# The direct inversion method for data assimilation using isentropic tracer advection 

M. N. Juckes

## To cite this version:

M. N. Juckes. The direct inversion method for data assimilation using isentropic tracer advection.

Atmospheric Chemistry and Physics Discussions, 2005, 5 (5), pp.8879-8923. hal-00301791

## HAL Id: hal-00301791

## https://hal.science/hal-00301791

Submitted on 18 Jun 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## The direct inversion method for data assimilation using isentropic tracer advection

M. N. Juckes

British Atmospheric Data Centre, SSTD, Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire, OX11 0QX, UK

Received: 15 June 2005 - Accepted: 13 July 2005 - Published: 19 September 2005
Title Page

Abstract

Conclusions
Tables

14

4
Back
Full Screen / Esc
Correspondence to: M. Juckes (m.n.juckes@rl.ac.uk)
© 2005 Author(s). This work is licensed under a Creative Commons License.

Close
Introduction
References
Figures

- I
- 

Print Version

EGU

## Abstract

A new data assimilation algorithm is applied to MIPAS and SBUV measurements of stratospheric ozone. The results are validated against HALOE, POAM III, SAGE II and III, OSIRIS and ozonesonde data. The new assimilation algorithm has the accuracy of 5 the Kalman smoother but is, for the systems studied here with up to 200000 variables per time step and 61 million control variables in total, many orders of magnitude less computationally expensive. The analysis produced minimises a single penalty function evaluated over an analysis window of over one month. The cost of the analysis is found to increase nearly linearly with the number of control variables. Compared with 850 profiles from Electrochemical Concentration Cell sondes at 29 sites the analysis is found to be merely $0.1 \%$ high at 420 K , rising to $0.4 \%$ at 650 K ( 813 sonde profiles). Comparison against the other satellites imply that the bias remains small up to 1250 K ( 38 km ) and then increases to around $-10 \%$ at $1650 \mathrm{~K}(44 \mathrm{~km}$ ). Between 20 and 35 km the root-mean-square difference relative to HALOE, SAGE II and III, and POAM is in the 5 to $10 \%$ range, with larger discrepancies relative to other instruments. Outside this height range rms differences are generally larger, though agreement with HALOE remains good up to 50 km .

## 1. Introduction

The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) instrument on the ENVISAT satellite measured ozone profiles with near global coverage from July 2002 to March 2004. Compared with the operational SBUV instruments, MIPAS has improvements both in vertical resolution and in its ability to make both night and day measurements. This paper assesses the quality of the MIPAS ozone observations, and some global gridded fields derived from them, by comparison with independent observations from ozonesondes, SAGE II and III, POAM III, HALOE, OSIRIS and SBUV.

# Direct inversion 

M. N. Juckes

Title Page

Abstract

Conclusions

Tables


Print Version

EGU

The global gridded fields are generated by a new algorithm to solve the variational
formulation of the problem. The new algorithm exploits all the information available in
ACPD
5, 8879-8923, 2005 tems which have been developed in an operational context (e.g. Bloom et al., 1996; Lorenc, 1996; Courtier, 1997; Talagrand, 1997; Courtier et al., 1998; Rabier et al., 5 1998; Jeuken et al., 1999; Khattatov et al., 2000; Chipperfield et al., 2002; Struthers et al., 2002; Serafy and Kelder, 2003; Eskes et al., 2003) designed to produce real-time analyses. In this situation the information content in the analysis comes predominantly from observation prior to the analysis time. The system used here, in contrast, makes full use of observations after the analysis time, providing a significantly increased information input. Cohn et al. (1994) have discussed the advantages of making use of additional data, but found marginal benefit in their sub-optimal implementation.

The Generalised Inversion Method (GIM, e.g. Bennett et al., 1998) is closer to the approach described below, in that it fits an extended series of observations without heavy reliance on a background field. For linear systems the Generalised Inversion Method gives the same solution as the Kalman Smoother, but it uses substantially less computational resources that the latter when dealing with systems with large numbers of points in the spatial domain. The cost of GIM scales as the product of the number of observations times the number of control variables. This is much more efficient than the Kalman Smoother for systems with large numbers of variables, and makes it usable for interesting physical applications Bennett et al. (1998), but it is still (in its current formulation) substantially more expensive that currently implemented operational

## Direct inversion

M. N. Juckes


Print Version

Interactive Discussion

EGU et al. (1997) used lower resolution and their algorithm could not propagate information back in time.

The large cost of optimal algorithms has led many authors to investigate sub-optimal
systems, e.g. reverse domain filling trajectories (Dragani et al., 2002) or low order modelling of the background error covariance (e.g. Riishøjgaard, 1998, 2000, 2001).

The method used is 'direct' in the sense that it uses neither a background field to convey information between analysis segments, as used in the above cited works on 3 and 4D-VAR, nor representers to derive the impact of individual observations, as in GIM. The new method preserves the quasi-elliptic nature of the problem which follows naturally from its specification in terms of a minimisation.

## 2. Methodology and data

### 2.1. Background

${ }_{10}$ Figure 1 shows results of an assimilation run at approximately $1 / 2$ degree resolution (768 points around the equator). The positions of MIPAS measurements occurring with 2 h of the time shown are marked as coloured crosses, and those occurring within 2 to 6 h are shown as black crosses. There is clearly more information in the field than could be obtained from simple interpolation of the observations. It will be shown

## Direct inversion

M. N. Juckes

Title Page

Abstract
Conclusions

Tables


Print Version

Interactive Discussion

EGU

The observations are constrained by assuming they satisfy, within error bars, an isentropic advection equation (that is, quasi-horizontal advection on surfaces of constant
potential temperature):
$M[\chi] \equiv \chi_{t}+\boldsymbol{u} \cdot \nabla \chi=\epsilon$,
where $\chi$ is the ozone mixing ratio, $\epsilon$ is a random process with spatial and temporal correlation scales smaller than the numerical discretization, and $\boldsymbol{u}$ is the horizontal wind taken from the ECMWF operational analyses (see Table 3 in Appendix A for full reference). That is:
$E\left[\epsilon(\lambda, \phi, t) \epsilon\left(\lambda_{0}, \phi_{0}, t_{0}\right)\right]=\sigma_{\text {ap }}^{2} f_{\text {ap }}\left(\lambda, \phi, t ; \lambda_{0}, \phi_{0}, t_{0}\right)$
where the structure function $f_{\text {ap }}$ is normalised so that
$\iiint f_{\mathrm{ap}}\left(\lambda, \phi, t ; \lambda_{0}, \phi_{0}, t_{0}\right) \cos (\phi) d \phi d \lambda d t=1$,
and assumed to be localised:
$f_{\text {ap }}\left(\lambda, \phi, t ; \lambda_{0}, \phi_{0}, t_{0}\right) \ll 1$
if $\left|t-t_{0}\right|>\tau_{\text {ap }}$
or $\left(\phi-\phi_{0}\right)^{2}+\cos (\phi)^{2}\left(\lambda-\lambda_{0}\right)^{2}>I_{\mathrm{ap}}$,
for constants $\tau_{\text {ap }}$ and $/_{\text {ap }}$.
Note that the definition of $M$ here includes the time derivative and acts on the full

ACPD
5, 8879-8923, 2005

Direct inversion
M. N. Juckes

Title Page

Abstract

Conclusions

Tables


Back
Introduction References

Figures

Full Screen / Esc

Print Version gained by discretising the problem at this stage and representing $M$ by a matrix. However, keeping the analytic form allows the structure which emerges below to be exploited in the choice of discretization.

If $\epsilon$ is a random process and if the errors in the observations can also be represented by random terms, the Bayesian formulation can be used to construct a joint likelihood density function (LDF) for the structure of the fields given the observations. The LDF can be expressed as the exponential of minus a penalty function so that the minimum of the penalty function corresponds to the maximum of the likelihood density. Appendix $B$ shows how the inverse of the noise auto-correlation function can be constructed for a representative class of such functions, and shows that the inverse can be approximated by a constant if the length scale of the analysis field is large compared to the autocorrelation length scale of $\epsilon$. This will be assumed to be the case here.

