

Some challenges of middle atmosphere data assimilation

By S. POLAVARAPU^{1,2*}, T. G. SHEPHERD², Y. ROCHON¹ and S. REN²

and similar papers at core.ac.uk

brou

provided by Central Archive at

(Received 19 May 2005; revised 12 September 2005)

SUMMARY

The assimilation of measurements from the stratosphere and mesosphere is becoming increasingly common as the lids of weather prediction and climate models rise into the mesosphere and thermosphere. However, the dynamics of the middle atmosphere pose specific challenges to the assimilation of measurements from this region. Forecast-error variances can be very large in the mesosphere and this can render assimilation schemes very sensitive to the details of the specification of forecast error correlations. An example is shown where observations in the stratosphere are able to produce increments in the mesosphere. Such sensitivity of the assimilation scheme to misspecification of covariances can also amplify any existing biases in measurements or forecasts. Since both models and measurements of the middle atmosphere are known to have biases, the separation of these sources of bias remains a issue. Finally, well-known deficiencies of assimilation schemes, such as the production of imbalanced states or the assumption of zero bias, are proposed explanations for the inaccurate transport resulting from assimilated winds. The inability of assimilated winds to accurately transport constituents in the middle atmosphere remains a fundamental issue limiting the use of assimilated products for applications involving longer time-scales.

KEYWORDS: Mesosphere Variational assimilation 3D-Var

1. INTRODUCTION

Data assimilation is employed at operational weather forecast centres to combine measurements and model forecasts to obtain an initial state for model integrations. Historically, this process has focused on the troposphere where much of the regular, globally available measurements are taken. In the past decade, the lids of numerical weather-prediction models have been raised into the mesosphere in order to better assimilate radiance measurements (typically from nadir sounders aboard operational satellites) whose sensitivities extend into the upper stratosphere or higher (Untch *et al.* 1999; Swinbank *et al.* 2002; McNally 2004; Derber 2005, personal communication). At the same time, climate modellers have raised their model lids into the mesosphere or higher in order to better depict the polar stratosphere as well as the transport of constituents. The meridional circulation in the middle atmosphere is primarily wave driven, therefore the depiction of wave propagation and breaking greatly influences the transport of species. In addition, an important component of downward motion in the polar stratosphere is due to breaking gravity waves in the mesosphere (Garcia and Boville 1994).

It is also becoming increasingly important to estimate the chemical state of the atmosphere for the purposes of forecasting (i.e. chemical weather), or understanding climate trends. At weather forecast centres, ozone is being assimilated to improve the assimilation of radiance measurements, to impact on wind analyses and to improve radiance calculations in addition to forecasting total column amounts (e.g. Derber and Wu 1998; Hólm *et al.* 1999; Dethof and Hólm 2002; Jackson and Saunders 2002; Struthers *et al.* 2002). In these studies, a full general circulation model (GCM) with parametrized ozone chemistry is employed. Ozone is also assimilated using off-line

* Corresponding author: Meteorological Service of Canada, Downsview, Ontario, Canada M3H 5T4.

e-mail: saroja.polavarapu@ec.gc.ca

© Royal Meteorological Society, 2005.

models with both parametrized chemistry (e.g. Riishøjgaard *et al.* 2000; Stajner *et al.* 2001; Eskes *et al.* 2002, 2003) and complex chemistry models (e.g. Khattatov *et al.* 2000; Errera and Fonteyn 2001; Chipperfield *et al.* 2002; Marchand *et al.* 2004). This work in chemical data assimilation is partly motivated by the recent availability of constituent measurements from instruments aboard research satellites including Odin launched in 2001, TIMED in 2001, ENVISAT in 2002, SCISAT-1 in 2003 and EOS-AURA in 2004.

The goal of this work is to present some of the challenges of assimilating data from the stratospheric and mesospheric regions, commonly known as the middle atmosphere. Of necessity, the account will be brief and therefore not complete. The choice of issues presented is naturally coloured by our experiences with the Canadian Middle Atmosphere Model. It is hoped that issues that are only briefly mentioned here will be expounded upon in other articles and fora.

The plan of the paper is as follows. The fundamentals of middle atmosphere dynamics are briefly reviewed in section 2, in order to provide a context for the subsequent discussion. The issue of the vertical spreading of information upwards to the mesosphere is discussed in section 3. The importance of correctly analysing winds in the stratosphere is underlined in section 4, while the implications for balance constraints in assimilation are considered in section 5. Section 6 provides a brief discussion of other important issues and section 7 offers a conclusion.

2. BRIEF OVERVIEW OF MIDDLE ATMOSPHERE DYNAMICS

The goal of this section is to provide a very simple overview of the middle atmosphere, mentioning only the most basic features and those that impact on the design of data assimilation systems for this region. A detailed examination of middle atmosphere dynamics is found in Andrews *et al.* (1987). Shorter treatments include those of Shepherd (2000, 2003).

The distinguishing feature of the stratosphere is its strong static stability which is due to the absorption by ozone of solar ultraviolet radiation. However, a radiative equilibrium temperature is much lower than what is observed in winter in polar regions. Furthermore, the polar vortex winds determined from gradient balance with such temperatures are far stronger than observed, with the overestimation increasing with altitude. Thus the middle atmosphere is not locally in radiative equilibrium, and the departure from such a balance is due to a mechanically driven circulation. That is, the breaking of planetary and gravity waves which propagate upwards from the troposphere is associated with a transfer of momentum to the zonal mean flow, generally decelerating it. The same circulation, together with mixing processes by the breaking waves, leads to a substantial transport of ozone and in particular to maximum ozone abundance in the winter extratropical stratosphere, thereby changing the radiative balance. From this general description, it is clear that middle atmosphere modelling involves a complex interplay of chemistry, radiation and dynamics.

In the mesosphere, the divergent kinetic energy is comparable to or larger than the rotational kinetic energy spectrum for wave numbers higher than 10 (Koshyk *et al.* 1999). This reflects the prevalence of gravity waves. Gravity waves are generated in the troposphere by topography and by deep convection, as well as other processes, and propagate upwards, increasing in amplitude due to the exponential decrease in density. Orographically generated gravity waves are stationary and usually deposit their momentum in the stratosphere. Convectively generated gravity waves with large phase speeds deposit their momentum in the mesosphere and lower thermosphere where the

wind speed is larger and when their critical levels are reached. Therefore the impact of gravity waves in driving the zonal mean flow is important to the representation of the mean mass circulation and the overall depiction of the middle atmosphere. Since models without non-orographic gravity-wave drag (GWD) parametrizations inevitably suffer from the ‘cold pole problem’ in the southern polar winter (Pawson *et al.* 2000), the drag due to resolved gravity waves alone is apparently insufficient to provide enough downwelling over the southern pole. Thus it is necessary to parametrize the effects of unresolved gravity waves.

