surface. Top: each curve is the PDF of all SH data, in disjoint 4 day intervals. Bottom: each curve is the PDF of the time-integrated data in disjoint 30° longitude slices. See example J in text.

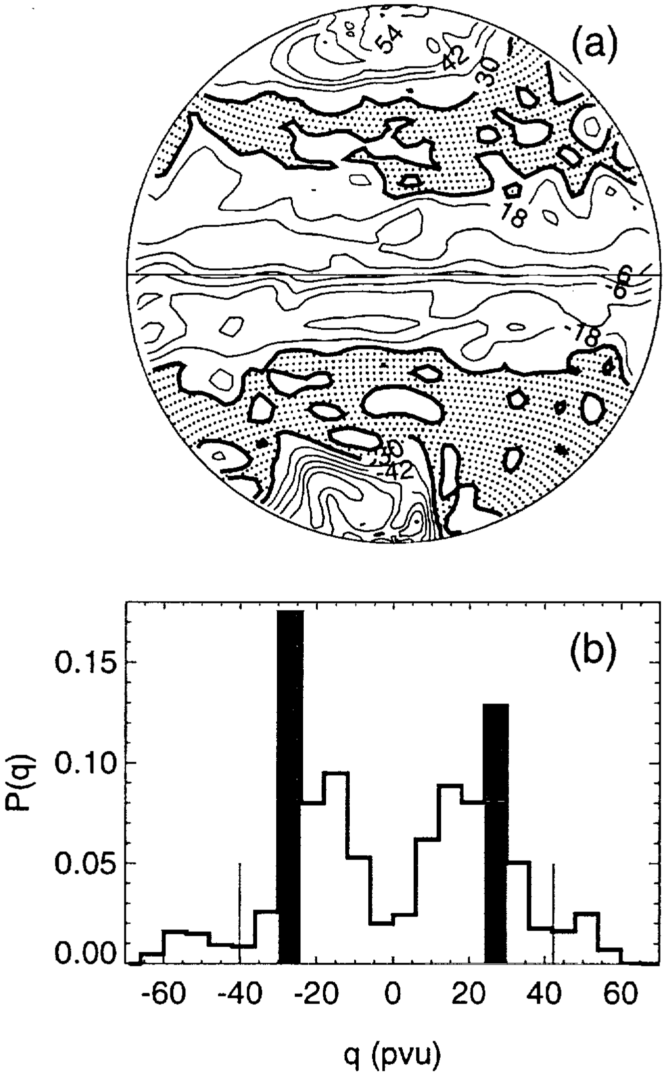

Figure 19. (a) Equal area map projection of the UKMO potential vorticity analysis on the 500K surface, Dec 1, 1995. The contour interval Δq=6 pvu (1pvu=10^{-6}m^{2}s^{-1}K kg^{-1}) and 90E is at the bottom of the graph. An equal area distribution of points covers the area enclosed by the $|q|$ =24 and 30 pvu contours. (b) Equal area PDF of the PV field in (a). P(q) is the fractional area between PV contours q and q+Δq, with bin size Δq=6 pvu. The PV bins with the largest area are in the northern and southern surf zones and are shaded; they correspond to the area of the dotted regions in (a). The thin vertical lines mark the vortex edge, as computed from the algorithm of *Nash et al.* [1996]. See example K in text.

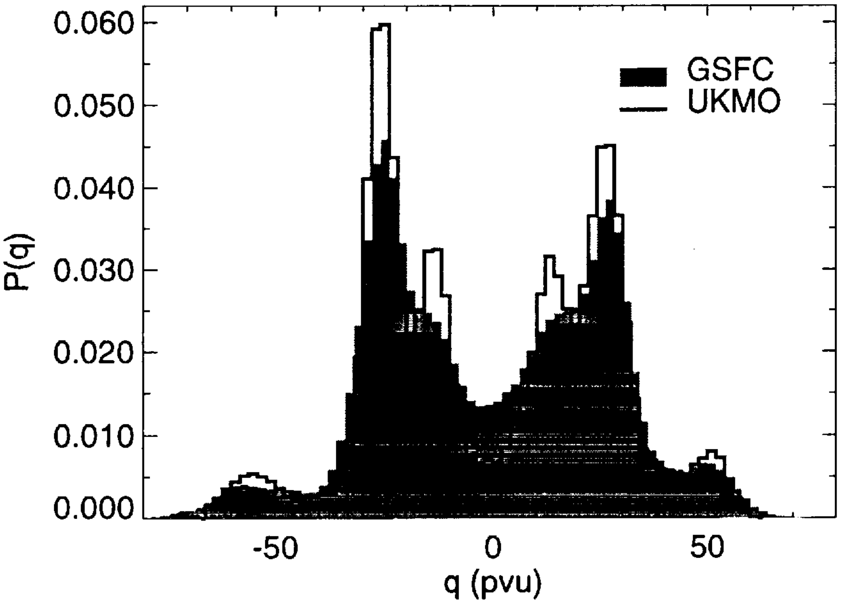

Figure 20. Comparison of GSFC (shaded) and UKMO (line) Potential vorticity PDFs on the 500K surface for the 4 day period starting on Nov 27, 1995. The UKMO 4-day PDF shows how the multimodal structure becomes clearer when more data is used (compare with the one-day PDF in Fig. 19b). See example L in text.