Here the penalty function implied by the observations and Eq. (1) is augmented by smoothing terms:

$$
\begin{align*}
\mathcal{J} & =\sum_{i} \sigma_{\mathrm{obs}: i}^{-2}\left[\chi_{\mathrm{obs}: i}-\chi\left(\lambda_{\mathrm{obs}: i}, \phi_{\mathrm{obs}: i}, t_{\mathrm{obs}:}\right)\right]^{2} \\
& +\iint\left\{w_{\mathrm{ap}}\left[\chi_{t}+\boldsymbol{u} \cdot \nabla \chi\right]^{2}\right. \\
& \left.+w_{\text {num }}\left|\left(c_{1} \frac{\partial}{\partial t}, \nabla\right)\left(c_{1}^{2} \frac{\partial^{2} \chi}{\partial t^{2}}+\nabla^{2} \chi\right)\right|^{2}\right\} d A d t \tag{5}
\end{align*}
$$

, predixed by $w_{\text {num }}$, imposes regularty near observations. $c_{1}$ is a constant determining the ratio between spatial and temporal smoothing. Here $c_{1}=0.5$ days per radian is used. The scalar weighting coefficient $w_{\text {ap }}$ would ideally be derived from the covariance of a random process model error. As the latter is generally not random, value of $w_{\text {ap }}$ will be determined empirically.

Applying the calculus of variations to the penalty function in Eq. (5) shows that it is minimised with respect to $\chi(\theta, \phi, t)$ when that field satisfies the following analysis equation:

## ACPD

5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

Title Page

Abstract
Conclusions
Tables


Print Version

Interactive Discussion

EGU
$\mathcal{A}[\chi] \equiv \mathcal{A}_{\text {obs }}+\mathcal{A}_{\mathrm{ap}}+\mathcal{A}_{\text {num }}=\mathcal{S}$
where

$$
\begin{align*}
\mathcal{A}_{\mathrm{obs}}= & \sum_{i} \frac{\chi\left(\lambda_{\mathrm{obs}: i}, \phi_{\mathrm{obs}: i}, t_{\mathrm{obs}: i}\right)}{\sigma_{\mathrm{obs}: i}^{2}} \times \\
& \delta\left(t-t_{\mathrm{obs}: i}\right) \delta\left(\lambda-\lambda_{\mathrm{obs}: i}\right) \delta\left(\phi-\phi_{\mathrm{obs}: i}\right),  \tag{7a}\\
\mathcal{A}_{\mathrm{ap}}= & w_{\mathrm{ap}}\left(-\frac{D^{2} \chi}{D t^{2}}\right),  \tag{7b}\\
\mathcal{A}_{\mathrm{num}}= & -w_{\mathrm{num}}\left(c_{1}^{2} \frac{\partial^{2}}{\partial t^{2}}+\nabla^{2}\right)^{3} \chi,  \tag{7c}\\
\mathcal{S}= & \sum_{i} \frac{\chi_{i}}{\sigma_{\mathrm{obs}: i}^{2}} \delta\left(t-t_{i}\right) \delta\left(\lambda-\lambda_{i}\right) \delta\left(\phi-\phi_{i}\right) .
\end{align*}
$$

The boundary conditions at the start and end times ( $t_{\text {start }}$ and $t_{\text {end }}$, respectively) are:

$$
\begin{align*}
& \frac{D \chi}{D t}, \frac{\partial}{\partial t}\left(c_{1}^{2} \frac{\partial^{2}}{\partial t^{2}}+\nabla^{2}\right)^{n} \chi \equiv 0  \tag{8}\\
& \quad \text { for } t=t_{\text {start }}, t_{\text {end }}, \text { and } n=1,2
\end{align*}
$$

Equations (6) and (7) can also be derived by forming the Euler-Lagrange equations from (1) (e.g. Bennett, 1992) and then eliminating the Lagrange multiplier. The relaxation algorithm described below cannot be applied to the Euler-Lagrange equations themselves because of the intrinsic non-localness in those equations. A small forcing applied to either equation will generally produce a global response because the

ACPD
5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

Title Page
Abstract
Introduction
Conclusions
References
Figures


Back
Close
Full Screen / Esc

Print Version

EGU

$$
\begin{equation*}
\sum_{i} \frac{\chi_{i}-\chi\left(\lambda_{\mathrm{obs}: i}, \phi_{\mathrm{obs}: i}, t_{\mathrm{obs}: i}\right)}{\sigma_{\mathrm{obs}: i}^{2}}=0, \tag{9}
\end{equation*}
$$

that is, the analysis is unbiased with respect to the observations. This also means that any bias in the observations is present, without modification, in the analysis. Equation (9) gives a simple relation between the analysis averaged over the observation points and the observations. Unfortunately, there is no simple expression relating the

The spatial mesh used is a latitude-longitude grid which thins towards the pole so that the longitudinal spacing does not decrease too drastically. It is described in more detail in Appendix C. This grid preserves some of the simplicity of the latitude-longitude grid but avoids the convergence of points at the poles which can create numerical conditioning problems (e.g. Thuburn and Li, 2000, and references therein). In the calculations described below the longitudinal spacing at the equator is equal to the latitudinal spacing. Five different resolutions will be employed having $12 \times 2^{N_{g r i d}}$ points around the equator for $N_{\text {grid }}=3$ to 7 . More details are listed in Table 1 below.

## ACPD

5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

Title Page


Full Screen / Esc

Print Version

Interactive Discussion

The solution method is a a multi-grid relaxation algorithm described in Appendices D and $E$. This is an iterative method, and the cost of each iteration is proportional to the number of control variables. The number of iterations appears to be independent of the resolution, so that the net cost of the solution algorithm increases only linearly with the problem size.

The second order time derivative is discretised using the standard 3-point formula. The algorithm does not require the integration of any first order time derivatives, so the usual problems of numerical stability found with forecast and adjoint models do not arise here.
The Lagrangian derivative is evaluated using a semi-Lagrangian scheme with an implicit definition of the parcel displacements:

$$
\begin{equation*}
\frac{D^{2} \chi}{D t^{2}} \stackrel{\text { def }}{=} \frac{\chi\left(\mathbf{x}_{m}, t_{m}\right)-2 \chi(\mathbf{x}, t)+\chi\left(\mathbf{x}_{p}, t_{p}\right)}{\Delta t^{2}} \tag{10}
\end{equation*}
$$

$$
\begin{equation*}
\mathbf{x}-\mathbf{x}_{m}=\frac{\Delta t}{2}\left[\mathbf{u}(\mathbf{x}, t)+\mathbf{u}\left(\mathbf{x}_{m}, t_{m}\right)\right] \tag{11}
\end{equation*}
$$

$$
\begin{equation*}
\mathbf{x}_{p}-\mathbf{x}=\frac{\Delta t}{2}\left[\mathbf{u}\left(\mathbf{x}_{p}, t_{p}\right)+\mathbf{u}(\mathbf{x}, t)\right] \tag{12}
\end{equation*}
$$

ACPD
5, 8879-8923, 2005

Direct inversion
M. N. Juckes

Title Page

Abstract
Conclusions
Tables


EGU

### 2.4. The smoothing term

The smoothing term ensures that the problem has a unique solution. The dynamical penalty will be zero for any field satisfying $D \chi / D t=0$. Since there are an uncountable infinity of such fields and only a finite number of observations, it is clear that the we will in general have an infinite number of solutions which fit all the observations and have zero dynamical penalty. These solutions are such that the value on any trajectory intersecting an observations is given by that observation. In the absence of any additional information the value on trajectories which do not intersect observations is undetermined.

The form of the smoothing term used is constrained by consideration of the structure of the solution near the observations. This structure is determined by a balance between the $\delta$-function forcing and the highest derivatives in the equation (because the $\delta$-function forcing generates small scales, and with small scales the higher derivatives have the largest magnitude). The highest derivatives here are, by construction, homogeneous in space and time when time is scaled by $c_{1}$. Let $r^{2}=\Delta \phi^{2}+\cos \phi^{2} \delta \lambda^{2}+c_{1}^{-2} \Delta t^{2}$ be the space-time distance from an observation. The solution near the observation is then
$\frac{r^{3}}{4 \pi}\left[\chi\left(\lambda_{i}, \phi_{i}, t_{i}\right)-\chi_{i}\right]+a_{1} r^{2}+a_{2}$
where $a_{1}$ and $a_{2}$ are constants determined by the larger scale solution. If the smoothing term in the analysis equation were 4th order, then the leading order term near the observations would be proportional to $r$. This would imply a singularity in the gradient at each observation. Having a 6th order term guarantees a sufficient degree of smoothness of the solution in the vicinity of the observations so that all the derivatives in the physically motivated a priori constraint can be accurately evaluated.