From the point of view of middle atmosphere modelling, important processes include short- and long-wave radiation, parametrization of momentum transfer by gravity waves, and a means of damping upward-propagating waves at the model lid. A simple Rayleigh friction employed near a model’s lid does not conserve momentum and can produce spurious upward and downward influences (Shepherd *et al.* 1996; Shepherd and Shaw 2004). Instead, a sponge layer should act only on the waves and not on the zonal mean flow.

3. VERTICAL INFORMATION TRANSFER

The power of data assimilation lies in its ability to spread the information of observations into regions of data gaps. If the lids of numerical models lie well above the observed regions (the troposphere and stratosphere), this leaves a data gap in the mesosphere. The propagation of information upwards from the well-observed regions can occur both through wave propagation during model forecasts and through vertical correlations in the background error statistics. Because four-dimensional (4D) data assimilation schemes often depend upon 3D background-error covariances, it is important to first consider the behaviour of a 3D assimilation scheme in spreading information vertically.

Figure 1 shows an analysis from an assimilation cycle which started from a snapshot from a climate simulation on 15 December 2001. The model used is the Canadian Middle Atmosphere Model (CMAM), a comprehensive GCM with fully interactive chemistry, radiation and dynamics, with a vertical domain extending from the surface of the earth to approximately 100 km (Beagley *et al.* 1997). The assimilation system used is described in Polavarapu *et al.* (2005). It is the same scheme that was used operationally by the Canadian Meteorological Centre (CMC) until March 2005 (Gauthier *et al.* 1999), except with modifications to use the CMAM vertical coordinate. The data used in the CMAM assimilation cycle was taken from CMC archives and is the standard meteorological input for the period (December 2001–February 2002). As a result, *in situ* measurements were available only below 10 hPa. From 10 to 1 hPa, the only observations in the system were from Advanced Microwave Sounding Unit (AMSU) channels 10–14. Thus, the measurements were available only below 1 hPa, yet Fig. 1(a) shows that the largest analysis increments are above 1 hPa. Because winds speeds are very large and fields are highly variable in the mesosphere, large increments due to mesospheric data would not be suspicious. However, in our case, observations are below 1 hPa only. A second concern is that the zonally averaged zonal mean wind analysis for the corresponding day has reduced easterlies at 1 hPa near 60°S (Fig. 1(b)). This corresponds to a local westerly wind anomaly. The summer stratosphere is a relatively quiescent place in terms of planetary wave activity, and the radiatively determined zonal wind is easterly. In climate simulations, CMAM summer stratopause winds are always easterlies, with very little variability (Wunch *et al.* 2005). While the GWD parametrization can produce westerly accelerations, this occurs at much higher altitudes.

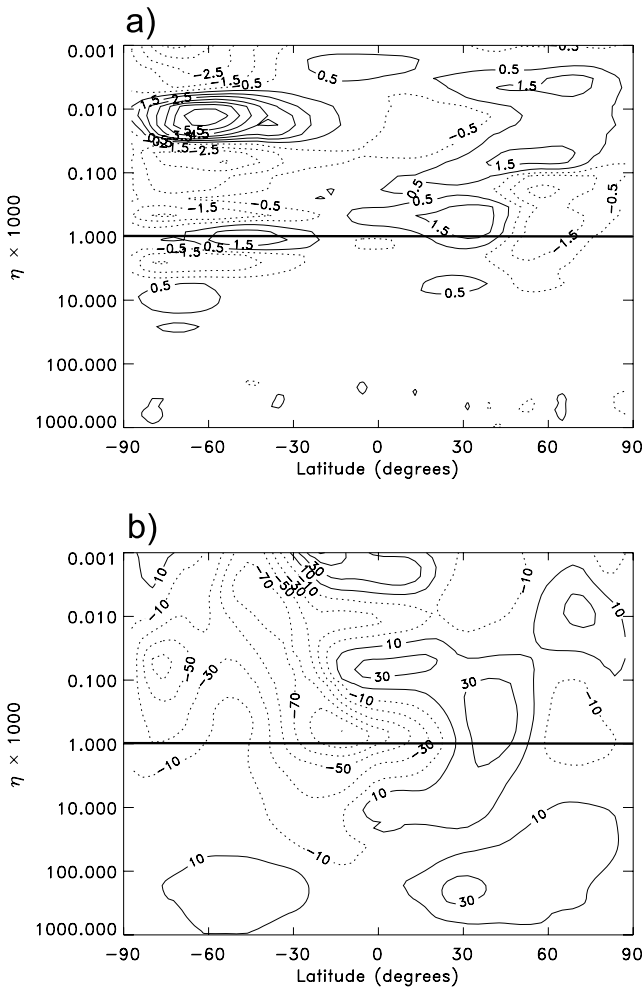


Figure 1. Zonally averaged zonal wind (a) analysis increments and (b) analysis for 12 UTC on 22 February 2002. Units are m s^{-1} , with negative contours dashed. η is the model level (see Laprise and Girard 1990) and 1000η gives an approximate pressure in hPa. *In situ* data are all below 10 hPa, while all measurements are below 1 hPa.

Furthermore, the model response to the increments in Fig. 1 was computed as the 6-hour forecast minus the initial state and the GWD forcing was analysed. The results showed that the GWD scheme was producing easterly acceleration near this location. Therefore, the reduced easterlies at 1 hPa and 60°S in Fig. 1(b) are a result of the data assimilation process and are unexpected based on the model behaviour. Figure 1 therefore points to two concerns about the assimilation: (i) the presence of large mesospheric analysis increments due to tropospheric and stratospheric observations, and (ii) the production of unphysical states in the upper stratosphere.

