**Atmospheric
Chemistry
and Physics
Discussions**

A new real-time Lagrangian diagnostic system for stratosphere-troposphere exchange: evaluation during a balloon sonde campaign in eastern Canada

M. S. Bourqui[1], A. Yamamoto[1], D. Tarasick[2], M. D. Moran[2], L.-P. Beaudoin[4],
I. Beres[2], J. Davies[2], A. Elford[2], W. Hocking[3], M. Osman[3], and R. Wilkinson[4]

[1]Department of Atmospheric and Oceanic Sciences, Department of Chemistry, McGill University, Montréal, Québec, Canada
[2]Air Quality Research Division, Environment Canada, Downsview, Ontario, Canada
[3]Department of Physics and Astronomy, University of Western Ontario, London, Ontario, Canada
[4]Canadian Space Agency, St. Hubert, Québec, Canada

Correspondence to: M. S. Bourqui (michel.bourqui@mcgill.ca)

"Evaluation of new STE diagnosis system

M. S. Bourqui et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

|◄ | ►|
◄ | ►
Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

"Evaluation of new STE diagnosis system

M. S. Bourqui et al.

Abstract

A new real-time Lagrangian diagnostic system for stratosphere-troposphere exchange (STE) developed for Environment Canada (EC) has been delivering daily archived data since 20 July 2010. The STE calculations are performed globally following the Lagrangian approach proposed in Bourqui (2006) using medium-range, high-resolution operational global weather forecasts. Following every weather forecast, trajectories are started from a dense three-dimensional grid covering the globe, and are calculated for six days of the forecast. All trajectories crossing either the dynamical tropopause (± 2 PVU) or the 380 K isentrope and having a residence time greater than 12 h are archived, and also used to calculate several diagnostics. This system provides daily global STE forecasts that can be used to guide field campaigns, among other applications. The archived data set offers unique high-resolution information on transport across the tropopause for both extra-tropical hemispheres and the tropics. This will be useful for improving our understanding of STE globally, and as a reference for the evaluation of lower-resolution models. This new data set is evaluated here against measurements taken during a balloon sonde campaign with daily launches from three stations in eastern Canada (Montreal, Egbert, and Walsingham) for the period from 12 July to 4 August 2010. The campaign found an unexpectedly high number of observed stratospheric intrusions: 79 % (38 %) of the profiles appear to show the presence of stratospheric air below 500 hPa (700 hPa). An objective identification algorithm developed for this study is used to identify layers in the balloon-sonde profiles affected by stratospheric air and to evaluate the Lagrangian STE forecasts. We find that the predictive skill for the overall intrusion depth is excellent for intrusions penetrating down to 300 and 500 hPa, while it becomes negligible for intrusions penetrating below 700 hPa. Nevertheless, the statistical representation of these deep intrusions is reasonable, with an average bias of 24 %. Evaluation of the skill at representing the detailed structures of the stratospheric intrusions shows good predictive skill down to 500 hPa, reduced predictive skill between 500 and 700 hPa, and none below. A significant low statistical

bias of about 30 % is found in the layer between 500 to 700 hPa. However, analysis of missed events at one site, Montreal, shows that 70 % of them coincide with candidate clusters of trajectories that pass through Montreal, but that are too dispersed to be detected in the close neighbourhood of the station. This allows us to expect a negligible bias throughout the troposphere in the spatially averaged STE frequency derived from this data set, for example in climatological maps of STE mass fluxes.

1 Introduction

Stratosphere-Troposphere Exchange (STE) is known to control the chemical composition of the lower stratosphere and the upper troposphere to a great extent. Many previous studies have investigated the mechanisms leading to extra-tropical STE in the northern hemisphere and have shown that STE typically occurs during the breaking of baroclinic waves at synoptic- to meso-scales and involves a variety of processes such as those associated with clouds, radiative relaxation, gravity wave breaking, wind shear turbulence, and numerical diffusion (Shapiro, 1978; Langford et al., 1996; Zierl and Wirth, 1997; Stohl et al., 2003; Gray, 2003; Mullendore et al., 2005; Bourqui, 2006). This complexity is unimportant when considering only the global net flux of mass across the tropopause (Holton et al., 1995; Appenzeller et al., 1996). However, because chemical gradients across the tropopause are large, it is the separate stratosphere-to-troposphere (S→T) and troposphere-to-stratosphere (T→S) mass fluxes that control the chemical composition around the tropopause, not the net mass flux. An accurate estimate of separate S→T and T→S mass fluxes requires to resolve the complex processes leading to STE with their underlying scales. Bourqui (2006) suggests that consideration of hourly meteorological fields with horizontal resolution of $0.5° \times 0.5°$ is necessary to resolve these processes adequately. Furthermore, the nature of the separate S→T and T→S mass fluxes is such that they become infinite at infinitely small scales (Hall and Holzer, 2003). In a finite-grid representation, these fluxes are essentially determined by the smallest resolved scales and therefore do not provide relevant

M. S. Bourqui et al.

information on STE. This ill-defined character is probably responsible for a good part of the large range of estimates of separated STE fluxes found in the literature (e.g., Hoerling et al., 1993; Gettelman and Sobel, 2000; Sprenger and Wernli, 2003). To render the fluxes well-defined, it is necessary to include an explicit scale constraint such as the *residence time* (Wernli and Bourqui, 2002; Hall and Holzer, 2003; Bourqui, 2006).

Wernli and Bourqui (2002) and Sprenger and Wernli (2003), using European Center for Medium-Range Weather Forecasts (ECMWF) analyses and ERA-15 reanalyses, respectively, provided such S→T and T→S mass flux estimates for the extra-tropical northern hemisphere using a Lagrangian approach with a residence time criterion. However, the absence of independent hemispheric estimates of separated STE fluxes in the literature makes it difficult to evaluate the uncertainties in their results. Wernli and Bourqui (2002) and Sprenger and Wernli (2003) suggested that S→T mass fluxes dominantly occur in regions of storm tracks in the winter and above continents in the summer. They also suggested that rapid transport can bring stratospheric air to the lower troposphere within a time scale of several days, and that in the northern extra-tropics, the lower tropospheric destinations of such rapid transport are concentrated on both sides of the north American continent. Since then, a few studies have been successful at modelling individual events (Cooper et al., 2005; Gerasopoulos et al., 2006). More recently Bourqui and Trépanier (2010) found an intense activity of such rapid transport, confirmed by ozone sonde measurements, during a summer month over North America. The mass flux associated with this deep S→T transport activity was found to be about one order of magnitude larger than that estimated by Sprenger and Wernli (2003). A similarly high contribution of stratospheric ozone in the lower troposphere was found by Lefohn et al. (2011) on average during the years 2004-2006. This suggests that Wernli and Bourqui (2002) and Sprenger and Wernli (2003) may significantly underestimate the frequency of deep STE events.

Estimates of cross-tropopause fluxes of ozone from chemistry transport models (CTMs) were summarised in Stevenson et al. (2006) with a focus on the tropospheric ozone budget. The CTMs had varying representations of stratospheric ozone, such as

interactive stratospheric chemistry, specified ozone climatologies, and specified ozone vertical fluxes. Their horizontal resolutions were of the order of $2° \times 2°$ or coarser. The stratospheric source of ozone was found on average to contribute at the same level as the net photochemistry, but the model-to-model variability was very large. S→T fluxes were predicted to increase by an average of 8 % by 2030 under a climate change scenario. Hegglin and Shepherd (2009) used a low-resolution Climate-Chemistry Model (CCM) and estimated that S→T fluxes of ozone may increase by 23 % between 1965 and 2095 as a result of climate change. Terao et al. (2008) used a CTM with similar resolution and interactive stratospheric chemistry, and estimated using correlations in the stratospheric and tropospheric ozone variability that the stratosphere contributes to 40 % and 30 % of the 500 hPa ozone budget in high and mid-latitudes, respectively. Using ozonesonde data, Tarasick et al. (2005) showed that both trends and interannual variability of tropospheric ozone values over Canada are highly correlated with lower stratospheric ozone amounts. This suggests that STE may be an important factor in determining tropospheric ozone trends. Similar results were found at European sites (Ordóñez et al., 2007). However, observations of ozone from balloon sondes reflect complex, inhomogeneous trends over the past three decades (Oltmans et al., 2006). Knowledge of STE is still too incomplete to disentangle the relative contributions to these trends from in-situ photochemistry and from the stratosphere.

A global, high-resolution climatology of separate S→T and T→S mass fluxes is clearly needed to help evaluate the accuracy of STE fluxes predicted in CTMs and CCMs and to improve our knowledge of a central contributor to the chemical composition of the troposphere and lower stratosphere and its variability on all time scales.

Global weather forecasts offer an interesting, quite unexplored avenue for developing global Lagrangian estimates of STE: (i) they are free of the non-physical corrections introduced by data assimilation in (re-)analyses; (ii) they offer higher frequency data than the typical 6-hourly (re-)analysis data; (iii) their resolution is now high enough for an accurate representation of the processes responsible for STE; (iv) they allow the real-time forecasting of STE; and (v) they allow evaluation of STE forecasts against

observations, something that is not generally possible when STE is estimated using re-analysis data sets. Their disadvantage is that they introduce an error related to the divergence of the global forecasts from the true atmospheric evolution, which increases with the length of the forecast. This error is however not devoid of interest, as it provides

5 a Lagrangian perspective on forecast errors.

Here, we introduce the first real-time Lagrangian STE data set based on global weather forecasts. These data have been calculated daily since 20 July 2010 at Environment Canada (EC) following the methodology introduced in Bourqui (2006). They consist of daily Lagrangian STE forecasts composed of all six-day trajectories cross-

10 ing either the ±2 PVU dynamic tropopause or the 380 K isentrope somewhere on the globe, within five successive 24 h time windows with a 12 h residence time. Including trajectories crossing either the ±2 PVU tropopause or the 380 K isentrope allows consideration of extra-tropical and tropical STE, respectively. In addition, this data set offers a diagnostic of the lower part of the stratospheric residual circulation in the extra-

15 tropics. The calculation of trajectories and necessary dynamical fields is made using the highest-available temporal resolution, namely hourly, native forecast data, and the spatial density of initial trajectories is taken to be of the same order as the spatial resolution of meteorological data. The complete set of selected trajectories is archived daily and several diagnostics are produced.

20 The goal of this paper is to introduce and provide a first evaluation of this new Lagrangian STE data set, against observations from a one-month balloon sonde campaign in summer 2010 involving three launch sites located in eastern Canada: Montreal (Quebec), Egbert (Ontario), and Walsingham (Ontario). We focus on the transport of stratospheric air into the troposphere, or in other words, on stratospheric intrusions.