The solution method described in Appendices D and E requires increasing numbers

## Direct inversion

M. N. Juckes

Title Page

Figures


Print Version of ancillary variables as the order of the equation is increased. In order to keep the computational cost down it is desirable to use the smallest suitable value, which is 6th

### 2.5. Experimental setup

5, 8879-8923, 2005
The main results will be presented at a resolution of $1.875^{\circ}$. The dynamical and numerical weighting will be non-dimensionalised by a rate constants $c_{2}=0.4 \mathrm{ppmv} /$ day and $c_{3}=0.2 \mathrm{ppmv} / \mathrm{radian}$, respectively, with

$$
\begin{align*}
w_{\mathrm{ap}} & =w_{\mathrm{ap}}^{*} c_{2}^{-2}  \tag{13}\\
w_{\text {num }} & =w_{\text {num }}^{*} c_{3}^{-6} \tag{14}
\end{align*}
$$

The results discussed below use $w_{\text {ap }}^{*}=8$ and $w_{\text {num }}^{*}=0.1 \times 2^{-6 N_{\text {grid }}}$ unless otherwise stated, where $N_{\text {grid }}$ is the number of levels in the multigrid hierarchy ( $N_{\text {grid }}=5$ for the resolution of $1.875^{\circ}$ ).

Convergence is expressed in terms of a tolerance, $T_{\text {conv }}=0.005$ unless otherwise stated. The iteration is stopped when: (i) the mean square residual in the assimilation equation is less then $T_{\text {conv }}$ times the mean square solution, (ii) the fractional changes, over 3 iterations, in the a priori, numerical and observational components of the cost function are less than 2,5 and 2 times $T_{\text {conv }}$, respectively.

### 2.6. Analysis time windows

The 6 month period discussed here is broken up into 6 overlapping segments, and a separate minimisation carried out for each segment. The six time periods: (1) 25 January to 12 March, (2) 25 February to 12 April, (3) 25 March to 12 May, (4) 25 April to 20 June, (5) 25 May to 12 July and (6) 25 June to 12 August, all in 2003. At points

Direct inversion
M. N. Juckes

Title Page

Abstract
Conclusions
Tables


Print Version
Interactive Discussion

EGU well away from the ends of the time windows the analysis benefits from information of both future and past observations. Near the ends of the time windows this is no longer the case and we can expect some loss of accuracy as a result. The differences between the analyses in the overlap period will be used below to assess the amplitude of the random error in the analysis.

The loss of accuracy at the ends of the time segment could be avoided by using a more sophisticated boundary condition at the end point, but this would require handling a large error covariance matrix. Here the period near the end of the assimilation is used only for quality control. In the forecast situation neither of these latter two options is available: we must simply accept that the error in a real-time assimilation will be larger than in a retrospective one.

## 3. Results

### 3.1. Hemispheric fields

The MIPAS and SBUV datasets provide near global coverage (the MIPAS dataset has can be used to generate global analyses.

Figure 2 shows fields for 00:00 GMT, 10 July 2003, for analyses of MIPAS and SBUV data. This is a period when a wave-breaking event is pulling a streamer of polar vortex air into mid-latitudes. These plots are taken from assimilations at $1.875^{\circ}$ resolution.

# Direct inversion 

M. N. Juckes

Title Page

Abstract
Conclusions

Tables


Print Version

### 3.2. Random error

We expect there to be a loss of accuracy near the end of the analysis time windows. This can be assessed by looking at the the difference between two overlapping assimilations. Figure 3 shows the logarithm of the RMS difference between the June and July segments:
$E(t)=\sqrt{\frac{1}{4 \pi} \int\left(\chi^{(5)}-\chi^{(6)}\right)^{2} d A}$,
where the superscripts (5) and (6) refer to the analysis time segment listed in Sect. 2.6. In the central region the difference levels out at around 0.05 ppmv . This level appears to be related to the convergence criteria used to determine when to stop the iteration towards a solution of Eqs. (6) and (7).

The dashed lines in Fig. 3 show results from experiments using different values of $w_{\text {ap }}$. A larger weighting causes information to persist in the assimilation for longer, and hence the loss of accuracy at the ends of the assimilation window persist for longer.

At the ends of the overlap period we see a near exponential increase in E. The differences at the ends of the analysis windows (around 0.22 ppmv and 0.14 ppmv on 25 June and 12 July respectively) reflect the difference in accuracy between the Kalman Filter and the Kalman Smoother. The e-folding timescale for information persistence is shown by Fig. 2 to be around 2 to 3 days, depending on the magnitude of $w_{\mathrm{ap}}$. This means that the information input to an analysis at any time, $t_{0}$ say, comes from a period of 4 to 6 days, and the impact of observations within the same time step as $t_{0}$ is relatively small. Consequently, the information content of the Kalman Smoother will be close to twice that of the Kalman Filter. In other words, the error variance at the end of the assimilation window will be twice that in the centre. This allows us to make a heuristic estimate of the random error in the assimilation: Let $K F, K S$ and $B F$ be the solutions of the Kalman Filter, Kalman Smoother and Backward (or reverse) Kalman Filter respectively. If the impact of observations shared by the forward and backward

5, 8879-8923, 2005

## Direct inversion

M. N. Juckes


Print Version
filters is neglected:

$$
\begin{align*}
\overline{(K F-K S)^{2}} & =\overline{\left[\frac{1}{2}(B F-K F)\right]^{2}}  \tag{16}\\
& =\frac{1}{2} \sigma_{K F}^{2}=\sigma_{K S}^{2}, \tag{17}
\end{align*}
$$

where $\sigma^{2}$ is the random error variance and the overbar denotes a global mean. This tion window, is around 0.23 ppmv . This value is consistent with the differences between the analyses and independent observations evaluated below. Figure 3 does not give any indication of possible systematic errors: a bias in the MIPAS measurements, for instance, would lead to a common error in the two assimilation segments which would used below to assess the systematic errors and give a second estimate of the random errors.

### 3.3. Numerical cost of the method

Solving the Laplace equation with a multigrid relaxation algorithm is known to have a numerical cost that scales with $N$, where $N$ is the number of variables in the solution. Table 1 shows how the numerical cost of solving the analysis Eq. (7d) varies with resolution using the current algorithm. The computations have been carried out on a single 2 GHz processor. To provide a means of comparing with results on different platforms, the cost is also presented in terms of the number of CPU cycles per control variable. The software is still at an early stage of development with many optimisation issues yet to be explored, so there is potential for improvement on the absolute values of these costs. The most important result at this stage is the ability of the algorithm to deliver near linear dependence of cost on problem size, hence making it a promising candidate for significantly larger data assimilation problems.

5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

Title Page

Abstract
Conclusions

Tables


Print Version

Interactive Discussion

EGU

Table 2 lists the biasses and variances of the differences between the different types of
5, 8879-8923, 2005

Direct inversion
M. N. Juckes

Title Page
of sondes. They conclude that the ECC sonde is the most precise. The table shows systematic differences between the analysis and this sonde well below $1 \%$ between 420 K and 650 K , averaged over a total of over 800 sonde profiles.

The Indian sondes give anomalously large differences, reading less than half the
The Indian sondes give anomalously large differences, reading less than half the
mixing ratio of the MIPAS analysis at 420 K . In the following diagrams, measurements from these sondes have been omitted.

The comparison with ozonesondes is in line with the results of Migliorini et al. (2004),
who found negligible biases in a sample of 30 near coincident measurements, and a difference variance of around 0.35 ppmv . Here, the rms difference from the ECC sondes, averaged over the 3 lowest levels, is 0.3 ppmv . The differences are greater at 850 K , but the comparison with other satellite instruments (below) suggests that this is due to loss of accuracy in the sondes rather than in the analysis.

Figure 5 below also shows comparisons against ozone sonde profiles which have been vertically smoothed over 2 km . This does not affect the bias, but it can be seen that it makes a significant reduction in the variance, implying that some of the disagreement between sondes and analysis is due to the lower vertical resolution in the latter.
where the subscript "inst" refers to the instrument used, the sum is taken over all observations at each level, $N_{\text {inst }}$ is the number of observations in the sum. All the sonde data is taken from the World Ozone and Ultraviolet Radiation Data Centre (WOUDC) website (see Appendix A for details), the complete list of the 37 stations used is in Appendix B.

Smit and Kley (1998) describe an extensive intercomparison between different types

A more accurate method of dealing with this problem is described in Migliorini et al.