(a) Specification of background-error covariances

Let us first consider the issue of analysis increments. In a 3D assimilation scheme, the background-error covariance matrix determines the spatial influence of the observations. To isolate the vertical spreading of information, it is useful to consider an assimilation experiment in one dimension of a single radiance measurement. A single

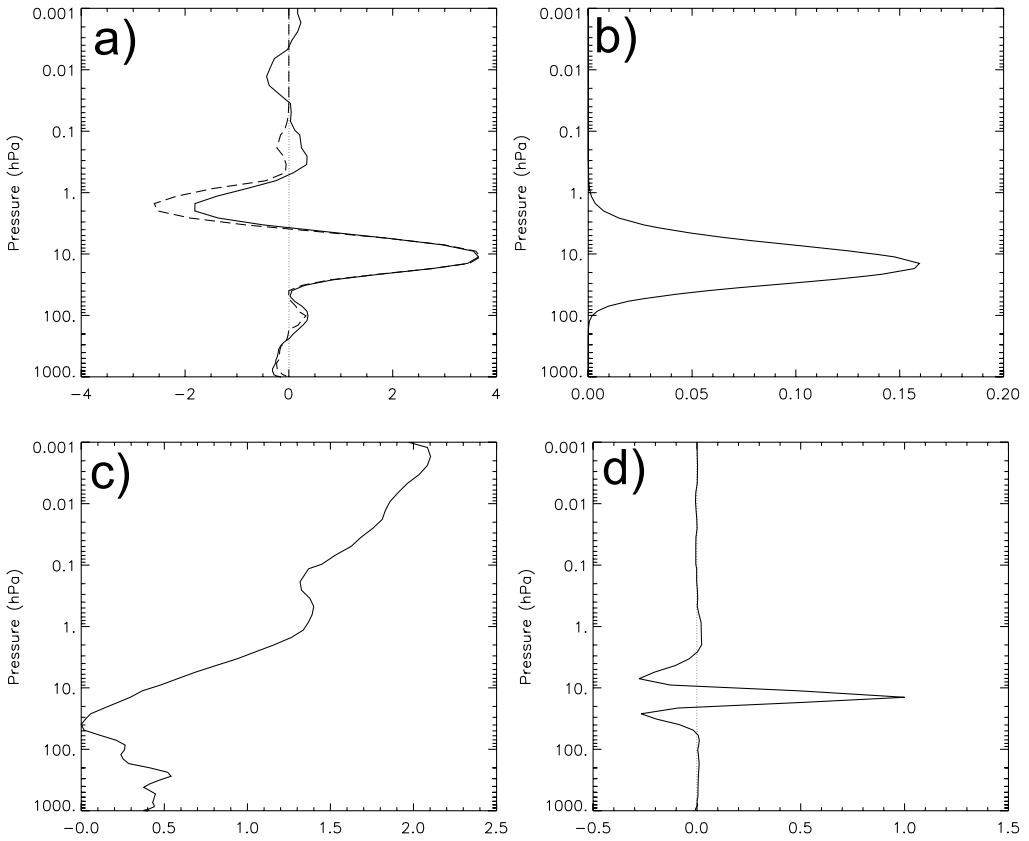


Figure 2. Results of a 1D-Var assimilation with a single observation of brightness temperature: (a) analysis increment (K) using the vertical correlation function shown in (d) (solid), and on setting small correlations to zero (dashed), (b) weighting function (K K^{-1}) corresponding to AMSU-A channel 11, (c) \log_{10} of background-error variances in K^2 , and (d) one sample vertical correlation function with a peak near 10 hPa.

observation corresponding to AMSU-A channel 11 is assimilated using an observation increment of 4 K and an observation-error variance of 0.25 K^2 . The implied background-error variance in observation space is 0.26 K^2 . Figure 2(b) shows that the weighting function for channel 11 peaks near 10 hPa but it senses temperature in the region between 100 and 1 hPa. Also shown, in Figs. 2(c) and (d), are the CMAM background-error variance for 60°S , and one of the 65 vertical correlation functions used. Note that the variance increases by a factor of 100 from the tropopause to the mesopause. When the large variance is combined with tiny but non-zero correlations, analysis increments are formed in the mesosphere (Fig. 2(a)). These tiny correlations far from the peak in Fig. 2(d) may not be justified statistically, since typically $\text{O}(10^5)$ points are used to obtain these values. Setting to zero the correlations above and below the second zero crossing from the peak in Fig. 2(d) yields the dashed curve in Fig. 2(a). Increments above 1 hPa are greatly reduced but still present due to the negative lobes retained in the correlations. The increments generally become more negative due to the removal of small positive values near the negative lobes. Note that the negative increment near 1 hPa has increased in magnitude to become nearly 75% of the positive peak near 10 hPa. The negative correlations in Fig. 2(d) appear due to the hydrostatic relation and are present in

background error statistics of operational systems (e.g. Figs. 7, 9 and 14 of Derber and Bouttier 1999; Figs. 12–13 of Gauthier *et al.* 1999; Figs. 3 and 5 of Daley and Barker 2001; Fig. 2 of Ingleby 2001). McNally (2004) also notes the tendency of the negative correlations to generate analysis increments that oscillate in sign in the vertical. Clearly, the large variances in the mesosphere make this region very sensitive to the details of the correlations specified. Are such large values realistic? First of all, the background-error covariances should reflect a 6-hour forecast error, because of the 6-hour interval used to insert observations. In the absence of forecasts, this is difficult to obtain. As a first guess, 6-hour differences of states from a climate run were used, after first removing diurnal and tidal signals (see Polavarapu *et al.* 2005). Thus the large variances seen here are due to the large variability of the fields in the CMAM mesosphere. This variability is supported by observations which suggest the ubiquitous presence of high-frequency gravity waves of large magnitude (Fritts and Alexander 2003). Part of this gravity-wave spectrum is resolved by CMAM (Manson *et al.* 2002). Since observations from the mesosphere will also have a large variance, the large variances of the background errors are justified.

The background error statistics also involve correlations. These are harder to justify, having come from model-only fields. When assimilating observations from the mesosphere, if the observations do not resolve gravity waves, these should be filtered from the analysis increments (but kept in the background state). Therefore the background-error covariances should not propagate information according to structures due to gravity waves. Removing such structures should be possible by filtering or smoothing fields prior to computation of statistics. However, the spectrum of divergent kinetic energy is quite shallow (Koshyk *et al.* 1999), so defining an ideal filter may be difficult. Tidal signals should also be removed from fields before computing statistics. This was done for the CMAM statistics. That is, stationary waves and the migrating and non-migrating diurnal and semi-diurnal tides were first removed. However, other tidal components may remain in the fields.

In summary, large error variances in the mesosphere render assimilation systems very sensitive to the details of the specification of the background-error covariances. How best to define these covariances in the mesosphere is not yet clear.