25 This evaluation covers only one summer season and is restrained spatially. Since the STE data set is based on meteorological forecasts, whose skill may vary in space and time, the characterisation of errors made here may not be automatically generalised to other seasons and regions. Nevertheless, this is a first step towards understanding the capabilities and limitations of this new data set.

"Evaluation of new STE diagnosis system

M. S. Bourqui et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◀ | ▶

◀ | ▶

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

The next Section introduces in detail the calculation of the Lagrangian STE data set and its diagnostics. Section 3 describes the balloon sonde campaign, provides an analysis of the observations and the detection of stratospheric intrusions. The evaluation of the Lagrangian STE data set is then discussed in Sect. 4, followed by an analysis of errors and missed intrusions in Sect. 5 and conclusions in Sect. 6.

2 Description of the new Lagrangian stratosphere-troposphere exchange data set

The approach proposed by Bourqui (2006) is applied separately to each EC multi-day 00:00 UTC operational global weather forecast. These global weather forecasts are calculated at EC's Canadian Meteorological Centre using the Global Environmental Multi-scale (GEM) numerical weather prediction model (Côté et al., 1998a,b) and provide hourly meteorological fields with horizontal grid spacing of $0.3° \times 0.3°$, and 80 hybrid vertical levels up to 0.1 hPa. Trajectories are then computed using an adapted version of the LAGRANTO Lagrangian trajectory model (Wernli and Davies, 1997), which is based upon an Eulerian iterative scheme. The resulting data set is archived in the NetCDF (Network Common Data Form, http://www.unidata.ucar.edu/software/netcdf) format and represents about 4 Gb per day of selected trajectory data.

A first set of three-dimensional forward trajectories is started at 00:00 UTC for up to six days from a 15 million point three-dimensional grid covering the entire globe with even horizontal spacing of 55 km (distance equivalent to $0.5°$ latitude), and 5 hPa vertically between 600 and 10 hPa. A selection is performed during the trajectory calculation in order to keep calculating only those trajectories that cross either the ± 2 PVU or the 380 K iso-surfaces during the period [12 h, 36 h[UTC, with a residence time of 12 h. The residence time criterion requires the trajectories to stay for at least 12 h in the stratosphere and in the troposphere, just before/after the exchange time. This criterion efficiently avoids the selection of trajectories that oscillate around the tropopause due to numerical noise (Bourqui, 2006), and at the same time it eliminates the singularity

problem of one-way STE fluxes (Hall and Holzer, 2003). This selection of trajectories reduces the number of trajectories that need to be considered by a factor of 100, from around 15 million to 150 000. The three-dimensional positions of the selected trajectories are archived every hour together with potential vorticity and potential temperature values along each trajectory. Note that these trajectories include both S→T and T→S events. This procedure is then repeated for the same forecast for new sets of trajectories started at 24:00, 48:00, 72:00, and 96:00 UTC, representing STE in the periods [36 h, 60 h[, [60 h, 84 h[, [84 h, 108 h[, and [108 h, 132 h[UTC, respectively. These additional trajectories are also included in the data set, as they offer a forecasting capability over the next 6 days, and they allow estimation of the degradation of STE estimates along the six first days of the meteorological forecast. In addition, trajectories crossing either ±2 PVU or 380 K in the last time window are extended backward in time to 00:00 UTC. This allows the analysis of (rapid) upward transport as well.

Two principal diagnostics are then calculated. Firstly, geographical maps of the density of either stratosphere-to-troposphere (S→T) or troposphere-to-stratosphere (T→S) exchanges are produced globally for the ±2 PVU and 380 K iso-surfaces (not shown). Secondly, the S→T trajectories that reach the lower troposphere (700 hPa) are organised as clusters following Bourqui and Trépanier (2010). Two trajectories are considered to belong to the same cluster if their three-dimensional Euclidean separation distance, averaged over the time period when both are located below the tropopause and above the 700 hPa level, is smaller than 222 km (equivalent to 2° latitude). Note that the calculation of the Euclidean distance is modified with a factor of 100 applied to the vertical separation as a simple way to account for the aspect ratio of the atmosphere. For each cluster, the three-dimensional position of the center of mass and the mean radius are calculated along the time axis.

In this evaluation paper, we use exclusively the trajectories crossing 2 PVU in the S→T direction since only those are relevant for the comparison with the stratospheric intrusions found in the balloon sonde observations. Moreover, we restrict the evaluation to the most accurate ones, namely the trajectories started at 00:00 UTC and

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

representing exchange events within the time window [12 h, 36 h[UTC.

3 Balloon sonde observations and analysis

Balloon sondes were launched daily at 13:00 local daylight savings time (17 UTC) from 12 July to 4 August 2010 from three sites in eastern Canada: Montreal (Quebec: 45.3° N, 73.4° W), Egbert (Ontario: 44.2° N, 79.8°W), and Walsingham (Ontario: 42.6° N, 80.6°W). Egbert is situated about 500 km southwest of Montreal (and about 100 km north of Toronto) and Walsingham is situated about 600 km southwest of Montreal (and about 100 km southwest of Toronto). The balloons were equipped with a radiosonde and an ECC ozonesonde (Komhyr, 1969) to measure vertical profiles of ozone partial pressure, relative humidity (RH), temperature, pressure, and wind speed and direction. The accuracy of ozone partial pressure and relative humidity measurements is expected to be better than 5 %. The vertical distance between successive measurements in the profiles is 0.3 hPa on average and 1.8 hPa maximum. The number of successful balloon-sonde launches over the 24 days of the campaign was 22 from Montreal, 17 from Egbert and 17 from Walsingham (see Table 1).

3.1 Temporal evolution of observations of ozone and relative humidity

Figure 1 shows the daily profiles of RH and ozone mixing ratio as a function of pressure, averaged in 10 hPa vertical bins, at Montreal for the entire campaign. The black solid line denotes the thermal tropopause estimated from the measured temperature profiles using the WMO definition (WMO, 1986). The RH profiles show the presence of dry air in the troposphere over Montreal persisting throughout the entire campaign, with RH values typically lower than 30 %. The bottom of this dry air region undulates between 600 and 800 hPa. The ozone molar mixing ratio profiles show a reverse pattern, with relatively elevated ozone concentrations in the regions of dry air. The anticorrelation between ozone and relative humidity is even more striking for individual profiles

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

⏮ | ⏭

◀ | ▶

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

(not shown). Such anomalously dry air located below the tropopause with a marked anticorrelation between RH and ozone is the typical signature of the presence of air of stratospheric origin in the troposphere. It is important to note, however, that layers of anomalously dry and ozone-rich air identified within single balloon-sonde profiles may
5 not always have their origin above the 2 PVU tropopause. They can have their origin above any level in the upper troposphere/lower stratosphere region with sufficiently low humidity and high ozone mixing ratios. Since the 2 PVU tropopause does not coincide with a unique set of humidity and ozone thresholds, it is impossible to remove this uncertainty here. In addition, the summer season of this campaign is associated with
10 enhanced in-situ formation of ozone in the planetary boundary layer, which could be transported upward, although in this case one would expect the ozone-rich air to show higher humidity.

With these caveats, Fig. 1 suggests that stratospheric intrusions reaching the lower troposphere were present above Montreal for the full length of the campaign, except
15 on 21 July and 27–28 July. The entire period can be separated into three large intrusion events, each lasting from four to seven days: (1) from 14 July to 20 July; (2) from 22 July to 26 July; and (3) from 29 July to 4 August. Based on the RH profile alone (the ozone profile is missing), the second multi-day event should be extended to 27 July, since dry air was found on that day down to 800 hPa too. RH and ozone
20 profiles at Egbert and Walsingham (not shown) display similar structures, suggesting a nearly continuous presence of stratospheric intrusions reaching the lower troposphere over eastern Canada during this period. The relative proximity of the three stations would suggest that some degree of correspondance should exist between them. But the larger number of missing profiles in Egbert and Walsingham makes it difficult to
25 distinguish multi-day intrusion events or a clear correspondance with the three large intrusion events seen at Montreal.

3.2 Objective identification of stratospheric intrusions

In order to objectively compare observations of ozone and humidity to the Lagrangian STE data set, it is necessary to set up a numerical algorithm to identify stratospheric intrusions in the measured profiles. The algorithm built for the purpose of this evalu-
⁵ ation uses the same criteria as a subjective determination of intrusion levels (such as that above), and aims for similar results. The results from this objective identification algorithm depend upon the values of the parameters used, and these may not be appropriate for other seasons and locations. These results are also naturally subject to the uncertainties discussed above that are intrinsic to identifying stratospheric intru-
¹⁰ sions from balloon sonde profiles. These are difficult to quantify. Nevertheless, use of this algorithm ensures that identical criteria are used for all profiles, thus providing consistency and reproducibility of results.

The algorithm screens an observed profile from the ground upwards by increments of 10 hPa, and identifies alternating bottoms and tops of intrusions using the following
¹⁵ criteria applied to a centered smoothed window of 50 hPa width. Fields are smoothed within this window using a seventh-order polynomial fit. All three criteria must be met for an intrusion bottom or top to be detected, and multiple intrusions layers may be identified in a single profile. Vertical gradients are expressed in pressure coordinates.
Detection of the bottom of the intrusion:

²⁰ (i) Positive vertical gradient in RH $\frac{d\text{RH}}{dP} > 1\,\%\,\text{hPa}^{-1}$;

(ii) RH < 60 % in layer above;

(iii) Negative vertical gradient in ozone $-\frac{d\text{O}_3}{dP} > 0.1\,\text{ppbv}\,\text{hPa}^{-1}$.

Detection of the top of the intrusion:

(i) Negative vertical gradient in RH $-\frac{d\text{RH}}{dP} > 1\,\%\,\text{hPa}^{-1}$;

²⁵ (ii) RH > 40 % in layer above;

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

(iii) Positive vertical gradient in ozone $\frac{dO_3}{dP} > 0.1\,\text{ppbv}\,\text{hPa}^{-1}$.