### 3.5. Validation against other satellite measurements

The scope of stratospheric observations is steadily increasing. This paper analyses ozone measurements from 7 satellite instruments: MIPAS (Tsidu et al., 2003), SBUV (Planet et al., 2001) SAGE II and III (Thomason and Taha, 2003), POAM III (Lumpe et al., 2003; Pierce et al., 2003), OSIRIS (von Savigny et al., 2003), and HALOE (Brühl et al., 1996). The use of a wide range of instruments provides insight, in some instances, into the source of discrepancies when a single instrument departs from the majority. Brief details of the instruments are given in Appendix A.

Figure 4 shows a sample comparison for July 2003, at 500 K and 850 K. For SBUV and OSIRIS the observations, measurement minus MIPAS analysis, evaluated at each profile location are averaged in $5^{\circ}$ latitudinal bins and plotted as a solid purple and green lines, respectively, with dashed lines showing plus and minus one standard deviation. For the other measurements a symbol is plotted, as detailed in the figure caption, for each measurement. There is little systematic variation of the residuals with latitude. The SBUV measurements show a low bias at 850 K and a high bias at 500 K .

Figure 5 shows the the means and variance of the residuals averaged over the 6 month assimilation period, February-July, 2003. Below 35 km the assimilation has

SAGE III measures slightly higher than the assimilation, SBUV is substantially higher, especially in the mid to upper stratosphere. The OSIRIS measurements show a large height dependent bias.

The bias in the SBUV measurements may be due to low vertical resolution: as the

Direct inversion
M. N. Juckes


Print Version

Interactive Discussion

EGU ozone number density has a concave profile in the mid and upper stratosphere a pos-
itive bias would be an expected consequence of vertical averaging associated with broad weighting functions.

Between 35 km and 50 km it appears that MIPAS is measuring low relative to all the
other instruments. In this region the difference between HALOE and MIPAS is substantially smaller than the difference with respect to other instruments. As HALOE and MIPAS both rely on infra-red measurements, while SAGE, POAM and SBUV exploit ultra-violet radiation, this may point to a problem with the underlying spectroscopy. These results do not, however, give any indication as to whether the UV or IR instruments are more reliable.

Figures 6 and 7 show some analogous plots created without any data assimilation. Instead, differences between nearby measurements have been evaluated. Pairs of measurements are included in the comparison if they fall within 6 h and 2.5 great circle degrees of each other. This constraint means that there are far fewer data points, but still enough. The same pattern of differences emerges, showing that the systematic differences seen in Fig. 5 are not generated by the assimilation system. The differences are generally larger in the "nearest neighbour" plots, showing that the gridded fields bring significant added value.

Two factors could account for this: firstly, the assimilation accounts for the spatial and temporal variation of the ozone field. This should improve the intercomparison, provided that the variations are represented with sufficient accuracy. Secondly, the value of the assimilation at any point is a weighted average of many MIPAS observations. This could lead to a reduction in the random error.

Figure 4a also gives an indication of the biases between the other instruments. These results are in line with Danilin et al. (2002), who show HALOE ozone measurements to be around $4 \%$ systematically lower than SAGE II.

At 30 km there is a cluster of instruments with $\mathcal{V}_{\text {inst }} \approx 0.35 \mathrm{ppmv}$. If this difference is equally partitioned between the assimilation and the verifying instruments, and if these differences are assumed to be indpendent, this implies a random error in the analysis, at this height, of 0.25 ppmv , which is very close to the estimate ( 0.23 ppmv ) derived in Sect. 2 from overlapping analyses.

# Direct inversion 

M. N. Juckes


Print Version

### 3.6. Dependence on resolution

Comparison of Figs. 1 and 2a suggest that there is little change in the analysis when
5, 8879-8923, 2005

## Direct inversion

A quantitative evaluation is shown in Fig. 8, with resolutions varying from $7.5^{\circ}$ down to $0.47^{\circ}$. The root-mean-square departure from observations is plotted, using the 850 K analyses and the entire 6 month study period.

There is a clear improvement between $N_{\text {eq }}=48$ and 96 , and a small but still significant improvement when the resolution is increased again to 192. After that point, however, there is no significant gain in accuracy as measured by the $\mathcal{V}_{\text {inst }}$ validation statistic.

This invariance to changes in resolution also shows that the smoothing term is not having a major influence on the solution for $N_{\text {eq }} \geq 192$. The coefficient of the smoothing term varies as the 6th power of the resolution, so there is a reduction by a factor of $4^{6}=4096$ in the amplitude of this term between the $N_{\text {eq }}=192$ and the $N_{\text {eq }}=768$ analyses.

### 3.7. Dependence on weighting of the process model

Figure 9 shows how $\nu_{\text {inst }}$ varies as the weighting of the process model, $w_{\text {ap }}$ is varied.
There is remarkably little sensitivity over a wide range of values.
It may be tempting to think of the process model term in the cost function as being "like" the background term in the widely used strong constraint version of 4D-VAR formulated by Talagrand (1997). In the latter, the cost function consists of two terms, the non-observation term being a background term. Thus, in so far as it is a nonobservation term, the process model term here is "like" the background term. However, unlike the background term used in 4D-VAR, the process model term here does not involve any empirical correlation structures. A further important point, which may
M. N. Juckes

Full Screen / Esc
Title Page
ntroduction

References

Figures


Print Version

EGU
explain the insensitivity of the results to $w_{\text {ap }}$, is that the process model term has a large null space: that is, there is a large family of solutions of $D \chi / D t=0$ for which $A_{\text {ap }} \equiv 0$. Increasing $w_{\text {ap }}$ will move the analysis closer to the null space of $\mathcal{A}_{\text {ap }}$. However, this might not be a very large change in absolute terms if the distance between $\chi$ and that null space is already small. With the standard form of 4D-VAR the background term has no null space - it can specify the solution completely. If the weighting of this term is increased the contribution of the observations will eventually become insignificant. This cannot happen here, there will always be a significant input of information from the observations, no matter how large $w_{\text {ap }}$ is made. It is, however, found that, with the present iteration algorithm, increasing $w_{\text {ap }}$ slows the convergence.

## 4. Conclusions

The direct inversion method for data assimilation, which applies the evolution equations as a weak constraint, has been demonstrated to work for isentropic analyses of stratospheric tracers. In the context of global geophysical data sets, this is a moderate sized system: much smaller than operational meteorological analysis systems but large enough for some of the problems of associated with large systems to be relevant. For the Kalman filter, for example, the cost scales with the square of the number of control variables held at each time level. With the present method the cost is nearer to being linearly dependent on the number of control variables.

The efficiency of the algorithm makes it possible to increase the spatial resolution to the point were the solution becomes essentially independent of the resolution.

The resulting analyses have been compared with radiosondes and a wide range of other satellite instruments. The bias relative to the ECC sonde data is extremely small, less than $0.5 \%$ below 650 K . Agreement with HALOE, SAGE II, POAM III and SAGE III lunar occultation retrievals is good up to around 40 km . Above that height the MIPAS analysis develops a significant negative bias. SAGE III measures high relative to the analysis throughout the stratosphere, as does SBUV above 25 km .

## Direct inversion

M. N. Juckes

Title Page

Figures


Print Version
Interactive Discussion

The random error estimates of 0.23 ppmv (about $4 \%$ ) at 850 K derived from comparison of overlapping analysis periods in Sect. 3.2 are consistent with the variance of differences from other instruments. This is closer than agreement found by analysing

ACPD
5, 8879-8923, 2005

## Direct inversion

M. N. Juckes
typical e-folding time for the decay of influence of an observation in the analysis is found to be 2.5 days. This time-scale depends on the assumed skill of the model, as reflected in the weighting given to the model constraint in the analysis.

The gridded ozone fields produced in this study are available from http://home.badc. rl.ac.uk/mjuckes/mista/ as Netcdf files (CF compliant).

## Appendix A: the instruments

There are three main classes of instruments in this study. Four occultation instruments (HALOE, POAM III, SAGE II and III) measure the absorption of solar radiation by viewing the sun through the atmospheric limb as the satellite moves into and out of darkness. This gives good vertical resolution and accurate results, but is restricted to around 28 profiles per day. This is not enough to represent the spatial structure in the fields. Two instruments measure scattered solar radiation. SBUV is nadir viewing, measuring backscattered UV. This provides good spatial coverage in the sunlit hemisphere, but the vertical resolution is poor. Osiris measures limb scattered UV, giving better vertical resolution, but the precision is poor. Lastly, MIPAS measures the emitted infrared spectrum with limb viewing geometry. This provides global coverage, day and night. The vertical resolution is moderate, better than the nadir viewing instruments but not as good as that achieved by the occultation instruments

SAGE III data were obtained from the NASA Langley Research Center EOSDIS Distributed Active Archive Center. A few details of the instruments are listed in Table 3, together with the websites which contain further information and access to the data.