(b) *Separation of model and measurement biases*

Returning to Fig. 1, the second issue concerned the generation of unphysical states. Comparing the pattern of the increments in Fig. 1(a) to the final analysis in Fig. 1(b) reveals little similarity. This is because the spurious features in the analysis developed in time. The analysis increment pattern was similar at each cycle and therefore repeatedly inserted. The persistence of analysis increments is indicative of a bias. This was verified by averaging analysis increments over 1 week (Fig. 3). Below 10 hPa where *in situ* data were available, there is no bias. However, above 10 hPa a clear bias is present. Moreover the bias grows in time. While the observations of radiance are related to model temperatures, multivariate relations implicitly defined in the background covariances produce a bias in wind analyses from a temperature bias. The biases in the mesosphere are due to biases in the stratosphere being spread upwards through the background error correlations. This was verified by setting the vertical correlations in the mesosphere to zero, and noting that the biases disappeared (not shown). A bias in analysis increments may be due to biases in the observations or in the model. In this case, both possibilities are likely. The AMSU radiances do have significant biases which are removed by comparing to 6-hour forecast fields (following Harris and Kelly 2001). Below 10 hPa where the forecasts are unbiased (Fig. 3), this procedure is effective. However, for the

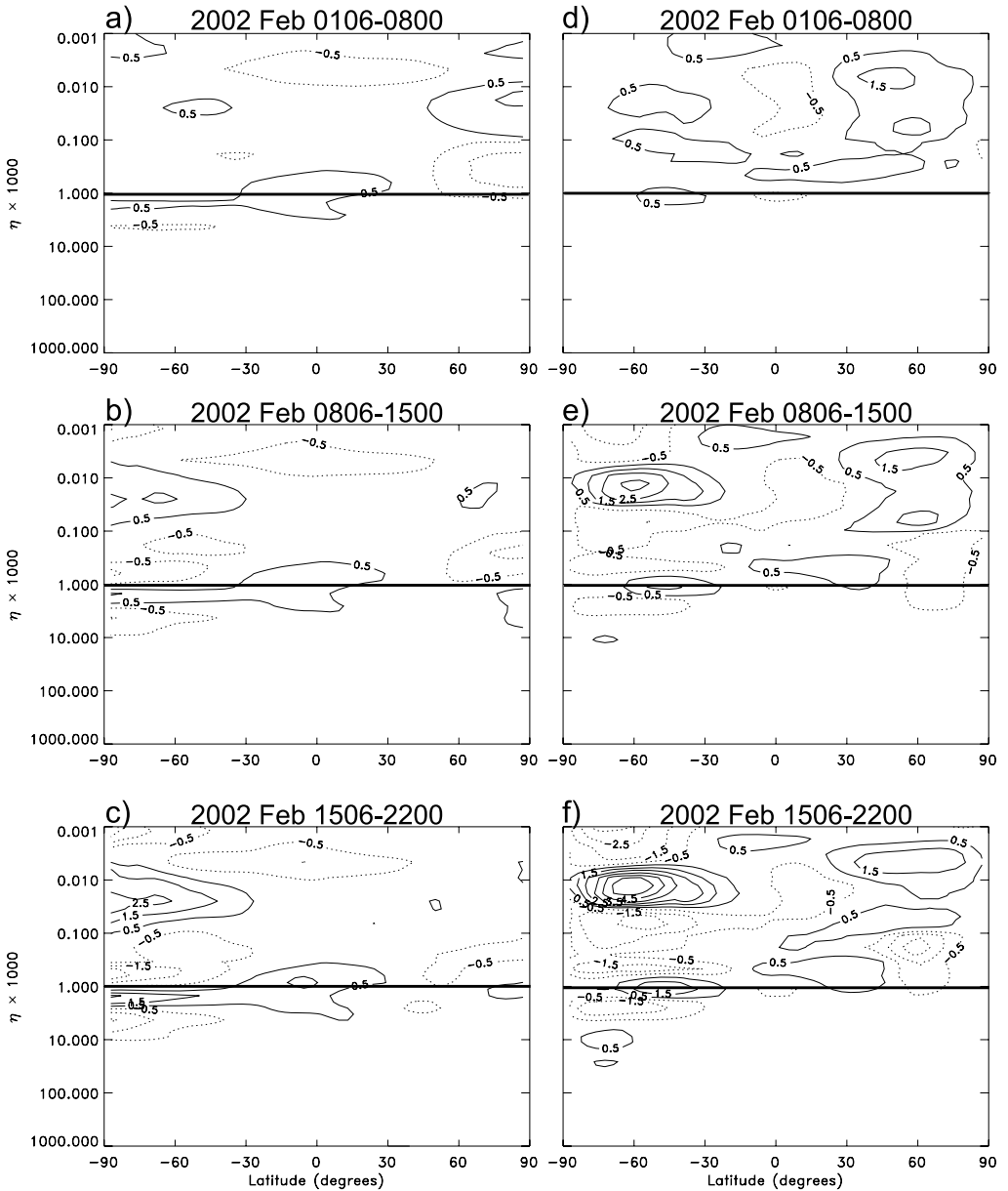


Figure 3. Weekly average analysis increments of (a, b, c) temperature (K) and of (d, e, f) zonal wind (m s^{-1}). All fields are zonally averaged and negative contours are dashed. The vertical scale is as Fig. 1. The dates used in averaging are given at the top of each panel in the form ddhh.

channels that peak above 10 hPa (AMSU-A channels 10–14), it is not clear what the best predictors of bias are. Thus far, only information from the tropospheric forecast is used to predict AMSU biases in recognition of the fact that models are well known to be biased in the stratosphere. Pawson *et al.* (2000) compared several middle-atmosphere climate models and found global mean biases in temperature, as well as zonally varying biases. Generally, models were too cold at the poles. At the equator, some models were too warm, while others were too cold. Clearly then, stratospheric forecasts can be

expected to be biased. Indeed, biases in ERA-40 analysis increments are largest in the stratosphere (Fig. 3 of Dee 2005). Therefore, an important issue in middle-atmosphere data assimilation is the identification and removal of both model and measurement biases. Bias correction schemes such as those proposed by Dee and da Silva (1998), Dee and Todling (2000), Ménard *et al.* (2004) and Dee (2005), should be useful in removing the bias. However, no scheme as yet can separate the two sources of bias (due to observations and models). Therefore, an important outstanding issue is how to separate model and measurement biases.

4. GETTING THE CORRECT BREWER–DOBSON CIRCULATION

If transport is well represented by models, then predicted distributions of chemical species can be compared with observations to evaluate the chemistry module. If a model's transport is itself erroneous, it is difficult to learn anything from such comparisons. Therefore, in order to use the vast array of species measurements now available, it is important to assess the transport implied by assimilated winds.

The distribution of species is closely linked to the Brewer–Dobson circulation. This is the slow, meridional circulation that is primarily driven by upward-propagating Rossby waves which break and exert an easterly torque on the zonal mean flow. This drives a poleward circulation within the stratosphere. By continuity, rising motion in the tropics and descending motion over the pole is implied. Because the Brewer–Dobson circulation is driven by Rossby waves, it is limited to the winter hemisphere where the mean zonal winds are westerly.