The threshold RH value of 60 % within a stratospheric intrusion allows inclusion of regions of intrusions that have undergone significant mixing with the surrounding air, consistent with the Lagrangian STE data set, which ignores subgrid-scale mixing. Af-
5 ter this first-pass identification of any intrusion layers, the algorithm next removes any thin layers by either ignoring intrusions that are thinner than 50 hPa or by merging intrusions that are closer to each other than 50 hPa. Finally, and except in Fig. 1 where top and bottom levels of intrusions are shown on a 10 hPa vertical grid, the top and bottom levels of intrusions are moved to the closest grid point of a 50 hPa-resolution
10 vertical grid starting at 1000 hPa. This simplifies the analysis and the comparison with the Lagrangian STE data set. The resulting stratospheric intrusions are only sensitive to the details of the choice of parameters in regions where the stratospheric signature is ambiguous. In such cases, the parameters are chosen such that most ambiguous cases are included. The presence of ambiguous cases constitutes a source of un-
15 certainty in the comparison with the Lagrangian STE data set (see Sect. 5). However, these ambiguous layers are usually located at either the bottom or the top of intrusions, and therefore it will be useful to distinguish between the inner part of intrusions and the layers surrounding their bottom and top levels.

Intrusions identified with this algorithm for Montreal balloon-sonde launches are
20 shown as a black hatching overlaid on each profile in Fig. 1. The hatching nicely separates the troposphere between two regions with markedly different relative humidity and ozone characteristics. The air below identified intrusions is markedly moister and typically, but not always, has less ozone. The hatched region coincides well with the zone identified above of dry, ozone-rich air whose bottom undulates over time. Intru-
25 sions penetrate to the middle troposphere around 600 hPa every day of the campaign except for July 21 and 28, when moist air is clearly present at higher altitudes, probably in relationship with clouds in the region. A few isolated layers that were not identified as stratospheric intrusions can also be noted surrounded by hatched areas (e.g., 16, 18, 20, and 24 July; 2 August). Most of these show clearly higher RH than in the

Title Page

Abstract   Introduction

Conclusions   References

Tables   Figures

◄◄   ►►

◄   ►

Back   Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

surrounding intrusions. Note that some layers with relatively moist air were identified as stratospheric intrusions (e.g., 13 July around 350 hPa, 31 July around 300 hPa, 4 August around 700 hPa). These cases are ambiguous, and as mentioned above, the algorithm parameters are chosen such that ambiguous cases are typically included in intrusions.

Table 1 summarizes the number of profiles where a stratospheric intrusion was detected somewhere below 300, 500 and 700 hPa for the three stations. Intrusions are detected below 300 hPa in 89 % of all profiles, below 500 hPa in 79 %, and below 700 hPa in 38 %. The Montreal station has an even higher number of intrusions at all levels, with an intrusion reaching below 700 hPa every other day. Walsingham has the lowest number of intrusions, with still one day in four showing an intrusion reaching below 700 hPa. This frequency of deep intrusions is remarkably high, though bearing in mind that this is a relatively short period. Note however from Fig. 1 that no intrusion has been detected below 800 hPa, where boundary-layer processes intervene. This also applies to the other two stations. Yet, it can be estimated from the total black shaded area below 700 hPa in Fig. 1 that around ten percent of the air below this level originates in the stratosphere, a number which is about two orders of magnitude larger than estimates from Sprenger and Wernli (2003).

3.3 Distributions of humidity and ozone in observed stratospheric intrusions

In order to characterise the composition of the stratospheric intrusions that were detected and to verify the appropriateness of the objective identification algorithm, it is useful to analyse the distributions of humidity and ozone with respect to the identified intrusions. As mentioned above, the layers near the bottom and top levels of intrusions are in some cases ambiguous. We distinguish between the different regions of a tropospheric column relatively to the intrusion by tagging each 50 hPa layer of a vertical profile with one of the following five categories with decreasing priority:

1. **Intermediate (Bottom):** The two 50 hPa layers on either side of the bottom of the intrusion.

2. **Intermediate (Top):** The two 50 hPa layers on either side of the top of the intrusion, if they have not already been assigned to the "Intermediate (Bottom)" category.

3. **Inside Intrusion:** The 50 hPa layers inside an intrusion that have not been assigned to one of the "Intermediate" categories.

4. **Above Intrusion:** The 50 hPa layers that lie above an intrusion and that have not been assigned to one of the "Intermediate" categories.

5. **Below Intrusion:** The 50 hPa layers that lie below an intrusion and that have not been assigned either to one of the "Intermediate" categories or to the "Above Intrusion" category.

An example of this categorisation is given in Fig. 2 for 24 July in Montreal.

Distributions of observations of RH (%), specific humidity (g kg^{-1}), and ozone (ppbv) are shown in Fig. 3 with respect to these intrusion categories for all three stations combined. Note that the corresponding distributions for individual stations are very similar (not shown). The height of the bars gives the percentage of the 50 hPa layers belonging to a given category. Consistent with the high frequency of intrusions mentioned before, the category "Inside Intrusion" dominates down to 700 hPa but becomes almost absent below this level. This is due to the fact that the bottom of intrusions never penetrated below 800 hPa during the entire campaign. RH (top row) inside stratospheric intrusions is very dry at all pressure levels: more than 60 % of the air has a RH lower than 20 %, and less than 20 % has RH larger than 40 %. The RH distribution is clearly shifted towards higher values in the "'Intermediate" categories, and even more so in the "Below Intrusion" and "Above Intrusion" categories. Interestingly, the RH distributions, in particular for the "Inside Intrusion" category, stay relatively constant throughout the four pressure ranges. This may be due to a combination of low specific humidity, indicating

memory of a stratospheric origin, and temperature. In order to disentangle these two contributions, the middle row of Fig. 3 shows the specific humidity distributions. The categories keep their relative differences, but the specific humidity clearly increases with increasing pressure in all categories, including the "Inside Intrusions" category. This shows the important role of the vertical profile of temperature in keeping the low RH in stratospheric intrusions throughout the troposphere. Assuming that air inside intrusions comes from the stratosphere, a likely cause for the increase in specific humidity inside intrusions with increasing pressure is that significant mixing takes place with the surrounding air during the descent. Another possible cause may be that the air identified here as from stratospheric origin has its actual origin in the upper troposphere for lower altitude intrusion layers. The overall pattern in ozone distributions is consistently similar to the specific humidity (but inverted), with larger ozone molar mixing ratios at lower pressures and in the "Inside Intrusion" category. This similarity highlights the passive nature of these two tracers.

4 Evaluation of the Lagrangian STE data set

4.1 Methodology

The finite spatial (55 km, 5 hPa) and temporal (24 h) intervals separating the trajectories at initial time allow for a mapping of STE mass fluxes with a spatio-temporal grid of typically coarser resolution (e.g. $2° × 2° × 50$ hPa $× 24$ h). In the case of this evaluation, we need to compare the Lagrangian STE data with instant, point-observed profiles. This requires us to artificially define an area and time-interval around the observation location and time. The presence of a trajectory within this area and time-interval is then assumed to represent the presence of air of stratospheric origin above the given station, and is compared to the observed profile at the station. Here, we have chosen this area to be within a radius of 55 km around the station, and the time-interval to be ±6 h around measurement time. The vertical grid was set with 50 hPa

M. S. Bourqui et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

spacings from 1000 hPa. The radius was chosen to allow for some stretching of the STE parcel. For the time-interval, we have verified that trajectories cross the 2 PVU tropopause uniformly within the [12 h, 36 h[window (not shown), such that taking a shorter time-interval does not lead to a statistical bias. We have chosen a ±6 h inter-

5 val as a compromise between including events that were not observed, but that would have been observed at a different time, and allowing for the sparseness of the initial intervals between trajectories. Since more than one trajectory can be located in each 50 hPa bin above a station (i.e. within a 55 km radius of the site and within a ±6 h interval), we consider that a 50 hPa bin with one or more trajectories has a stratospheric

10 origin, while the absence of trajectories represents a tropospheric origin.

It is important to note that this finite radius and time-interval unavoidably leads to mismatches between STE data and observations, and that this error disappears in geographical density maps of STE (not shown). We have tested the sensitivity of the results to these parameters by varying the radius between 27.5 and 555 km and the

15 time-interval between ±1 and ±12 h, but the results could not be improved. When increasing the size of the spatial and/or temporal window, more trajectories were detected in each vertical 50 hPa bin, but the missed events (i.e., no trajectories found in a layer where a stratospheric intrusion was observed) were not reduced significantly. Similarly, when reducing the size of the spatial and/or temporal window, fewer trajec-

20 tories were detected in each vertical bin, but the overforecasts or false alarms (i.e., trajectories present in a layer where no stratospheric intrusion was observed) were not significantly reduced. This relative insensitivity most likely comes from the fact that intrusions translate to clusters of trajectories, and an increase in the analysis window size essentially leads to the inclusion of more trajectories from the same cluster.

25 The skill of the Lagrangian STE forecast at capturing the stratospheric intrusions identified in observations is quantified using 2 × 2 contingency tables and standard categorical scores (Jolliffe and Stephenson, 2003). A 2 × 2 contingency table simply provides the number of profiles with/without stratospheric intrusions observed and/or forecasts (see illustration in Table 2). The diagonal of the table represents the correct

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

forecasts (with/without intrusions), while the off-diagonal elements reflect the overforecasts or false alarms (top right) and the missed events (bottom left). The four categorical scores used in this study are described below. Variables a, b, c, d and n are defined in Table 2.

PC: Proportion Correct $= (a + d)/n$ (1 = perfect). This score represents the proportion of correct forecasts, whether with or without intrusions. A sequence of perfect guesses would result in a PC = 1.0. An infinite sequence of random guesses would result in a PC = 0.5. A finite sequence of random guesses would result in a PC belonging to a binomial probability distribution centered in 0.5, whose width depends on the size of the sequence. Hereafter, we consider that the PC is statistically significant if it is not within the central 90 % of the binomial probability distribution representing random guesses. In other words, a statistically significant PC has a probability of representing a sequence of random guesses that is lower than 10 %. In general, the PC score artificially increases in cases of low event frequency.

FB: Frequency Bias $= (a + b)/(a + c)$ (1 = perfect). This score measures the bias in the forecast. It is equal to 1 for an unbiased forecast, larger than 1 for an overforecast, and smaller than 1 for an underforecast.

HR: Hit Rate $= a/(a + c)$ (1 = perfect). This score measures the ratio of the number of correct forecast events to the total number of observed events. It thus quantifies the proportion of missed events. It is equal to one when there are no missed events, and 0 when all events have been missed. Since it only counts the correct forecast events, it may be artificially improved when there is a tendency for overforecasting.

FAR: False Alarm Ratio $= b/(a + b)$ (0 = perfect). This score measures the ratio of the number of overforecast or false alarm events to the total number of forecast events. It thus quantifies how trustable the forecast events are. It is equal to 0 when no false alarm are forecast and 1 when all forecasts are false alarms.