The ozonesonde data used comprises profiles from 37 stations. WOUDC is one of five World Data Centres which are part of the Global Atmosphere Watch programme of the World Meteorological Organization. The WOUDC is operated by the Experimental Studies Division of the Meteorological Service of Canada.

## Appendix B: Noise correlation structure functions and their inverses

This appendix derives the form of the process model error covariance used in the analysis equation. If the problem is discretised, the steps below would be essentially trivial, amounting to no more than assuming that the process model error covariance is diagonal under the assumed discretization. In the present work it is advantageous to delay the discretization of the problem, so the equivalent steps must be worked through in the continuous form of the problem. This does not produce any surprises, but it helps to clarify the relation between the continuous noise process in Eq. (1) and the discretised form.

The correlation structure is first approximated by a simple analytic form which can be inverted, in the sense described below. The inverse can then be used to construct the Bayesian likelihood distribution function.

The Bayesian expression for the joint likelihood function is made up of terms containing the inverses of error covariances. Here, the error covariance has been expressed as a continuous function, $f_{\text {ap }}$, rather than as a matrix. The inverse of the function may then be expressed as a differential operator with the property that:

ACPD
5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

Title Page


Back
Introduction

References

Figures
$L\left[f_{\mathrm{ap}}\right]=\cos ^{-1}(\phi) \delta\left(t-t_{0}\right) \delta\left(\lambda-\lambda_{0}\right) \delta\left(\phi-\phi_{0}\right)$.
As the spherical geometry introduces some algebraic complexity which is not relevant here, the rest of this appendix will deal with Cartesian geometry. Let
$\tau=c_{1} t$
and suppose that the covariance structure function depends only on the space-time distance between two points given by:
$r^{2}=\left(x-x_{0}\right)^{2}+\left(y-y_{0}\right)^{2}+\left(\tau-\tau_{0}\right)^{2}$.
The Bayesian formulation requires the inverse of the error covariance. In the continuous representation the inverse should be interpreted as an operator which produces a delta function when applied to the covariance function. The following three equations present families of operators and associated functions with this property.
$F_{1}=\frac{1}{4 \pi r l_{\mathrm{ap}}^{2}} \exp \left(\frac{-r}{l_{\mathrm{ap}}}\right) ;\left(1-l_{\mathrm{ap}}^{2} \nabla_{*}^{2}\right) F_{1}=\delta_{*}$,
$F_{2}=\frac{1}{8 \pi l_{\mathrm{ap}}^{3}} \exp \left(\frac{-r}{l_{\mathrm{ap}}}\right) ;\left(1-l_{\mathrm{ap}}^{2} \nabla_{*}^{2}\right)^{2} F_{2}=\delta_{*}$,
$F_{3}=\frac{l_{\mathrm{ap}}+r}{32 \pi I_{\mathrm{ap}}^{4}} \exp \left(\frac{-r}{I_{\mathrm{ap}}}\right) ;\left(1-l_{\mathrm{ap}}^{2} \nabla_{*}^{2}\right)^{3} F_{3}=\delta_{*}$,
where
$\nabla_{*}^{2}=\frac{\partial^{2}}{\partial x^{2}}+\frac{\partial^{2}}{\partial y^{2}}+\frac{\partial^{2}}{\partial \tau^{2}}$
$\delta_{*}=\delta\left(x-x_{0}\right) \delta\left(y-y_{0}\right) \delta\left(\tau-\tau_{0}\right)$,
and $l_{\mathrm{ap}}$ is a correlation length scale. In Eq. (B1a) the differential operator is a Laplacian,

ACPD
5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

Title Page

Abstract

Conclusions
Tables


Print Version but has a discontinuity in the gradient. Higher order operators can be used (Eq. B1c) to generate functions with more smoothness: here a continuous gradient at $r=0$.
$f_{\text {ap }}=\sigma_{\text {ap }}^{2} F_{2}$,
5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

Title Page

Abstract

Conclusions

Tables


Back
Introduction

References

Figures
shows that, under the assumption that $l_{\mathrm{ap}}$ is small compared to the grid spacing, the shows that, under the assumption that ${ }_{\text {ap }}$ is small compared to the grid spacing, the
value of the variance needed in the discretised equations is $\sigma_{a p}^{2}$. This is equal to the volume integral of the error covariance structure function.

## Appendix C: The binary-thinned latitude-longitude spherical grid

In order to arrive at the analysis equation which will be solved it is now assumed that
$I_{\mathrm{ap}} \ll I_{\mathrm{fd}}$,
where $I_{\mathrm{fd}}$ is the minimum spacing of the discretization grid. With this assumption the inverse error covariance can be approximated:
$C_{\mathrm{ap}}^{-1} \approx \sigma_{\mathrm{ap}}^{-2}$.
10 Note that the variance of the model error is $f_{\mathrm{ap}}(0)=\sigma_{\mathrm{ap}}^{2} /\left(8 \pi l_{\mathrm{ap}}^{3}\right)$. The above calculation

The spherical grid used here is a compromise between the convenience of a simple latitude longitude grid and the efficiency of grids with near uniform node distributions. The nodes lie on a set of equidistant latitudes, including both poles and the equator. The grid is defined by five parameters listed in Table 5.

The number of latitudes, including both poles, is $\left(N_{\text {lat0 }}-1\right) \times 2^{N_{\text {grid }}-1}+1$, and they are equi-spaced. The number of grid points around the equator is $N_{\text {eq }}=N_{\text {eq } 0} \times 2^{N_{\text {grid }}-1}$. Moving polewards from the equator, the number of nodes around a latitude circle is halved every time the node spacing falls below $\gamma_{\text {ref }}$ times the node spacing at the
equator. The thinning stops before the number of points falls below $N_{\text {lon min }}$. It can be shown that each finer grid includes all the points of the coarser grids in the hierarchy.

Grid sophistication can be measured by the ratio of the maximum to minimum node
5, 8879-8923, 2005

## Direct inversion

 ing as close as possible to the equatorial value. This gives $\gamma_{\text {mesh }} \approx 1$, but nodes do not line up meridionally, so that the grid is not suitable for finite difference applications. Hav10 ing a high value of $\gamma_{\text {mesh }}$ is inefficient, because accuracy is limited by the largest grid spacing and having more densely packed nodes at the poles adds to computational cost without contributing to improved accuracy. The indirect effects of the inhomogeneous grid spacing are generally more significant than the direct effects of having more nodes than necessary. With $\gamma_{\text {mesh }}=2$ the inhomogeneity introduced by the mesh is less than the inhomogeneity in typical satellite observations, where the ratio of along track to across track observational spacing may be a factor of 20 . This suggests that the mesh defined above will be sufficiently homogeneous for present purposes.A key factor in the implementation of the multigrid algorithm is the consistency of the transformation between mesh refinements with the discretization of the Laplacian. It is easily shown that $\int \nabla^{2} \psi d A=0$, for any field $\psi$, where the integral is taken over the surface of the sphere. Once a finite difference representation of $\nabla^{2}$ has been chosen, the representation of integrals should be chosen to respect the above constraint. This constrains the choice of weights given to each node. This may imply, for instance, using $\cos (\theta-d \theta / 2)-\cos (\theta+d \theta / 2)$ instead of $d \theta \sin \theta$. The two expressions give the same formal accuracy, but the first also satisfies the integral condition exactly. It has been found that using such weights in the mesh refinement process leads to an efficient multigrid algorithm, whereas using weights with the same formal accuracy but not satisfying the integral condition slows convergence substantially.
M. N. Juckes


EGU

## Appendix D: The relaxation algorithm

The relaxation algorithm is a standard iterative solution method for second order elliptical partial differential equations. In its simplest form, if $A$ is the discretization of the analysis equation and $S$ the source term, the iteration is defined by:
$A u^{(k)}-S=r^{(k)}$,

$$
\begin{equation*}
u^{(k+1)}=u^{(k)}-A_{\text {diag }}^{-1} r^{(k)}, \tag{D1}
\end{equation*}
$$

where $A_{\text {diag }}$ is the diagonal matrix obtained by setting the off-diagonal components of $A$ to zero. The ( $k$ ) superscript here denotes the $k$ th approximation in an iterative approach to the solution. If $A$ is diagonally dominant (that is, if every diagonal term is greater than the sum of absolute values of off diagonal terms in the corresponding row) it can be shown that $\left|r^{(k)}\right| \rightarrow 0$ as $k$ increases. Unfortunately, this condition cannot be fulfilled by any discretization of the high order smoothing term in the analysis equations being solved here.