Because air irreversibly enters the stratosphere through the tropical tropopause, the concentration of a long-lived tracer with tropospheric sources and a steady tropospheric trend can be used to indicate the average time that a parcel of air has spent in the stratosphere. That is, the mixing ratio of a species such as SF₆ or annually averaged CO₂ minus its value at the tropical tropopause, divided by the trend, gives an estimate of the mean age of air. The spatial distributions of mean age tend to coincide with those of all long-lived species. From observations, mean ages of several years are inferred in the extratropical upper stratosphere. Unfortunately, the transport derived from GCM simulations is too rapid. In a comparison of a number of 2D and 3D models, Hall *et al.* (1999) found that all models had mean ages that were far too low throughout the stratosphere. They also found that models had too rapid ascent of tracers in the tropics. At least for 3D GCMs, the problem is not likely to be the mean diabatic circulation, as tropical temperatures in models are not unreasonable and in many cases suggest an insufficient upwelling (Pawson *et al.* 2000). Furthermore, Hall *et al.* (1999) found too much mixing between midlatitude and tropical air; the tropical isolation was very weak. From this one can infer that horizontal winds in the subtropics and vertical winds in the tropics are too noisy.

The situation is even worse when species are driven by assimilated winds. The age of air is then even less than when the same model is driven by GCM winds. For example observations suggest an age of 4–5 years at 65°N and 20 km, while GCM winds provide an age of 4 years and assimilated winds give only 2 years (Douglass *et al.* 2003). Weaver *et al.* (1993) showed that when performing off-line transport, the vertical velocity is far more accurate when calculated using the diabatic heating than when obtained using the horizontal eddy heat flux convergence and the mean vertical heat advection. In particular, using the horizontal eddy heat flux convergence and the mean vertical heat advection yields too much upwelling in the tropics and too much

downwelling at midlatitudes. Yet this procedure is closer to how assimilated winds are used to obtain vertical velocity.

Weaver *et al.* (1993) argue that during a forecast the winds and potential temperature satisfy the thermodynamic equation and are in thermal balance. However the insertion of data upsets this balance. Errors in the observations and in the model can excite spurious gravity waves which will dissipate (in about 1 day of simulation) but can lead to erroneous instantaneous temperature tendencies. Then excessive vertical velocities develop to balance these tendencies. In the midlatitudes, the covariances used in 3D schemes typically reflect a simple balance such as geostrophy. Since the correlations actually filter the analysis increments (see Daley 1991), this may help control the imbalance problem. Unfortunately, in the tropics no similar balance is defined.

The source of the transport problems associated with assimilated winds is still unresolved. Schoeberl *et al.* (2003) argue that tropical data assimilation can lead to spurious potential vorticity anomalies that lead to excessive ventilation of the tropics. Douglass *et al.* (2003) note that most assimilation schemes assume unbiased observation and forecast errors and that this could lead to erroneous transport when biases are present. In fact, several or even all of these hypotheses may be correct. It remains to be seen which of the deficiencies of assimilation schemes turn out to be the largest contributors to transport errors.

5. BALANCE

Due to the existence of observation and model errors on all time-scales, analyses are not balanced and spurious gravity waves are launched when integrating a model. This results in noisy forecasts of fields such as divergence, vertical velocity and surface pressure tendency. In addition, when forecasts corrupted by noise are used in quality-control procedures, the rejection of good observations or the acceptance of bad observations can occur. In 4D-Var schemes, initialization is incorporated within the assimilation step through the addition of balance constraints. However, in 3D schemes a separate initialization step is still necessary to remove spurious gravity waves from the analysis. Because of the importance of tides in the middle atmosphere, it is important to filter analysis increments only, in order to keep the tidal and diurnal signals in the background state. Moreover, because of the prevalence of gravity waves in the mesosphere, filtering the background state is not justified. These gravity waves are generated by the model and not by the assimilation procedure and are therefore presumed to be physical.

In 3D schemes, the background-error covariance also performs a filtering of analysis increments (see Daley 1991). The choice of geostrophic departure and divergence as analysis variables (as in Parrish and Derber 1992) is well suited for analysing flow in stratospheric midlatitudes where the flow is well balanced (on average). However in the mesosphere, where gravity waves significantly contribute to the kinetic energy spectrum, it is not clear whether the same choice of analysis variable is appropriate. The same argument can be made for the tropics, where geostrophic balance breaks down and vertical winds are crucial for representing the Brewer–Dobson circulation.

Typically, in the tropics and in the mesosphere, no balance is imposed. Thus winds are not derived from temperature (or radiance) in these regions. At the same time, wind observations are few, leaving winds and temperature largely unbalanced in these regions. Therefore, new wind measurements, particularly in the tropics, could be very important to stratospheric analyses. Indeed the need for wind measurements, particularly in the tropics, is a prime motivation for new satellite instruments that aim to provide such measurements: ADM-Aeolus which will be launched in 2007 (Stoffelen *et al.* 2005)

and the proposed Stratospheric Wind Interferometer for Transport Studies (SWIFT) instrument (Shepherd *et al.* 2001).

Noisy horizontal winds in the subtropics and noisy vertical winds have been implicated in the transport errors discussed in section 4. Thus insufficiently balanced analyses may contribute to these errors, as suggested by Weaver *et al.* (1993). In this case, since 4D-Var analyses tend to be in better balance than 3D-Var analyses (Gauthier and Thépaut 2001) they should produce smaller transport errors.

6. OTHER ISSUES

Many models do not capture the Quasi-Biennial Oscillation (QBO), an important source of variability in the middle atmosphere (Randel *et al.* 2004). In fact, even when models do capture the QBO, the reasons are not always clear. Capturing the QBO involves the following factors: vertical resolution, vertical diffusion, GWD parametrization and deep convective parametrization. Therefore, the modelling of the QBO is not yet well understood. This has important implications for data assimilation. If an assimilating model does not capture the QBO, tropical wind biases will be periodically present. If the assimilation scheme is not designed to cope with biases, poor tropical analyses can result. Indeed, Douglass *et al.* (2003) suggest that the presence of biases together with the use of bias-blind assimilation schemes may be the source of the too-rapid Brewer–Dobson circulation obtained from assimilated winds.

Water vapour in the upper troposphere has a very large dynamic range, making it difficult to analyse. Furthermore, errors in specific humidity are not Gaussian. To deal with this problem, the analysis variable is often chosen to be the log of specific humidity or relative humidity. However, neither variable has sufficiently Gaussian errors and other choices are better (Hólm 2002; Dee and da Silva 2003). Moisture is often analysed independently of other variables. That is, no cross correlations between moisture and other dynamic variables are specified in the background-error covariances. However, moisture is coupled to temperature in the forward models for radiance measurements. For example, AMSU channels that are sensitive to moisture are also sensitive to temperature. Therefore, an accurate temperature analysis is important for a good moisture analysis. If temperature biases are present and not corrected, the moisture analysis will be affected. Since all air irreversibly enters the middle atmosphere through the cold tropical tropopause, the saturation mixing ratio of water vapour at the tropopause gets carried into the stratospheric overworld. Moisture in the lowermost stratosphere is affected by transport from the overworld above, as well as by direct mixing with the upper troposphere. Clearly then, an accurate moisture analysis in the upper troposphere will be important to the representation of stratospheric water vapour.