4.2 Comparison of stratospheric intrusions in the STE data set and in observations at the three stations

The number of trajectories of stratospheric origin (i.e. crossing the 2 PVU tropopause in the S→T direction) found in the data set that approached one of the three balloon-sonde sites within a radius of 55 km, and within ±6 h of the measurement time, is shown in Fig. 4 for each day for the three stations, with the stratospheric intrusions identified from observed profiles overlaid as black hatching. At Montreal (Fig. 4, top row), the first two multi-day intrusion events seen in observations, namely 14–20 July and 22–26 July, coincide well with distinct multi-day intrusion events in the forecasts. During these two periods, differences occur mainly below 600 hPa in both directions (i.e., both missed or overforecast events), but with a tendency for the forecasts to miss events. At Egbert and Walsingham, the general structures of the intrusions seem also well captured, but again with a few events missed, such as 14 and 27 July at Egbert and 27 July at Walsingham. Other mismatches mostly concern pressures larger than 600 hPa, as at Montreal. The third multi-day intrusion event identified in observations at the three stations (29 July–4 August) is also found in the STE forecasts. Interestingly, the forecasts of this event are very similar for the three stations, with deep intrusions found in the beginning and end of the period, and shallower intrusions confined above 450 hPa found inbetween. Observations suggest that this event was deeper in its core above Montreal than above the other stations. The forecast therefore captures accurately the beginning and end of the event at all three stations, but misses the deep part of its core above Montreal. This may be either due to a spatial shift of the whole event in the forecasts, or to a misrepresentation of a deeper part of its core which was only observed at Montreal. Note that the thermal tropopause (black solid lines) estimated from observations and the dynamical tropopause as shown by the top of the blue shaded columns in Fig. 4 agree well at Montreal, but show a few days of mismatch at Egbert and Walsingham.

4.3 Skill at representing the overall depth of intrusions

Table 3 provides the contingency tables and categorical scores for the skill of the STE forecast at capturing events, defined as the presence of a stratospheric intrusion penetrating below 300, 500, and 700 hPa. These figures can be linked to the number of intrusions identified from observations given in Table 1. Note that in Table 3, forecast and observed intrusions are considered as matching as long as they both include an intrusion layer located somewhere below the pressure threshold. This is an allowance for uncertainties associated with the relative sparseness of the trajectory's initial positions. In the upper troposphere, the STE data set tends to overforecast intrusions penetrating below 300 hPa, such that all profiles but one are forecast with an intrusion, including 5 non-observed events. This translates into a small positive bias varying between 5 and 14 % at individual stations and estimated to 10 % over the full campaign (Table 3, left column). These overforecasts translate into a proportion of correct forecasts estimated to 91 %. Consistently, the hit rate is equal to 100 %, while the false alarm ratio is equal to 9 %. In the mid-troposphere, the STE data set both misses and overforecasts a few intrusions penetrating below 500 hPa, resulting in a slight underestimate of the frequency of intrusions of 11 %. The proportion of correct forecasts is reduced to 77 %, with 20 % missed events, partially compensated by 10 % false alarms. For deep intrusions penetrating below 700 hPa, the proportion of correct forecasts is not statistically significant, i.e. the forecasts have no predictive skill at all. The STE data miss 66 % of events, which are partially compensated by 56 % false alarms, resulting in a small low bias of 24 % (FB = 0.76). This low bias shows however a large variability among the three stations (FB = 0.54, 0.75 and 1.14), owing to the overall small number of events. The STE data is therefore only useful in a statistical sense for intrusions penetrating lower than 700 hPa.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

4.4   Skill at representing the detailed structure of intrusions

Table 4 provides the contingency tables and categorical scores for the skill of the STE data at capturing the detailed structure of the intrusions. Here an event is defined as the presence of air of stratospheric origin within a 50 hPa vertical bin, and the analysis is performed separately for four ranges of pressures. Above 300 hPa, the predictive skill is excellent, with a slight tendency for false alarms and related positive bias, consistent with the discussion on Table 3. In the pressure range 300 to 500 hPa, the predictive skill degrades slightly, with increased missed events and false alarms. However these almost compensate, and result in a very small positive bias. Hence, down to 500 hPa, consistent with the analysis of Table 3, the STE data has both very good predictive skill and a small statistical bias. In the pressure range from 500 to 700 hPa, the proportion correct drops to 0.58, reflecting limited predictive skill that is however still statistically significant. The forecasts miss about half of the events, and are composed of 28 % false alarms. This results in a low bias estimated at 31 % (FB = 0.69). It is interesting to note that this degradation is dampened when considering skill at representing overall intrusion depth of 500 hPa (Table 3), due to a partial correction by the layers below 700 hPa. This suggests a greater difficulty in capturing the details of these intrusions in the 500–900 hPa region than their overall presence. In the pressure range between 700 and 900 hPa, the probability of observing an event drops to 15 %, resulting in only 30 events through the entire campaign, all located above 800 hPa. The probability of forecasting an event is very realistic at 13 %, and includes events down to 900 hPa. The STE forecasts miss 77 % of events, that are well compensated by 73 % false alarms, yielding a slight positive bias of 13 % (FB = 87). But despite this large number of missed events and false alarms, the proportion correct remains high (PC = 0.81). This is artificial, and due to the large number of non-observed, non-forecast events, reflecting the statistical predominance of no-events below 700 hPa rather than the forecast's predictive skill. Hence, the STE data shows a very good statistical representation of events below 700 hPa, but does not provide useful predictive skill.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

"Evaluation of new
STE diagnosis
system

M. S. Bourqui et al.

In an attempt to see where missed events and false alarms are located with respect to the observed events, Fig. 5 shows the percentages of 50 hPa layers with or without trajectories of stratospheric origin in the same five categories as discussed for Fig. 3. Note that the "Intermediate (Top)" and "Intermediate (Bottom)" categories are ambiguous in the observations and will be ignored in the evaluation of the STE data made here. In the region above 300 hPa, the STE forecasts miss about 15 % of the 50 hPa layers belonging to the "Inside Intrusion" category. This is due to the mismatch between the thermal tropopause and the 2 PVU tropopause. The categories "Above Intrusion" and "Below Intrusion" are quite systematically filled with overforecasts, consistent with the discussion on Table 4. This error is also found in the next range of pressures from 300 to 500 hPa, with systematic overforecasts in the category "Above Intrusion". This means that the STE data set systematically misses intermediate layers of tropospheric air embedded between stratospheric intrusion layers. In this pressure range, the STE data set represents the "Inside Intrusion" well, with about 15 % missed events, consistent with the discussion on Table 4. It still fills the "Below Intrusion" category with about 50 % overforecasts. In the region between 500 and 700 hPa, the forecast again fills the "Above Intrusion" category with about 50 % overforecasts. It represents the "Inside Intrusion" category less well, where it misses about a half of the intrusion layers, again consistent with Table 4. It represents the "Below Intrusion" category well, with about 20 % overforecasts. Finally, in the region below 700 hPa, about half of the "Inside Intrusion" layers are missed by the forecast, but this represents only a very small number of events. The only significant category in this pressure range is the "Below Intrusion", which is well captured with only about 5 % overforecasts. Overall, this shows that most overforecasts above 500 hPa take place within tropospheric layers that lie between stratospheric intrusions, and that the hit rate scores discussed in Table 4 almost exactly apply to the inner part of intrusions. This suggests that ambiguous layers, likely isolated in the "Intermediate (Top)" and "Intermediate (Bottom)" categories, do not account for the missed events in the forecasts.

5 Analysis of STE forecast errors

The previous section showed a smooth degradation, from the upper to the lower tro-posphere, of the predictive skill of the STE data for the 50 hPa intrusion layers. The predictive skill was found excellent in the upper troposphere, and good in the middle troposphere, but was deemed negligible below 700 hPa. The statistical bias was found to be slightly positive in the upper troposphere, and negative below 500 hPa. The re-gion 500–700 hPa had the strongest bias, with a frequency bias (FB) equal to 0.69 for the full campaign. This low bias was also found in the analysis of the STE data skill at representing overall intrusion depths. The predictive skill is not only sensitive to errors in the STE data, but also to uncertainties in the method of comparison with the observations. It is likely that the latter will have a lesser effect on the statistical bias of the STE data, due to the averaging effect. In this section, we analyse the five possible causes for this underestimate of stratospheric intrusions in the middle to lower troposphere in the Montreal area. This station has statistics that are well representative of the overall statistics for the full campaign, and has the largest number of observed balloon-sonde profiles.

1. As noted in Sects. 3.1 and 3.2, the identification of intrusions in the observed profiles is subject to uncertainties related to (i) the upper-troposphere acting as a source of dry, ozone-rich air, and not only the region above 2 PVU; (ii) the pres-ence of ambiguous layers in observed profiles; and (iii) the additional errors due to in-situ ozone formation. These uncertainties are difficult to estimate in this study, and are expected to contaminate the evaluation of the STE data made here. Using the objective algorithm developed for the purpose of this study, a better understanding of these uncertainties could be gained in a future balloon campaign by calculating additional backward trajectories started from the nodes of the measured profiles. Additional chemical species data (e.g. from aircraft measurements) would also be very helpful. However, this source of error does not affect the real skill of the STE forecasts.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

2. The relative spatio-temporal sparseness of the trajectory's starting grid affects the accuracy of the evaluation made here, as discussed in Sect. 4.1. It obviously limits the evaluated predictive skill of the STE data. It may also, however, introduce an artificial bias when a set of trajectories representing a stratospheric intrusion becomes too dispersed to be captured within the 55 km radius of a station. These missed events can be detected by screening the clusters of trajectories that descend into the middle or lower troposphere. Note that these clusters are part of the routine calculations of the Lagrangian STE data set. The next subsection offers a detailed screening of candidate clusters for the missed events and shows that too-dispersed trajectories do significantly contribute to the missed events. Note that this error source does not apply to climatological density maps, for which incidences are summed over spatio-temporal intervals.