To overcome this problem the differential equation is split into a multi-component ${ }_{5}$ system. Auxiliary variables $u_{\mathrm{a}}^{(m: k)}, m=0,1,2$, are introduced such that

$$
\left(\begin{array}{ccc}
\rho+w_{\mathrm{ap}} \mathcal{A}_{\mathrm{ap}} & 0 & -w_{\mathrm{num}} \nabla^{2}  \tag{D3}\\
-\nabla^{2} & 1 & 0 \\
0 & -\nabla^{2} & 1
\end{array}\right)\left(\begin{array}{l}
u_{\mathrm{a}}^{(0: k)} \\
u_{\mathrm{a}}^{(1: k)} \\
u_{\mathrm{a}}^{(2: k)}
\end{array}\right)=\left(\begin{array}{c}
r^{(0: k)}+s r c \\
r^{(1: k)} \\
r^{(2: k)}
\end{array}\right),
$$

where $r^{(m: k)}$ is an analysis residual and the superscript now includes both the index of the auxilliary variables and of the iteration number. $u_{\mathrm{a}}^{(0: k)} \equiv u^{(k)}$. The source term is

20 and the observation density is $\rho=w_{\text {obs }} \sum_{i} \delta\left(x-x_{i}^{(\mathrm{obs})}\right) \delta\left(y-y_{i}^{(\mathrm{obs})}\right)$.

Interactive Discussion

EGU

## Print Version

## Direct inversion

M. N. Juckes

Title Page

Abstract

Conclusions

Tables
Introduction

Figures


Full Screen / Esc

The aim is to construct an iteration such that $r^{(.: k)} \rightarrow 0$ as $k \rightarrow \infty$ and hence $u_{\mathrm{a}}^{(0: k)} \rightarrow u_{\mathrm{a}}$ and $u_{\mathrm{a}}^{(1: k)} \rightarrow \nabla^{2} u_{\mathrm{a}}$. This is done by defining an adjustment $a^{(m: k)}$ by solving an approximate form of Eq. (D3) obtained by neglecting all non-local coefficients. This leaves a $3 \times 3$ matrix equation at each mesh node. Let
$A_{\text {block-diag }} \equiv\left(\begin{array}{ccc}\rho_{0}+w_{\text {ap }} \gamma_{\text {ap }} & 0 & -w_{\text {num }} \gamma \\ \gamma & 1 & 0 \\ 0 & \gamma & 1\end{array}\right)$,
where $\rho_{0}$ is the discretised observation density and $\gamma=\frac{2}{c_{1}^{2 d t}}+\frac{2}{d x^{2}}+\frac{2}{d y^{2}}$ for a Cartesian grid with spacing $d x$ and $d y$ in the $x$ and $y$ directions respectively. This $3 \times 3$ matrix can easily be inverted, and the solution estimate is then updated as follows:
$u_{\mathrm{a}}^{(m: k+1)}=u_{\mathrm{a}}^{(m: k)}-\sum_{n=0,2}\left(A_{\text {block-diag }}^{-1}\right)_{m n} r^{(n: k)}$.

Appendix E: Multigridding
A multigridding algorithm has been employed to speed up the relaxation algorithm. This involves a hierarchy of grids with successively reduced resolution. This approach is motivated by the fact that the relaxation algorithm can handle small scale structures very efficiently but is slow to adjust large scale structures. The lower resolution grids provide information about the large scale structures which can be used to speed up convergence towards a solution on the finest grid. Let
$u_{\mathrm{a}}^{(: k: g)} \equiv\left(u_{\mathrm{a}}^{(0: k: g)}, u_{\mathrm{a}}^{(1: k: g)}, u_{\mathrm{a}}^{(2: k: g)}\right)$
be an approximation to $u_{\mathrm{a}}$ and associated ancillary variables, where $k$ is an iteration number and $g$ is a grid number, with $g=0$ corresponding to the highest resolution grid.

The analysis operator must then be available in a form which can be applied to any of

ACPD
5, 8879-8923, 2005

Direct inversion
M. N. Juckes

Title Page

Abstract

Conclusions

Tables


Back
Introduction
References

Figures

Print Version

Interactive Discussion

EGU the grids. If this operator is defined in differential form this does not pose any problems.

We define an analysis residual on the analysis grid:
$r_{\mathrm{a}}^{(: k: 0)}=\mathcal{A}\left[u_{\mathrm{a}}^{(: k: 0)}\right]$
and introduce a smoothing operator to transfer the residual fields onto the coarser grids:
${ }_{5} \quad r_{\mathrm{a}}^{(: k: g)}=\mathcal{S}\left[r_{\mathrm{a}}^{(: k: g-1)}\right], \quad g=1, \ldots g_{\max }$.
The relaxation algorithm can now be applied to generate an adjustment on the coarsest grid:
$a^{\left(: k: g_{\max }\right)}=A_{\text {block-diag }}^{-1}\left[r^{\left(: k: g_{\max }\right)}\right]$.
We then introduce an interpolation operator $U$ which moves fields from coarse to finer grids and define a preliminary adjustment:
$a_{*}^{(m: k: g)}=U\left[a^{(m: k: g+1)}\right]$,
The adjustment on finer grids is then defined by
$a^{(m: k: g)}=\mathcal{R}\left[r^{(m: k: g)}-\mathcal{A}\left[a_{*}^{(m: k: g)}\right]\right]+a_{*}^{(m: k: g)}$.
Acknowledgements. This work was funded by NERC through BADC. I am grateful for all the 5 data used in this study, which was of high quality and provided freely from the sources cited. I would like to thank A. Dhudhia and C. Piccolo for help with the MIPAS data.

ACPD
5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

Title Page

Abstract
Conclusions
Tables
4

Back
Close
Full Screen / Esc

Print Version

Bennett, A. F.: Inverse methods in physical oceanography, Cambridge University Press, Cambridge, UK, 1992. 8885

Bennett, A. F., Chua, B. S., Harrison, D. E., and McPhaden, M. J.: Generalized inversion of tropical-atmosphere-ocean (TAO) data and a coupled model of the tropical pacific, J. Climate, 11, 1768-1792, 1998. 8881
Bloom, S. C., Takacs, L. L., Silva, A. M. D., and Ledvina, D.: Data Assimilation Using Incremental Analysis Update, Mon. Weather Rev., 124, 1256-1271, 1996. 8881
Brühl, C., Drayson, S. R., II, J. M. R., Crutzen, P. J., Mclnerney, J. M., Purcell, P. N., Claude, H., Gernandt, H., McGee, T. J., McDermid, I. S., and Gunson, M. R.: Halogen Occultation Experiment ozone channel validation, J. Geophys. Res., 101, 10217-10240, 1996. 8894
Chipperfield, M., Khattatov, B., and Lary, D. J.: Sequential assimilation of stratospheric chemical observations in a three-dimensional model, J. Geophys. Res., 107(D21), 4585, doi:10.1029/2002JD002110, 2002. 8881
Cohn, S. E., Sivakumaran, N. S., and Todling, R.: A fixed-lag Kalman smoother for retrospective data assimilation, Mon. Wea. Rev., 122, 2838-2867, 1994. 8881
Courtier, P.: Dual formulation of four-dimensional variational assimilation, Quart. J. Roy. Met. Soc., 123, 2449-2462, 1997. 8881
Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F., and Fisher, M.: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation, Quart. J. Roy. Met. Soc., 124, 1783-1807, 1998. 8881
Danilin, M. Y., Ko, M. K. W., Bevilacqua, R. M., Lyjak, L. V., Froidevaux, L., Santee, M. L., Zawodny, J. M., Hoppel, K. W., Richard, E. C., Spackman, J. R., Weinstock, E. M., Herman, R. L., McKinney, K. A., Wennrg, P. O., Eisele, F. L., Stimpfle, R. M., Scott, C. J., Fahey, D. W., Elkins, J. W., Webster, C. R., Wilson, J. C., Perkins, K. K., and Bui, T. V.: Comparison of ER2 Aircraft and POAM-III, MLS, and SAGE-II Measurements during SOLVE Using Traditional Correlative Analysis and Trajectory Hunting Technique, J. Geophys. Res., 107(D20), 4420, doi:10.1029/2001JD002012, 2002. 8895
Dragani, R., Redaelli, G., Visconti, G., Mariotti, A., Rudakov, V., MacKenzie, A. R., and Stefanutti, L.: High resolution stratospheric tracer fields reconstructed with lagrangian techniques: a comparative analysis of predictive skill, J. Atmos. Sci., 59, 1943-1958, 2002. 8882
30 Eskes, H., van Velthoven, P. F. J., Valks, P. J. M., and Kelder, H.: Assimilation of GOME total ozone satellite observations in a three dimensional tracer transport model, Quart. J. Roy. Met. Soc., 129, 1663-1681, 2003. 8881
Jeuken, A. B. M., Eskes, H. J., van Velthoven, P. F. J., Kelder, H. M., and Hólm, E. V.: Assimila-