A primary motivation for the assimilation of middle atmosphere measurements is the recent availability of constituent observations from missions such as Odin, TIMED, ENVISAT, SCISAT-1 and EOS-AURA. Ozone and other constituents are being assimilated for various reasons: to improve the assimilation of radiance measurements (e.g. Derber and Wu 1998), to provide analyses and forecasts of constituents (e.g. Riishøjgaard *et al.* 2000; Stajner *et al.* 2001; Dethof and Hólm 2002; Eskes *et al.* 2002, 2003), to provide input to retrieval algorithms for satellite instruments (e.g. Riishøjgaard *et al.* 2000), and to provide correlative or value-added products for new satellite instruments (e.g. Swinbank and O'Neill 1994; Errera and Fonteyn 2001; Stajner *et al.* 2001; Dethof and Hólm 2002; Struthers *et al.* 2002; Eskes *et al.* 2002, 2003). Value-added products refer to gridded fields that have filled in the gaps between measurements or estimated unobserved species. Thus far, little has been said about

the direct assimilation of constituents (as opposed to the influence of dynamic variable assimilation on species). However, some of the issues that are important for dynamic variables are clearly also important for chemical species, most notably the issue of bias. Biases in model transport will directly lead to biases in the distribution of chemical constituents. Another issue concerns the preservation of correlations among long-lived species in analyses. Such correlations are present in models (Sankey and Shepherd 2003), but the insertion of data can destroy these relations (Chipperfield *et al.* 2002). Finally, the correlation of dynamic variables and species has yet to be utilized in the assimilation context. Temperature and ozone are correlated positively in the lower stratosphere, but negatively in the upper stratosphere (e.g. Ward *et al.* 2000). Whether these relations can be exploited in assimilation systems remains to be determined.

7. CONCLUSION

The main feature of the middle atmospheric circulation is the Brewer–Dobson circulation. This slow, thermally indirect, meridional circulation is forced by vertically propagating waves generated in the troposphere, and governs the distribution of temperature and constituents. Ideally, assimilated variables (both dynamic variables and constituents) could be compared with measurements to infer model deficiencies, estimate unobserved variables, or to compute geophysical budgets. However, the value of assimilated fields for transporting constituents has been called into question not only in an Eulerian framework (eg. Schoeberl *et al.* 2003) but also in a Lagrangian one (Stohl *et al.* 2004). Furthermore, the inherent violation of geophysical constraints (such as the conservation of mass or potential vorticity) during the assimilation process can have implications for diagnostic studies that rely on consistent, closed budgets (Rood 2005).

The poor depiction of the Brewer–Dobson circulation by assimilated winds is associated with excessive mixing in the vertical in the tropics and in the horizontal between the tropics and midlatitudes. The reasons postulated for this are excess noise in analyses in the form of spurious gravity waves (Weaver *et al.* 1993), biases in models and measurements that are unaccounted for (Douglass *et al.* 2003), and spurious wave activity in the tropics (Schoeberl *et al.* 2003). Since these hypotheses all point to known deficiencies of assimilation schemes, improving transport by assimilated winds can also result in a general improvement of assimilated products. The middle atmosphere may better highlight deficiencies in schemes because of the importance of long time-scales for the relevant scientific problems (such as stratospheric ozone depletion or climate change) and the accumulation of errors in time. In addition, the large variability in the mesosphere was shown here to challenge the robustness of assimilation schemes, magnifying small uncertainties in the specified background-error covariances.

Rood (2005) notes that improvements in numerical models and assimilation schemes have led to improvements in the transport depicted by assimilated winds but nevertheless raises the issue of a fundamental limit to the use of assimilated winds for transport studies. If the hypotheses of Weaver *et al.* (1993), Douglass *et al.* (2003) and Schoeberl *et al.* (2003) are right, then improvements in transport are possible through better balance of assimilated products, by accounting for biases in observations and in model forecasts and by improving tropical analyses. Indeed, preliminary results from ECMWF (Monge Sanz, Chipperfield and Simmons 2005, personal communication) reveal greater ages of air with a 4D-Var scheme than with a 3D-Var scheme. This is to be expected if imbalance is detrimental to transport since 4D-Var analyses are expected to be in better balance than 3D-Var analyses (Gauthier and Thépaut 2001). In addition, the improved bias correction of TOVS (TIROS Operational Vertical Scanner)

radiances may be an important factor in further improvements in age-of-air calculations at ECMWF (Monge Sanz, Chipperfield and Simmons 2005, personal communication). Thus the way forward may be to consider the numerous assumptions made by most assimilation schemes, in turn, in the context of long time-scales. At the same time, more physically based, or process-oriented, diagnostics such as age of air, trajectory calculations, Eliassen–Palm fluxes, etc., applied to assimilated fields may help elucidate the consequences of these assumptions. For example, Tan *et al.* (2004) show that the insertion of data weakens potential vorticity gradients in the subtropics rendering that region of the flow more barotropically unstable. At the same time, data insertion leads to excess wave activity with phase speeds near the zonal mean wind speed, increasing the meridional dispersion of air parcels across the subtropical transport barrier. As a more physical understanding of the process of data insertion emerges, the issue of a fundamental limit to the use of assimilated fields for transport can hopefully be clarified.

ACKNOWLEDGEMENTS

This work was undertaken as part of the ‘Modelling of Global Chemistry for Climate’ project which is funded by the Natural Sciences and Engineering Research Council, the Canadian Foundation for Climate and Atmospheric Sciences, the Canadian Space Agency and the Meteorological Service of Canada (through in-kind support).