3. The temporal limit of 6 days for the trajectory length may lead to missed events, and a statistical bias, since all older stratospheric air will be missing in the STE data. Figure 6 shows the distribution of the age of trajectories as they reach one of the three stations within the "Inside Intrusion" category. The age is determined as the difference between the time of crossing the 2 PVU and the average time while passing through the 55 km radius disk around the station. Since trajectories cross the 2 PVU tropopause approximately uniformly within the time window [12 h, 36 h[(not shown), and have a total length of 144 h, their maximum possible age is uniformly distributed between 108 h and 132 h (grey shaded zone in Fig. 6). The shapes of the distributions, although likely under-sampled, suggest a maximum frequency of the age above 300 hPa at around 48 h, between 300 and 500 hPa at around 48 h–60 h, and between 500 and 700 hPa at around 84 h. Below 700 hPa, only three events were found and no inference can be made. It is interesting to see that all distributions (except the one for below 700 hPa) show a decrease in frequency towards the age of 108 h, i.e. before that any limitation due to the trajectory length arises. This suggests that the trajectory length of 6 days allows for the highest frequency ages. Nevertheless, the layers 500–700 hPa and below

700 hPa are likely to be affected by the age cut-off. In an attempt to evaluate further this cause of error, the next subsection discusses candidate clusters that, if temporally extended, may account for missed events.

4. Another type of error is related to the divergence of the weather forecasts with respect to the true atmospheric state, which is likely to increase with forecast length. Such error can be expressed as geographical shifts in stratospheric intrusions, as misrepresentations in their vertical location and extent, or simply as an absence of stratospheric intrusions in the meteorological forecasts. They may translate into climatological biases. Furthermore, such a Lagrangian measure of meteorological forecast errors, which focuses on non-conservative flows that link the lower stratosphere to the lower troposphere, offers an interesting new perspective on forecast skill. This perspective is useful for understanding the accuracy of regional and global air-quality models and also the dynamical interaction between the upper and lower troposphere. Some insights into this source of error will be drawn in the next subsection by identifying candidate clusters that could have accounted for missed events if spatially shifted.

5. The last possible cause of error is related to the absence of subgrid scale parameterisations such as convection and turbulent diffusion in the kinematic trajectories. Bourqui and Trépanier (2010) suggested that descent from the stratosphere is typically cloud-free and quasi-adiabatic until a parcel reaches the lower troposphere, where clouds and boundary-layer processes intervene. It is therefore expected that such errors most likely contaminate the STE forecasts in the lower troposphere, typically below 700 hPa. Moreover, since subgrid scale processes affect grid scale winds in the weather model, it is expected to see some indirect effect of subgrid scale convection or turbulence on the trajectories, such as sudden jumps or disorganised motions. It is difficult to estimate implications on the predictive skill resulting from this source of error. However, it is likely that these errors cancel on a climatological average.

5.1 Cluster analysis of missed events

Here, we first identify clusters of trajectories that descend from the tropopause down to 700 hPa. This clustering is discussed in Sect. 2 and only rejects about 10 % of all trajectories reaching the 700 hPa level. The clusters whose centers of mass approach the station within a radius of 1,111 km (distance equivalent to 10° latitude) within ±6 h of the measurement time are then selected, and the trajectories contained in each cluster are plotted on individual maps (see Fig. 7 for a sample). Additionally, the clusters that enter within the same radius around the station within their last 12 h are also selected and plotted on individual maps (see Fig. 7 for a sub-sample). The first are analysed to identify possible clusters that were potentially shifted with respect to the station, or clusters that could not be seen within the 55 km radius around the station because their trajectories were too dispersed. The second are analysed to identify clusters that would have reached the station, were they calculated for a slightly longer period of time. Table 5 gives the list of missed intrusions during the campaign in Montreal, with the pressure range of the missed part of the intrusion. The right column gives the clusters, identified by the number of member trajectories, that would be candidates for the missed events if they were slightly shifted geographically (sh), if the density of trajectories were higher (de) or if they had been slightly extended in time (ex). Note that the same cluster may be a candidate for different events. Table 5 shows that all missed events, except the 22 July one, have at least one candidate cluster, and that in most cases, several distinct candidates clusters exist. In about 70 % of all missed events, there is at least one candidate cluster that passes over the station but is missed due to its low trajectory density. This explanation reduces the number of missed events by a factor of three at least, with the remaining missed events (except for one case) being explainable by a shift in the cluster due either to the spatio-temporal sparseness of trajectory starting grid leading to a mismatch between trajectory clusters and measurements, or to a geographical shift in the meteorological fields.

Figure 7 provides examples of typical candidate clusters that could account for three missed events: First row for 30 July , second and third rows for 15 July , fourth and fifth rows for 16 July . The candidate cluster for the 30 July event (first row) clearly crosses Montreal in the beginning of the ±6 h window around the measurement time, but because of significant dispersion, the trajectory density was too low and no trajectory was detected within the 55 km radius around the station. The cluster spans pressures between 700 and 800 hPa when it traverses the Montreal column, which coincides well with the bottom of the missed intrusion layer (500–750 hPa, see Table 5). In the second example (Fig. 7, second row), a large cluster passes at a few hundreds of kilometers east of the station around the 15 July measurement time. Its pressure range (500-750 hPa) coincides well with the observed intrusion (600–750 hPa). Another cluster (third row) that started three days earlier approaches Montreal while the trajectories disperse and end just above it. Although the cluster spans a large range of pressures at the time it arrives above the station (300–900 hPa), the particular trajectories that enter within the 55 km radius around Montreal are located above 600 hPa. These two candidate clusters illustrate well the complex time-dependent Lagrangian structures of certain stratospheric intrusions. The fourth and fifth rows in Fig. 7 are candidates for the missed event on 16 July , the fourth cluster with a southeastward shift and the fifth cluster with an extension of the trajectories. Both clusters together span the pressure range of the missed intrusion layer (600–750 hPa).

As shown in Table 5, several clusters that fall within the time and pressure ranges of the event but that are either shifted or too dispersed to be detected within the 55 km radius around the station can account for a missed event. In view of this, it seems tempting to increase the size of this radius. However, as discussed above, doing so would add more overforecasts and would not reduce significantly the number of missed events.

6 Conclusions

We have developed a new real-time Lagrangian diagnostic system for stratosphere-troposphere exchange. The Lagrangian STE data set produced by this new system offers daily, high-resolution global data on both the downward (S→T) and upward (T→S) transport of mass across the dynamical ±2 PVU tropopause and the 380 K isentrope. The data set is calculated using the methodology proposed in Bourqui (2006) using global meteorological forecasts from Environment Canada's operational Global Environmental Multiscale (GEM) numerical weather prediction model. Routine calculations with the new STE diagnostic system began in 20 July 2010 and have now delivered more than one year of data. Archived data include trajectories, daily maps of STE mass flux densities, diagnosis of deep intrusions, and details on clusters of trajectories. Once its errors are evaluated and characterized, this data set will be useful for improving our basic understanding of STE in the northern and southern hemispheres as well as in the tropics. It will also offer a reference data set for evaluating the cross-tropopause transport predicted by various chemistry-transport models and climate-chemistry models.

The present study is the first evaluation of this new data set, using balloon-sonde measurements from a field campaign in eastern Canada in summer 2010. During the campaign, balloon sondes were launched daily from three sites (Montreal, Egbert, and Walsingham). Measurements of ozone, relative humidity, temperature, pressure, and wind speed and direction were made through the troposphere and lower stratosphere. Analysis of the observational data suggested an unexpectedly high frequency of (deep) stratospheric intrusions during the campaign period. Over the three stations, 89 %, 79 % and 38 % of the days showed the signature of stratospheric air below 300 hPa, 500 hPa, and 700 hPa, respectively, though with significant variability among the three stations. In Montreal, every second measured profile showed clear signatures of stratospheric air below 700 hPa.

"Evaluation of new STE diagnosis system

M. S. Bourqui et al.

An objective algorithm was developed to identify stratospheric intrusions in the measured profiles in order to help the comparison with the STE data set. This objective algorithm also allowed characterization of the observations with respect to the intrusions. Distributions of RH within these different categories showed clear differences between the Inside Intrusion and the Above or Below Intrusion categories, but the distributions remained similar through the troposphere from the ground to above 300 hPa. This similarity with height is shown to be due to the effect of the increase of temperature with increasing pressure, which counterbalances the increase in specific humidity. The latter clearly increases at lower altitudes but is much dryer inside intrusions than outside. The ozone molar mixing ratio is found to behave very similarly to the specific humidity, but in an inverted fashion. It should be noted however, that an objective algorithm based only on individual profiles has inherent flaws. In particular, it is not able to distinguish descents of dry upper tropospheric air from descents of air from above the 2 PVU tropopause. Summertime low-level in-situ ozone production and upward vertical transport may add further errors. Hence, this increase in specific humidity and decrease in ozone mixing ratios with increasing pressure inside intrusions suggests that there is significant mixing taking place during the descent of stratospheric air. Evaluation of the STE data set at representing the overall depth of stratospheric intrusions identified in the observed profiles shows excellent predictive skill for intrusions penetrating below 300 hPa and 500 hPa, respectively. The bias is estimated to be smaller or equal to 10 % in both cases. The predictive skill vanishes for intrusions penetrating below 700 hPa, while the statistical skill remains reasonable with an estimated bias of 24 %. Evaluation of the STE data set at capturing the detailed structure of stratospheric intrusions shows good predictive skill above 500 hPa, with an estimated bias smaller than 15 %. The predictive skill degrades in the layer 500–700 hPa, with only 58 % correct forecasts, and an associated 31 % estimated low bias. The bias improves in the layer below 700 hPa while the predictive skill vanishes.

Five possible causes of mismatches between the Lagrangian STE data set and the stratospheric intrusions identified from observed profiles were discussed:

1. Uncertainties in the identification of stratospheric intrusions in observed profiles.

2. Relative spatio-temporal sparseness of trajectory starting grid.

3. The temporal limit of 6 days on trajectories.

4. Errors in the weather forecasts.

5. Lack of subgrid-scale processes in the trajectory calculations.

It is impossible to quantify accurately the separate contributions of these errors in this study, since the identification of stratospheric intrusions from observations is itself uncertain. Nevertheless, a qualitative understanding can be gained by screening individual missed intrusions for candidate clusters that would account for the missed event, were it to be slightly shifted in space, extended in time, or if the density of trajectories were to be increased. For all missed events except one, we found candidate clusters of trajectories approaching Montreal in the correct range of pressures. About 70 % of the missed events could be explained by a cluster of trajectories that dispersed such that no trajectory was detected within the 55 km radius around Montreal but several trajectories were found on either side of this station. This reduces the bias found in the layer 500–700 hPa by a factor three at least. Examples were given of candidate clusters for three missed intrusions.

From the five causes of errors listed above, only two (3, and 4) are anticipated to affect a climatology based on this STE data set. Trajectory age distributions furthermore suggest that error source 3 only affects the layers below 500 hPa. With a maximum bias of about 30 % in the layer between 500 and 700 hPa, composed of missed events up to 70 % of which can be attributed to the low density of trajectories in dispersed intrusions, the STE data set can be anticipated to have a reasonably small climatological bias.