# Direct inversion 

M. N. Juckes

Title Page

| Abstract | Introduction |
| :---: | :---: |
| Conclusions | References |
| Tables | Figures |



Print Version
Interactive Discussion

EGU
tion of total ozone satellite measurements in a three-dimensional tracer transport model, J. Geophys. Res., 104(D5), 5551-5564, 1999. 8881
Khattatov, B. V., Lamarque, J.-F., Lyjak, L. V., Menard, R., Levelt, P. F., Tie, X. X., Gille, J. C., and Brasseur, G. P.: Assimilation of satellite observations of chemical species in global chemistry-transport models, J. Geophys. Res., 105(D23), 29 135-29 144, doi:10.1029/2000JD900466, 2000. 8881
Lorenc, A. C.: Development of an Operational Variational Assimilation Scheme, J. Met. Soc. Japan, 75, 339-346, 1996. 8881

Lumpe, J. D., Fromm, M., Hoppel, K., Bevilacqua, R. M., Randall, C. E., Browell, E. V., Grant, W. B., McGee, T., Burris, J., Twigg, L., Richard, E. C., Toon, G. C., Sen, B., Boesch, H., Fitzenberger, R., and Pfeilsticker, K.: Comparison of POAM III ozone measurements with correlative aircraft and balloon data during SOLVE, J. Geophys. Res., 108(D5), 8316, doi:10.1029/2001JD000472, (2002), 2003. 8894
Lyster, P. M., Cohn, S. E., Menard, R., Chang, L.-P., Lin, S.-J., and Olsen, R.: An Implementation of a Two Dimensional Kalman Filter for Atmospheric Chemical Consituent Assimilation on Massively Parallel Computers, Mon. Wea. Rev., 125, 1674-1686, 1997. 8881
Migliorini, S., Piccolo, C., and Rodgers, C.: Intercomparison of direct and indirect measurements: Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) versus sonde ozone profiles, J. Geophys. Res., 109(D19), D19316, doi:10.1029/2004JD004988, 2004. 8893, 8894
Pierce, R. B., Al-Saadi, J., Fairlie, T. D., Natarajan, M., Harvey, V. L., Grose, W. L., Russell, J. M., Bevilacqua, R., Eckermann, S. D., Fahey, D., Popp, P., Richard, E., Stimpfle, R., Toon, G. C., Webster, C. R., and Elkins, J.: Large-scale chemical evolution of the arctic vortex during the 1999-2000 winter: HALOE/POAM3 Lagrangian photochemical modelling for the SAGE III ozone loss and validation experiment (SOLVE) campaign, J. Geophys. Res., 108(D5), 8317, doi:10.1029/2001JD001063, (2002), 2003. 8894
Planet, W. G., Miller, A. J., Horvath, K., Nagatani, R., Flynn, L., Hilsenrath, E., Kondragunta, S., Cebula, R. P., and DeLand, M. T.: Total ozone determinations from NOAA operational SBUV/2 observations: an update, J. Geophys. Res., 106(D15), 17471-17478, 2001. 8894 Rabier, F., Thépaut, J. N., and Courtier, P.: Extended assimilation and forecast experiments with a four-dimensional variational assimilation system, Quart. J. Roy. Met. Soc., 124, 18611888, 1998. 8881
Riishøjgaard, L. P.: A direct way of specifying flow-dependent background error correlations for

# Direct inversion 

M. N. Juckes

Title Page


Print Version

Interactive Discussion
meteorological analysis systems, Tellus, 50A, 42-57, 1998. 8882
Riishøjgaard, L. P.: A method for estimating the analysis error variance in a physical space data assimilation system, Quart. J. Roy. Met. Soc., 126, 1367-1385, 2000. 8882
Riishøjgaard, L. P.: Deterministic prediction of the ensemble variance for a barotropic vorticityequation model, Quart. J. Roy. Met. Soc., 127, 1761-1773, 2001. 8882
Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific Publishing, London, 2000. 8886
Serafy, G. Y. E. and Kelder, H. M.: Near-real-time approach to assimilation of satellite-retrieved 3D ozone fields in a global model using a simplified Kalman filter, Quart. J. Roy. Met. Soc., 129, 3099-3120, 2003. 8881
Smit, H. and Kley, D.: JOSIE: The 1996 WMO International intercomparison of ozonesondes under quasi flight conditions in the environmental simulation chamber at Jülich, WMO/IGACReport, WMO Global Atmosphere Watch report series, No. 130 (Technical Document No. 926), World Meteorological Organization, Geneva, 1998. 8893, 8910
Struthers, H., Brugge, R., Lahoz, W. A., O'Neill, A., and Swinbank, R.: Assimilation of ozone profiles and total column measurements into a global general circulation model, J. Geophys. Res., 107(D20), 4438, doi:10.1029/2001JD000957, 2002. 8881
Talagrand, O.: Assimilation of Observations, an Introduction, J. Meteor. Soc. Japan, 75, 191209, 1997. 8881, 8896
Thomason, L. W. and Taha, G.: SAGE III Aerosol Extinction Measurements: Initial Results, Geophys. Res. Lett., 30, 33-1-33-4, doi:10.1029/2003GL017317, 2003. 8894
Thuburn, J. and Li, Y.: Numerical simulations of Rossby-Haurwitz waves, Tellus, 52A, 181-189, 2000. 8886, 8902

Tsidu, G. M., Kiefer, M., von Clarmann, T., Fischer, H., Funke, B., Grabowski, U., Hase, F., Höpfner, M., López-Puertas, M., and Stiller, G. P.: Validation of MIPAS/ENVISAT Level-1B data products, edited by: Huang, H.-L., Lu, D., and Sasano, Y., Optical Remote Sensing of the Atmosphere and Clouds III, Proc. of SPIE, 4891, 483-496, 2003. 8894
von Savigny, C., Haley, C. S., Sioris, C. E., McDade, I. C., Llewellyn, E. J., Degenstein, D., Evans, W. F. J., Gattinger, R. L., Griffoen, E., Lloyd, N., McConnell, J. C., McLinden, C. A., Murtagh, D. P., Solheim, B., and Strong, K.: Stratospheric Ozone Profiles Retrieved from Limb Scattered Sunlight Radiance Spectra Measured by the OSIRIS Instrument on the Odin Satellite, Geophys. Res. Lett., 30, 1755-1758, doi:10.1029/2002GL016401, 2003. 8894

# Direct inversion 

M. N. Juckes

Title Page


Print Version

## ACPD

5, 8879-8923, 2005

Table 1. Variation of the cost of the numerical algorithm with problem size, for a 50 day analysis, on one isentropic level. The columns show: the number of grid refinements used in the multigrid method ( $N_{\text {grid }}$ ); the CPU time required ( $T$ ) [seconds]; the number of spatial mesh points ( $N_{\text {mesh }}$ ); the number of control variables - the number of mesh points times the number of time levels $\left(N_{\mathrm{cv}}\right)$; the number of CPU cycles per control variable (cycles/cv); and the number of iterations on the finest grid ( $N_{\mathrm{it}}$ ).

| $N_{\text {grid }}$ | $T$ | $N_{\text {mesh }}$ | $N_{\text {cv }}$ | cycles/cv | $N_{\text {it }}$ |
| ---: | ---: | ---: | ---: | ---: | ---: |
| 3 | 62 | 890 | 257210 | $4.9 \times 10^{5}$ | 33 |
| 4 | 236 | 3482 | 1006298 | $4.7 \times 10^{5}$ | 45 |
| 5 | 961 | 13658 | 3947162 | $4.9 \times 10^{5}$ | 42 |
| 6 | 4672 | 53594 | 15488666 | $6.0 \times 10^{5}$ | 56 |
| 7 | 26051 | 213338 | 61654682 | $8.5 \times 10^{5}$ | 74 |

## Direct inversion

M. N. Juckes

Title Page
Abstract
Introduction
References
Figures


Back
Close
Full Screen / Esc

## ACPD

5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

Title Page

Abstract
Introduction

References

Figures


Back
Close

Full Screen / Esc

Print Version

Interactive Discussion

EGU

## ACPD

5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

Title Page


Print Version
Interactive Discussion

EGU

Table 4. Name, position, country and instrument type of sondes used in this study. The total number of ascents in the 6 month period presented here is 860 .

| Number | Name | \# Profiles | Latitude | Longitude | Country | Type |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 191 | Samoa | 12 | -14.22 | -170.55 | American Samoa | ECC |
| 323 | Neumayer | 40 | -70.64 | -8.25 | Antarctica | ECC |
| 450 | Davis | 10 | -68.57 | 77.97 |  | ECC |
| 328 | Ascension Island | 19 | -7.97 | -14.41 | Ascension Island | ECC |
| 29 | Macquarie Island | 28 | -54.49 | 158.95 | Australia | ECC |
| 394 | Broadmeadows | 21 | -37.67 | 144.95 |  | ECC |
| 18 | Alert | 31 | 82.5 | -62.32 | Canada | ECC |
| 21 | Stony Plain | 24 | 53.55 | -114.1 |  | ECC |
| 24 | Resolute | 15 | 74.71 | -94.96 |  | ECC |
| 76 | Goose Bay | 25 | 53.31 | -60.35 |  | ECC |
| 77 | Churchill | 16 | 58.74 | -94.06 |  | ECC |
| 315 | Eureka | 28 | 79.99 | -85.93 |  | ECC |
| 242 | Praha | 37 | 50.02 | 14.45 | Czech Republic | ECC |
| 434 | San Cristobal | 18 | -0.91 | -89.59 | Ecuador | ECC |
| 438 | Suva | 14 | -18.12 | 178.4 | Fiji | ECC |
| 99 | Hohenpeissenberg | 62 | 47.8 | 11.02 | Germany | Brewer-Mast |
| 174 | Lindenberg | 27 | 52.21 | 14.12 |  | ECC |
| 316 | De Bilt | 24 | 52.1 | 5.18 | Holland | ECC |
| 10 | New Delhi | 11 | 28.65 | 77.22 | India | Indian-sonde |
| 187 | Poona | 7 | 18.53 | 73.85 |  | Indian-sonde |
| 205 | Thivandrum | 8 | 8.48 | 76.95 |  | Indian-sonde |
| 437 | Watukosek (Java) | 26 | -7.49 | 112.6 | Indonesia | ECC |
| 336 | Isfahan | 1 | 32.51 | 51.7 | Iran | ECC |
| 318 | Valentia Observatory | 7 | 51.93 | -10.24 | Ireland | ECC |
| 7 | Kagoshima | 22 | 31.6 | 130.6 | Japan | Carbon-lodine |
| 12 | Sapporo | 24 | 43.1 | 141.3 |  | Carbon-lodine |
| 14 | Tateno (Tsukuba) | 22 | 36.1 | 140.1 |  | Carbon-lodine |
| 101 | Syowa | 37 | -68.99 | 39.6 |  | Carbon-Iodine |
| 190 | Naha | 21 | 26.2 | 127.7 |  | Carbon-lodine |
| 443 | Sepang Airport | 12 | 2.73 | 101.7 | Malaysia | ECC |
| 256 | Lauder | 20 | -45.03 | 169.68 | New Zealand | ECC |
| 89 | Ny -Ålesund | 32 | 78.93 | 11.95 | Norway | ECC |
| 221 | Legionowo | 29 | 52.4 | 20.97 | Poland | ECC |
| 265 | Irene | 12 | -25.89 | 28.22 | South Africa | ECC |
| 401 | Santa Cruz | 17 | 28.46 | -16.25 | Spain | ECC |
| 435 | Paramaribo | 25 | 5.81 | -55.2 | Suriname | ECC |
| 156 | Payerne | 76 | 46.49 | 6.57 | Switzerland | ECC |
| $8912$ |  |  |  |  |  |  |

## ACPD

5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

Title Page

## Abstract

Introduction
Conclusions
References
Tables
Figures


Back
Close

## Full Screen / Esc

## Print Version

Interactive Discussion

EGU

ACPD
5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

Table 5. Parameters describing the grid used in this study, with variable longitudinal spacing. See text for details.

| Parameter | values used | description |
| :--- | :---: | :--- |
| $N_{\text {grid }}$ | $3,4,5,6,7$ | The number of grid refine- <br> ments |
| $N_{\text {lato }}$ | 7 | The number of latitudes <br> on the coarsest grid |
| $N_{\text {eq0 }}$ | 12 | The number of nodes <br> around the equator on the <br> coarsest grid <br> Determines latitudes at <br> which node spacing is <br> doubled <br> Minimum number of longi- <br> tudinal nodes |
| $V_{\text {ref }}$ | 0.7 |  |





Fig. 1. Southern hemisphere, isentropic fields of ozone, on the 850 K surface, for 10 July 2003, derived from MIPAS observations.

ACPD
5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

Title Page


Print Version
Interactive Discussion

EGU


ACPD
5, 8879-8923, 2005


## Direct inversion

M. N. Juckes


Print Version
(a)

```
Interactive Discussion
```

Fig. 2. Southern hemisphere, isentropic fields of ozone, on the 850 K surface, for 10 July 2003. (a) MIPAS, (b) SBUV.


## ACPD

5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

| Title Page |  |
| :---: | :---: |
| Abstract | Introduction |
| Conclusions | References |
| Tables | Figures |
| $\mathbf{I 4}$ |  |
| $\mathbf{~ B a c k}$ |  |
| Full Screen / Esc |  |

## Print Version

(b)

```
Interactive Discussion
```

Fig. 2. Continued.

## ACPD

5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

Title Page

Abstract

Conclusions

Tables


Full Screen / Esc

Print Version

Fig. 3. Global root-mean-square difference between analyses from time windows 5 and 6 , on EGU

## ACPD






| Sage III | Soge ॥ $\Delta \Delta \Delta \Delta \Delta$ | $\underset{\times \times \times \times x}{\text { Poam } I I I}$ | Osiris |
| :---: | :---: | :---: | :---: |
| Mipos | Soge III（L） | Holoe | SBUV |

WOUDC Ozonesondes（smoothed）
वロロロロ
WOUDC Ozonesondes

5，8879－8923， 2005

## Direct inversion

M．N．Juckes

Title Page

## Abstract

Introduction
Conclusions
References
Tables
Figures


Print Version

Fig．4．Validation data，showing differences between independent observations and the gridded
EGU fields．


(b)

| WOUDC Ozonesondes (smoothed) |  |
| :---: | :---: |
| WOUDC Ozonesondes |  |
| Osiris | Haloe |
| Sage III | Poam III |
| Sage III(L) | SBUV |
| Sage II |  |

## ACPD

5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

Title Page


Full Screen / Esc

Print Version
Interactive Discussion

EGU

Fig. 5. (a) Mean residual, (b) root-mean-square residuals, February to July 2003. the mean value of the verifying observation minus the MIPAS assimilation. The shading in the background shows the $10 \%, 5 \%$ and $1 \%$ lines.

## ACPD

5, 8879-8923, 2005

(b)

Fig. 6. As Fig. 5, except using nearly coincident profiles, with co-location criteria of spacing less than $2.5^{\circ}$ and 6 h .

## Direct inversion

M. N. Juckes

Title Page

Abstract

Conclusions
References

Figures


Full Screen / Esc

Print Version

EGU

## ACPD

5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

Title Page

Abstract

Conclusions

Tables


Full Screen / Esc

Print Version

EGU

ACPD
5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

Title Page

Abstract

Conclusions
References

Figures


Print Version

EGU

ACPD

5, 8879-8923, 2005

## Direct inversion

M. N. Juckes

Title Page

Abstract
Introduction
Conclusions
References

Tables
Figures


Full Screen / Esc

## Print Version

Fig. 9. As for Fig. 8, but varying $w_{\text {ap }}$.

Interactive Discussion

EGU