REFERENCES

- Andrews, D. G., Holton, J. R. and Leovy, C. B. 1987 *Middle atmosphere dynamics*. Academic Press
- Beagley, S. R., de Grandpré, J., Koshyk, J. N., McFarlane, N. A. and Shepherd, T. G. 1997 Radiative-dynamical climatology of the first-generation Canadian middle atmosphere model. *Atmos.–Ocean*, **35**, 293–331
- Chipperfield, M. P., Khatatov, B. V. and Lary, D. J. 2002 Sequential assimilation of stratospheric chemical observations in a three-dimensional model. *J. Geophys. Res.*, **107**, 4585, doi: 10.1029/2002JD002110
- Daley, R. 1991 *Atmospheric data analysis*. Cambridge University Press
- Daley, R. and Barker, E. 2001 NAVDAS: Formulation and diagnostics. *Mon. Weather Rev.*, **129**, 869–883
- Dee, D. P. 2005 Bias and data assimilation. *Q. J. R. Meteorol. Soc.*, **131**, 3323–3343
- Dee, D. P. and da Silva, A. M. 1998 Data assimilation in the presence of forecast bias. *Q. J. R. Meteorol. Soc.*, **124**, 246–295
- 2003 The choice of variable for atmospheric moisture analysis. *Mon. Weather Rev.*, **131**, 155–171
- Dee, D. P. and Todling, R. 2000 Data assimilation in the presence of forecast bias: the GEOS moisture analysis. *Mon. Weather Rev.*, **128**, 3268–3282
- Derber, J. C. and Bouttier, F. 1999 A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, **51A**, 195–221
- Derber, J. C. and Wu, W.-S. 1998 The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Weather Rev.*, **126**, 2287–2299
- Dethof, A. and Hólm, E. 2002 ‘Ozone in ERA-40: 1991–96’. Tech. Memo. No. 377, ECMWF, Reading, UK
- Douglass, A. R., Schoeberl, M. R. and Rood, R. B. 2003 Evaluation of transport in the lower tropical stratosphere in a global chemistry and transport model. *J. Geophys. Res.*, **108**, doi: 10.1029/2002JD002696
- Errera, Q. and Fonteyn, D. 2001 Four-dimensional variational chemical assimilation of CRISTA stratospheric measurements. *J. Geophys. Res.*, **106**(D11), 12253
- Eskes, H. J., van Velthoven, P. F. J. and Kelder, H. M. 2002 Global ozone forecasting based on RES-2 GOME observations. *Atmos. Chem. Phys.*, **2**, 271–278

- Eskes, H. J., Van Velthoven, P. F. J., Valks, P. J. M. and Kelder, H. M. 2003 Assimilation of GOME total ozone satellite observations in a three-dimensional tracer transport model. *Q. J. R. Meteorol. Soc.*, **129**, 1663–1682
- Fritts, D. C. and Alexander, M. J. 2003 Gravity wave dynamics and effects in the middle atmosphere. *Rev. Geophys.*, **41**, doi: 10.1029/2001RG000106
- Garcia, R. R. and Boville, B. A. 1994 ‘Downward Control’ of the mean meridional circulation and temperature distribution of the polar winter stratosphere. *J. Atmos. Sci.*, **51**, 2238–2245
- Gauthier, P. and Thépaut, J.-N. 2001 Impact of the digital filter as a weak constraint in the pre-operational 4D-Var assimilation system of Météo-France. *Mon. Weather Rev.*, **129**, 2089–2102
- Gauthier, P., Buehner, M. and Fillion, L. 1999 ‘Background-error statistics modelling in a 3D variational data assimilation scheme: estimation and impact on the analyses’. Pp. 131–145 in Proceedings of Seminar on diagnosis of data assimilation systems, 6–10 September 1999, ECMWF, Reading, UK
- Hall, T. M., Waugh, D. W., Boering, K. A. and Plumb, R. A. 1999 Evaluation of transport in stratospheric models. *J. Geophys. Res.*, **104**, 18815–18839
- Harris, B. A. and Kelly, G. 2001 A satellite radiance bias-correction scheme for data assimilation. *Q. J. R. Meteorol. Soc.*, **127**, 1453–1468
- Hólm, E. V. 2002 ‘Revision of the ECMWF humidity analysis: construction of a Gaussian control variable’. Pp. 1–6 in Proceedings of ECMWF/GEWEX Workshop on humidity analysis, 8–11 July 2002. ECMWF, Reading, UK
- Hólm, E. V., Untch, A., Simmons, A., Saunders, R., Bouttier, F. and Andersson, E. 1999 ‘Multivariate ozone assimilation in four-dimensional data assimilation’. Pp. 89–94 in Proceedings of SODA Workshop on chemical data assimilation. Pub. 188, KNMI, de Bilt, the Netherlands
- Ingleby, N. B. 2001 The statistical structure of forecast errors and its representation in the Met Office global 3-D variational data assimilation scheme. *Q. J. R. Meteorol. Soc.*, **127**, 209–231
- Jackson, D. and Saunders, R. 2002 ‘Ozone data assimilation: preliminary system’. Forecasting Research Technical Report No. 394, Met Office, Exeter, UK
- Khattatov, B. V., Lamarque, J.-F., Lyjak, L. V., Ménard, R., Levelt, P., Tie, X., Brasseur, G. P. and Gille, J. C. 2000 Assimilation of satellite observations of long-lived chemical species in global chemistry transport models. *J. Geophys. Res.*, **105**(D23), 29135–29144
- Koshyk, J. N., Boville, B. A., Hamilton, K., Manzini, E. and Shibata, K. 1999 Kinetic energy spectrum of horizontal motions in middle-atmosphere models. *J. Geophys. Res.*, **104**(D22), 27177–27190
- Laprise, R. and Girard, C. 1990 A spectral general circulation model using a piecewise-constant finite-element representation on a hybrid vertical coordinate system. *J. Climate*, **3**, 32–52
- McNally, A. 2004 ‘The assimilation of stratospheric satellite data at ECMWF’. Pp. 103–106 in Proceedings of the ECMWF/SPARC Workshop on modelling and assimilation for the stratosphere and tropopause, 23–26 June 2003, ECMWF, Reading, UK
- Manson, A. H., Meek, C. E., Koshyk, J. N., Franke, S., Fritts, D. C., Riggan, D., Hall, C. M., Hocking, W. K., MacDougall, J., Igarashi, K. and Vincent, R. A. 2002 Gravity-wave activity and dynamical effects in the middle atmosphere (60–90 km): observations from an MF/MLT radar network and results from the Canadian Middle Atmosphere Model (CMAM). *J. Atmos. Solar-Terr. Phys.*, **64**, 65–90
- Marchand, M., Bekki, Hauchecorne, A. and Bertaux, J.-L. 2004 Validation of the self-consistency of GOMOS NO₃, NO₂ and O₃ data using chemical data assimilation. *Geophys. Res. Lett.*, **31**, L10107, doi: 10.1029/2004GL019631
- Ménard, R., Yang, Y. and Polavarapu, S. 2004 ‘Model error estimation. Its application to chemical data assimilation’. Pp. 137–146 in Proceedings of ECMWF/SPARC Workshop on modelling and assimilation for the stratosphere and troposphere, 23–26 June 2003, ECMWF, Reading, UK
- Parrish, D. F. and Derber, J. C. 1992 The National Meteorological Center’s spectral-statistical interpolation analysis system. *Mon. Weather Rev.*, **120**, 1747–1763

- Pawson, S., Kodera, K., Hamilton, K., Shepherd, T. G., Beagley, S. R., Boville, B. A., Farrara, J. D., Fairlie, T. D. A., Kitoh, A., Lahoz, W. A., Langematz, U., Manzini, E., Rind, D. H., Scaife, A. A., Shibata, K., Simon, P., Swinbank, R., Takacs, L., Wilson, R. J., Al-Saadi, J. A., Amodei, M., Chiba, M., Coy, L., de Grandpré, J., Eckman, R., Fiorino, S. M., Grose, W. L., Koide, H., Koshyk, J. N., Li, D., Lerner, J., Mahlman, J. D., McFarlane, N. A., Mechoso, C. R., Molod, A., O'Neill, A., Pierce, R. B., Randel, W. J., Rood, R. B. and Wu, F. 2000 The GCM-Reality intercomparison Project for SPARC (GRIPS): Scientific issues and initial results. *Bull. Am. Meteorol. Soc.*, **81**, 781–796
- Polavarapu, S., Ren, S., Rochon, Y., Sankey, D., Ek, N., Koshyk, J. and Tarasick, D. 2005 Data assimilation with the Canadian Middle Atmosphere Model. *Atmos.–Ocean*, **43**, 77–100
- Randel, W., Udelhofen, P., Fleming, E., Geller, M., Gelman, M., Hamilton, K., Karoly, D., Ortland, D., Pawson, S., Swinbank, R., Wu, F., Baldwin, M., Chanin, M., Keckhut, P., Labitzke, K., Remsberg, E., Simmons, A. and Wu, D. 2004 The SPARC intercomparison of middle atmosphere climatologies. *J. Climate*, **17**, 986–1003
- Riishøjgaard, L. P., Stajner, I. and Lou, G.-P. 2000 The GEOS ozone data assimilation system. *Adv. Space Res.*, **5**, 1063–1072
- Rood, R. 2005 'Assimilation of stratospheric meteorological and constituent observations: A review'. *SPARC Newsletter*, **25**, 31–37. Available from <http://www.atmosph.physics.utoronto.ca/SPARC>
- Sankey, D. and Shepherd, T. G. 2003 Correlations of long-lived chemical species in a middle atmosphere general circulation model. *J. Geophys. Res.*, **108**, doi: 10.1029/2002JD002799
- Schoeberl, M. R., Douglass, A. R., Zhu, Z. and Pawson, S. 2003 A comparison of the lower stratospheric age spectra derived from a general circulation model and two data assimilation systems. *J. Geophys. Res.*, **108**, doi: 10.1029/2002JD002652
- Shepherd, G. G., McDade, I. C., Gault, W. A., Rochon, Y. J., Scott, A., Rowlands, N. and Buttner, G. 2001 The Stratospheric Wind Interferometer for Transport Studies (SWIFT). *Adv. Space Res.*, **27**, 1071–1079
- Shepherd, T. G. 2000 The middle atmosphere. *J. Atmos. Solar-Terr. Phys.*, **62**, 1587–1601
- 2003 Large-scale atmospheric dynamics for atmospheric chemists. *Chem. Rev.*, **103**, 4509–4531
- Shepherd, T. G. and Shaw, T. A. 2004 The angular momentum constraint on climate sensitivity and downward influence in the middle atmosphere. *J. Atmos. Sci.*, **61**, 2899–2908
- Shepherd, T. G., Semeniuk, K. and Koshyk, J. N. 1996 Sponge-layer feedbacks in middle atmosphere models. *J. Geophys. Res.*, **101**, 23447–23464
- Stajner, I., Riishøjgaard, L. P. and Rood, R. B. 2001 The GEOS ozone data assimilation system: specification of error statistics. *Q. J. R. Meteorol. Soc.*, **127**, 1069–1094
- Stoffelen, A., Pailleux, J., K allén, E., Vaughan, J. M., Isaksen, I., Flamant, P., Wergen, W., Andersson, E., Schyberg, H., Culoma, A., Meynart, R., Endemann, M. and Ingmann, P. 2005 The Atmospheric Dynamics Mission for global wind field measurement. *Bull. Am. Meteorol. Soc.*, **86**, 73–87

- Stohl, A., Cooper, O. R. and James, P. 2004 A cautionary note on the use of meteorological analysis fields for quantifying atmospheric mixing. *J. Atmos. Sci.*, **61**, 1446–1453
- Struthers, H., Brugge, R., Lahoz, W. A., O'Neill, A. and Swinbank, R. 2002 Assimilation of ozone profiles and total column measurements into a global general circulation model. *J. Geophys. Res.*, **107**(D20), doi: 10.1029/2001JD000957
- Swinbank, R. and O'Neill, A. 1994 A stratosphere–troposphere data assimilation system. *Mon. Weather Rev.*, **122**, 686–702
- Swinbank, R., Ingleby, N. B., Boorman, P. M. and Renshaw, R. J. 2002 'A 3D variational data assimilation system for the stratosphere and troposphere'. Forecasting Research Sci. Paper No. 71, Met Office, Exeter, UK
- Tan, W. W., Geller, M. A., Pawson, S. and da Silva, A. 2004 A case study of excessive subtropical transport in the stratosphere of a data assimilation system. *J. Geophys. Res.*, **109**(D11102), doi: 10.1029/2003JD004057
- Untch, A., Simmons, A., Hortal, M. and Jakob, C. 1999 'Increased stratospheric resolution in the ECMWF forecasting system'. Pp. 45–52 in Proceedings of SODA Workshop on chemical data assimilation, 9–10 December 1998. Pub. 188, KNMI, de Bilt, the Netherlands
- Ward, W. E., Oberheide, J., Riese, M., Preusse, P. and Offermann, D. 2000 'Planetary wave two signatures in CRISTA 2 ozone and temperature data'. Pp. 319–324 in *Atmospheric Science across the Stratopause*. Geophysics Monograph 123, American Geophysical Union, Washington DC, USA
- Weaver, C. J., Douglass, A. R. and Rood, R. B. 1993 Thermodynamic balance of three-dimensional stratospheric winds derived from a data assimilation procedure. *J. Atmos. Sci.*, **50**, 2987–2993
- Wunch, D., Tingley, M. P., Shepherd, T. G., Drummond, J. R., Moore, G. W. K. and Strong, K. 2005 Climatology and predictability of the late summer stratospheric zonal wind turnaround over Vanscoy, Saskatchewan. *Atmos.–Ocean*, **43**, 301–313