This study represents the first evaluation of this new Lagrangian STE data set. It is however limited to eastern Canada in one summer season, and therefore its results may depend to some extent on the variation with season and location of the accuracy

"Evaluation of new STE diagnosis system

M. S. Bourqui et al.

of the driving meteorological forecasts. Further evaluations in different seasons and locations around the world will be useful in order to characterise its errors in different parts of the world.

References

Appenzeller, C., Holton, J. R., and Rosenlof, K. H.: Seasonal variation of mass transport across the tropopause, J. Geophys. Res., 101, 15071–15078, 1996. 27969

Bourqui, M. S.: Stratosphere-troposphere exchange from the Lagrangian perspective: a case study and method sensitivities, Atmos. Chem. Phys., 6, 2651–2670, doi:10.5194/acp-6-2651-2006, 2006. 27968, 27969, 27970, 27972, 27973, 27993

Bourqui, M. S. and Trépanier, P.-Y.: Descent of deep stratospheric intrusions during the IONS August 2006 campaign, J. Geophys. Res., 115, D18301, doi:10.1029/2009JD013183, 2010. 27970, 27974, 27990

Cooper, O., Stohl, A., Hubler, G., Hsie, E., Parrish, D., Tuck, A., Kiladis, G., Oltmans, S., Johnson, B., Shapiro, M., Moody, J., and Lefohn, A.: Direct transport of midlatitude stratospheric ozone into the lower troposphere and marine boundary layer of the tropical Pacific Ocean, J. Geophys. Res., 110, D23310, doi:10.1029/2005JD005783, 2005. 27970

Côté, J., Desmarais, J., Gravel, S., Methot, A., Patoine, A., Roch, M., and Staniforth, A.: The operational CMC-MRB Global Environmental Multiscale (GEM) model. Part II: Results, Mon. Wea. Rev., 126, 1397–1418, 1998a. 27973

Côté, J., Gravel, S., Méthot, A., Patoine, A., Roch, M., and Staniforth, A.: The Operational CMCMRB Global Environmental Multiscale (GEM) Model. Part I: Design considerations and formulation, Mon. Wea. Rev., 126, 1373–1395, 1998b. 27973

Gerasopoulos, E., Zanis, P., Papastefanou, C., Zerefos, C. S., Ioannidou, A., and Wernli, H.: A complex case study of down to the surface intrusions of persistent

stratospheric air over the Eastern Mediterranean, Atmos. Environ., 40, 4113–4125, doi:10.1016/j.atmosenv.2006.03.022, 2006. 27970

Gettelman, A. and Sobel, A. H.: Direct diagnoses of Stratosphere-Troposphere Exchange, J. Atmos. Sci., 57, 3–16, 2000. 27970

Gray, S. L.: A case study of stratosphere to troposphere transport: The role of convective transport and the sensitivity to model resolution, J. Geophys. Res., 108, 4590, doi:10.1029/2002JD003,317, 2003. 27969

Hall, T. M. and Holzer, M.: Advective-diffusive mass flux and implications for stratosphere-troposphere exchange, Geophys. Res. Lett., 30, 1222, doi:10.1029/2002GL016,419, 2003. 27969, 27970, 27974

Hegglin, M. I. and Shepherd, T. G.: Large climate-induced changes in ultraviolet index and stratosphere-to-troposphere ozone flux, Nat. Geosci., 2, 687–691, 2009. 27971

Hoerling, P. M., Schaack, T. K., and Lenzen, A. J.: A global analysis of stratospheric-tropospheric exchange during northern winter, Mon. Wea. Rev., 121, 162–172, 1993. 27970

Holton, J. R., Haynes, P. H., McIntyre, M. E., Douglass, A. R., Rood, R. B., and Pfister, L.: Stratosphere-Troposphere Exchange, Rev. Geophys., 33, 403–439, 1995. 27969

Jolliffe, I. T. and Stephenson, D. B.: in: Forecast Verification: A Practitioner's Guide, Atmospheric Science, Wiley and Sons, Chichester, 2003. 27982

Komhyr, W. D.: Electrochemical concentration cells for gas analysis, Ann. Geophysicae, 25, 203–210, 1969. 27975

Langford, A. O., Masters, C. D., Proffitt, M. H., Hsie, E.-Y., and Tuck, A. E.: Ozone measurements in a tropopause fold associated with a cut-off low system, Geophys. Res. Letters, 23, 2501–2504, 1996. 27969

Lefohn, A. S., Wernli, H., Shadwick, D., Limbach, S., Oltmans, S. J., and Shapiro, M.: The importance of stratospheric-tropospheric transport in affecting surface ozone concentrations in the western and northern tier of the United States, Atmos. Environ., 45, 4845–4857, doi:10.1016/j.atmosenv.2011.06.014, 2011. 27970

Mullendore, G. L., Durran, D. R., and Holton, J. R.: Cross-tropopause tracer transport in midlatitude convection, J. Geophys. Res., 110, doi:10.1029/2004JD005,059, 2005. 27969

Oltmans, S., Lefohn, A., Harris, J., Galbally, I., Scheel, H., Bodeker, G., Brunke, E., Claude, H., Tarasick, D., Johnson, B., Simmonds, P., Shadwick, D., Anlauf, K., Hayden, K., Schmidlin, F., Fujimoto, T., Akagi, K., Meyer, C., Nichol, S., Davies, J., Redondas, A., and Cuevas, E.: Long-term Changes in Tropospheric Ozone, Atmos. Environ., 40, 3156–3173, 2006. 27971

"Evaluation of new STE diagnosis system

M. S. Bourqui et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

|◄ | ►|

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Ordóñez, C., Brunner, D., Staehelin, J., Hadjinicolaou, P., Pyle, J. A., Jonas, M., Wernli, H., and Prévôt, A. S. H.: Strong influence of lowermost stratospheric ozone on lower tropospheric background ozone changes over Europe, Geophys. Res. Letters, 34, L07805, doi:10.1029/2006GL029113, 2007. 27971

Shapiro, M. A.: Further evidence of the mesoscale and turbulent structure of upper level jet stream – frontal zone systems, Mon. Wea. Rev., 106, 1100–1111, 1978. 27969

Sprenger, M. and Wernli, H.: A northern hemispheric climatology of cross-tropopause exchange for the ERA-15 time period, J. Geophys. Res., 108, 8521, doi:10.1029/2002JD002,636, 2003. 27970, 27979

Stevenson, D., Dentener, F., Schultz, M., Ellingsen, K., van Noije, T., Wild, O., Zeng, G., Amann, M., Atherton, C., Bell, N., Bergmann, D., Bey, I., Butler, T., Cofala, J., Collins, W., Derwent, R., Doherty, R., Drevet, J., Eskes, H., Fiore, A., Gauss, M., Hauglustaine, D., Horowitz, L., Isaksen, I., Krol, M., Lamarque, J., Lawrence, M., Montanaro, V., Muller, J., Pitari, G., Prather, M., Pyle, J., Rast, S., Rodriguez, J., Sanderson, M., Savage, N., Shindell, D., Strahan, S., Sudo, K., and Szopa, S.: Multimodel ensemble simulations of present-day and near-future tropospheric ozone, J. Geophys. Res., 111, D08301, doi:10.1029/2005JD006338, 2006. 27970

Stohl, A., Wernli, H., Bourqui, M., Forster, C., James, P., Liniger, M., Seibert, P., and Sprenger, M.: A new perspective of stratosphere-troposphere exchange, B. American Meteorol. Soc., 84, 1565–+, 2003. 27969

Tarasick, D. W., Fioletov, V. E., Wardle, D. I., Kerr, J. B., and Davies, J.: Changes in the vertical distribution of ozone over Canada from ozonesondes: 1980-2001, J. Geophys. Res., 110, doi:{10.1029/2004JD004643}, 2005. 27971

Terao, Y., Logan, J. A., Douglass, A. R., and Stolarski, R. S.: Contribution of stratospheric ozone to the interannual variability of tropospheric ozone in the northern extratropics, J. Geophys. Res., 113, D18309, doi:{10.1029/2008JD009854}, 2008. 27971

Wernli, H. and Bourqui, M.: A Lagrangian "1-year climatology" of (deep) cross-tropopause exchange in the extratropical Northern Hemisphere, J. Geophys. Res., 107, D2, doi: 10.1029/2001JD000812, 2002. 27970

Wernli, H. and Davies, H. C.: A Lagrangian-based analysis of extratropical cyclones. Part I: The method and some applications, Quart. J. Roy. Meteor. Soc., 123, 467–489, 1997.

"Evaluation of new STE diagnosis system

M. S. Bourqui et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

◀◀   ▶▶

◀   ▶

Back   Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

WMO: Atmospheric Ozone 1985, World Meteorological Organization (WMO), Geneva, Switzerland, 1986. 27973

Zierl, B. and Wirth, V.: The role of radiation for stratosphere-troposphere exchange in an upper tropospheric anticyclone, J. Geophys. Res., 102, 23883–23894, 1997. 27975

5 27969

Table 1. Number of profiles where a stratospheric intrusion was found somewhere at pressures higher than 300, 500 and 700 hPa, for each site during the campaign. P is the pressure in hPa. Numbers in brackets next to the station names are the total number of successful launches. Percentages in brackets in the other three columns give the proportion of profiles with a stratospheric intrusion relative to the number of successful launches.

	Nr observed events (% w.r.t. Nr profiles)		
Station (Nr profiles)	$P > 300$ hPa	$P > 500$ hPa	$P > 700$ hPa
Montreal (22)	21 (95 %)	20 (91 %)	11 (50 %)
Egbert (17)	15 (88 %)	14 (82 %)	6 (35 %)
Walsingham (17)	14 (82 %)	10 (59 %)	4 (24 %)
Total (56)	50 (89 %)	44 (79 %)	21 (38 %)

"Evaluation of new STE diagnosis system

M. S. Bourqui et al.

Table 2. Definition of the contingency table. See text for details.

Number of events that are	Observed	Not observed	Total
Forecast	a	b	a + b
Not forecast	c	d	c+d
Total	a + c	b + d	a + b + c + d = n

"Evaluation of new STE diagnosis system

M. S. Bourqui et al.

Table 3. Statistical evaluation of the skill of the Lagrangian STE data at representing intrusions penetrating below 300, 500 and 700 hPa. An event is here defined as the presence of stratospheric air anywhere below the given level. Contingency tables are for the three stations together. Categorical skill scores are for the three stations together, with individual stations given in bracket (in the order: Montreal, Egbert, Walsingham). PC=Proportion Correct (1.0: all correct; 0.5: randomly correct); FB = Frequency Bias (>1: high bias; <1: low bias); HR=Hit Rate (1.0: all hit; 0.0: none hit); FAR = False Alarm Ratio (0.0: no false alarm; 1.0: all false alarms).

Depth of intrusion	$P > 300$ hPa			$P > 500$ hPa			$P > 700$ hPa		
Contingency table	50	5	55	35	4	39	7	9	16
	0	1	1	9	8	17	14	26	40
	50	6	56	44	12	56	21	35	56
PC	0.91[a] (0.95, 0.88, 0.88)			0.77[a] (0.73, 0.82, 0.76)			0.59[b] (0.50, 0.59, 0.71)		
FB	1.10 (1.05, 1.13, 1.14)			0.89 (0.80, 0.93, 1.00)			0.76 (0.54, 1.17, 0.75)		
HR	1.00 (1.00, 1.00, 1.00)			0.80 (0.75, 0.86, 0.80)			0.33 (0.27, 0.50, 0.25)		
FAR	0.09 (0.05, 0.12, 0.13)			0.10 (0.06, 0.08, 0.20)			0.56 (0.50, 0.57, 0.67)		

[a]: Statistically significant at the 10 % bilateral level. [b]: Not statistically significant at the 10 % bilateral level.

"Evaluation of new STE diagnosis system

M. S. Bourqui et al.

Table 4. Statistical evaluation of the skill of the Lagrangian STE data at representing the detailed structure of intrusions. An event is here defined as the presence of stratospheric air in a 50 hPa vertical bin. The evaluation is made separately for the four pressure ranges. Contingency tables are for the three stations together. Categorical skill scores are for the three stations together, with individual stations given in bracket (in the order: Montreal, Egbert, Walsingham). PC = Proportion Correct (1.0: all correct; 0.5: randomly correct); FB=Frequency Bias (>1: high bias; <1: low bias); HR=Hit Rate (1.0: all hit; 0.0: none hit); FAR=False Alarm Ratio (0.0: no false alarm; 1.0: all false alarms).

Pressure range	$P < 300$ hPa			$300 < P < 500$ hPa			$500 < P < 700$ hPa			$700 < P < 900$ hPa		
Contingency table	129	22	151	138	43	181	68	26	94	7	19	26
	2	4	6	29	14	43	68	62	130	23	175	198
	131	26	157	167	57	224	136	88	224	30	194	224
PC	0.85* (0.95, 0.84, 0.75)			0.68* (0.73, 0.66, 0.63)			0.58* (0.60, 0.49, 0.65)			0.81* (0.77, 0.79, 0.88)		
FB	1.15 (1.06, 1.20, 1.24)			1.08 (0.97, 1.28, 1.07)			0.69 (0.58, 0.67, 1.00)			0.87 (0.56, 1.57, 1.00)		
HR	0.98 (1.00, 1.00, 0.95)			0.83 (0.83, 0.89, 0.76)			0.50 (0.52, 0.44, 0.56)			0.23 (0.22, 0.29, 0.20)		
FAR	0.15 (0.05, 0.16, 0.24)			0.24 (0.15, 0.30, 0.29)			0.28 (0.11, 0.33, 0.44)			0.73 (0.60, 0.82, 0.80)		

*: Statistically significant at the 10 % bilateral level.

Table 5. List of missed stratospheric intrusions in Montreal and clusters found in the Lagrangian STE data set that could account for the missed events if they were slightly shifted in space (sh), extended in time (ex), or if the density of trajectories were higher (de). An observed stratospheric intrusion is considered to be missed if no trajectory is found within 100 hPa and a radius of 55 km around the station. The intrusions are determined using the objective algorithm described in Sect. 3. Each cluster is identified with its number of member trajctories, and one cluster can be a candidate for more than one missed event. The clusters listed in boldface are illustrated in Fig. 7.

Event date	Missed intrusion layer [hPa]	Number of trajectories per cluster (label)
12 July	350–600	26(sh), 25(sh)
14 July	450–650	**86(de)**
15 July	600–750	**195(sh)**, **86(de)**, 27(sh)
16 July	600–750	296(de), 195(sh), 172(de), 87(ex), 57(sh), 50(sh), 27(sh), **26(sh)**, 13(ex), **11(ex)**, 10(sh)
22 July	450–650	none
23 July	650–750	14(sh)
25 July	700–800	46(de), 32(sh), 24(ex), 13(sh)
26 July	600–800	32(ex)
30 July	500–750	104(sh), 67(sh), **41(de)**, 29(sh)
1 August	700–800	129(de), 75(de), 40(sh), **16(sh)**
2 August	450–650	129(de), 75(de), 40(sh), 16(sh)
3 August	550–650	129(de), 75(de), 40(sh)
4 August	700–800	129(de)
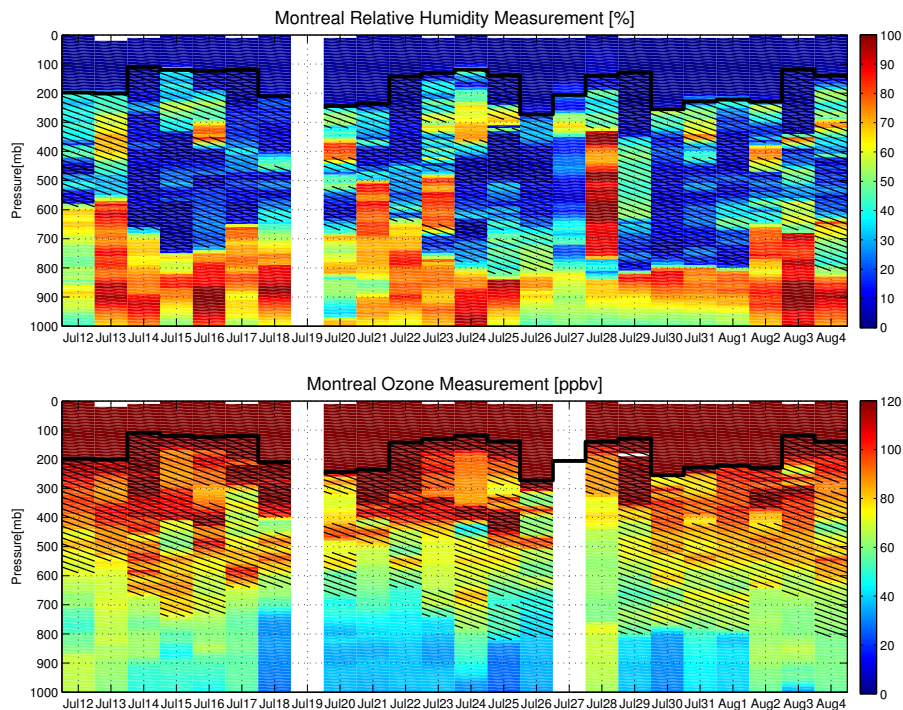
Fig. 1. Time series of daily measured profiles above Montreal. Colour shading represents relative humidity (top row, units of %) and ozone molar mixing ratio (bottom row, units of ppbv). The black thick line shows the thermal tropopause determined from the observed temperature and the black hatching shows the regions identified as stratospheric intrusions (the region between the bottom and the top intrusion levels) in the observed profiles (see Sect. 3). Blank columns correspond to missing data.
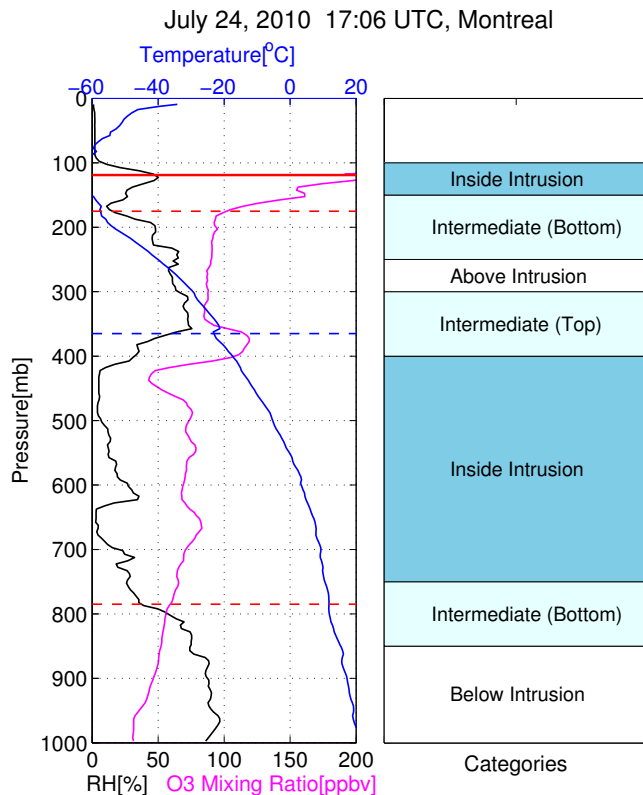
Fig. 2. Illustration of a measured profile on 24 July , 2010 above Montreal, and the corresponding categories based on the objective identification of top and bottom levels of stratospheric intrusions (see Sect. 3 for details). Left: Measurements of ozone molar mixing ratio (red line, in ppbv), relative humidity (black line, in %), and temperature (blue line, in °C), with horizontal lines marking bottom and top intrusion levels (red and blue dashed lines, respectively), and thermal tropopause (red solid line). Right: The corresponding categories assigned to each 50 hPa layer in the troposphere.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

I◀	▶I
◀	▶
Back	Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion
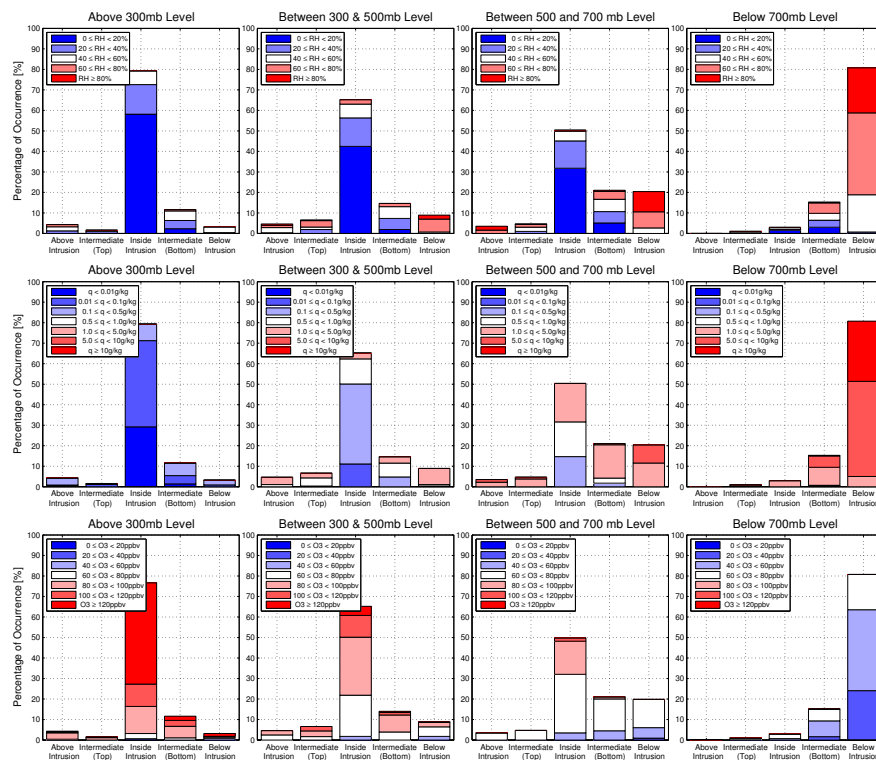
Fig. 3. Distributions of the average relative humidity (top, in %), specific humidity (middle, in g kg^{-1}), and ozone molar mixing ratio (bottom, in ppbv) within 50 hPa bins of 56 measured profiles, depending on the category to which they belong, and separately for four different pressure ranges: above 300 hPa, between 300 and 500 hPa, between 500 and 700 hPa and below 700 hPa (from left to right). Note the non-linear colour shading for the specific humidity. These distributions include all measurements from the three sites.
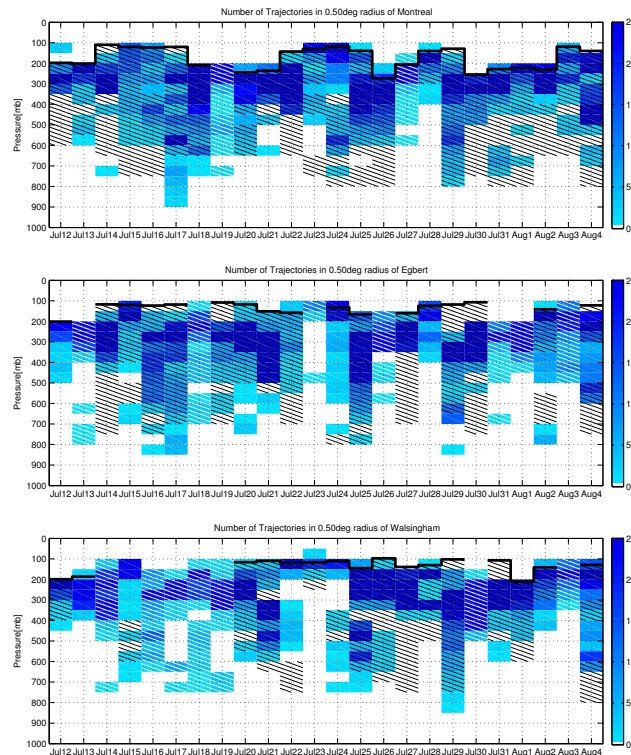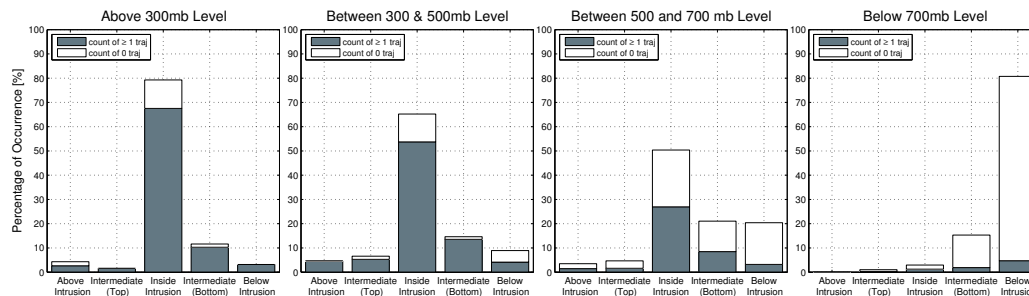
Fig. 4. Time series of the number of trajectories of stratospheric origin in the Lagrangian STE data set detected within 55 km of Montreal (top), Egbert (middle), and Walsingham (bottom) and within ± 6 h from observation time (coloured shading). The black thick line shows the thermal tropopause determined from the observed temperature and the black hatching shows the regions identified as stratospheric intrusions (the region between the bottom and the top intrusion levels shifted to the closest gridpoint on the 50 hPa grid) in the observed profiles (see Sect. 3). The white hatching denotes missing observations.

"Evaluation of new STE diagnosis system

M. S. Bourqui et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Fig. 5. Percentage of occurrence of the presence/absence of trajectories in 50 hPa bins within 55 km of the station, depending on the categories they belong, separately for four different pressure ranges: above 300 hPa; between 300 and 500 hPa; between 500 and 700 hPa; and below 700 hPa (from left to right). These distributions include all measurements from the three sites.
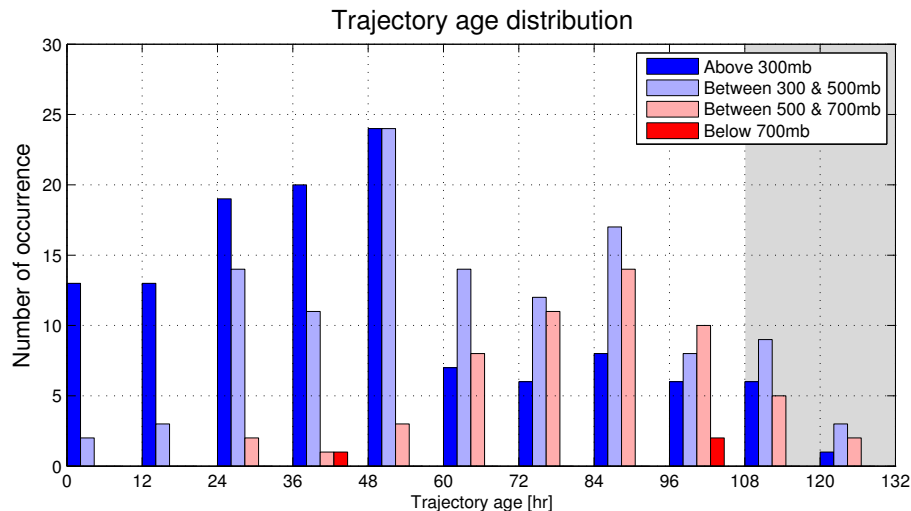
Fig. 6. Distribution of the age of trajectories when reaching one of the three stations in the "Inside Intrusion" category (see Sect. 3.2). The age is calculated as the difference between the time of crossing the 2 PVU tropopause and the average time while passing through the 55 km radius disk around the station. The distributions are separated with respect to the pressure range reached by trajectories within the 55 km radius disk around the station. The vertical axis provides the number of trajectories in each bin, without normalisation. A grey shading is applied in the region of ages between 108 and 132 h to indicate uncompleteness of the distributions caused by the limitation to 6 days of the trajectory length.

Title Page

Abstract Introduction

Conclusions References

Tables Figures
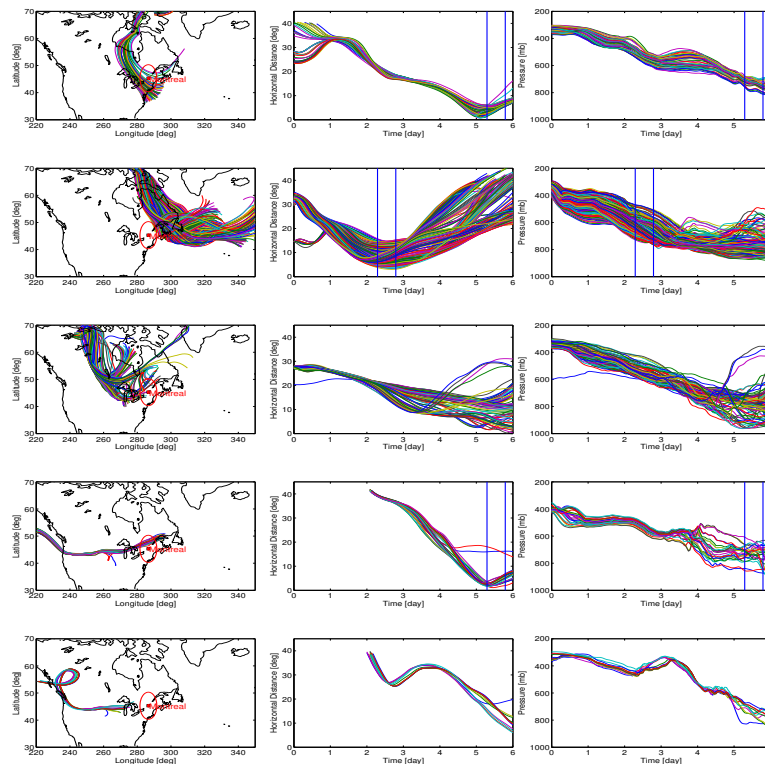
◄◄ ►►

◄ ►

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Fig. 7. Examples of typical clusters of trajectories that could be candidates for the missed events listed in bold face in Table 5. First row: Candidate **41(de)** for 30 July event, starting time is July 25 00 UTC. Second row: Candidate **195(sh)** for 15 July event, starting time is 13 July 00 UTC. Third row: Candidate **86(de)** for 15 July event, starting time is July 9 00 UTC. Fourth row: Candidate **26(sh)** for 16 July event, starting time is July 11 00 UTC. Fifth row: Candidate **11(ex)** for 16 July event, starting time is July 11 00 UTC. Left column: Geographical location of the cluster's trajectories with respect to a radius of 555 km around Montreal station (equivalent to 5° latitude). Middle column: Horizontal distance of each cluster's trajectory with respect to Montreal station (units of distance equivalent degrees latitude) as a function of time elapsed since trajectory starting time. Right column: Pressure of cluster's trajectories (hPa) as a function of time elapsed after starting time. The black vertical lines in the middle and right columns show the window ±6 h around measurement time. The trajectories are coloured with six different colours.

"Evaluation of new STE diagnosis system

M. S. Bourqui et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion