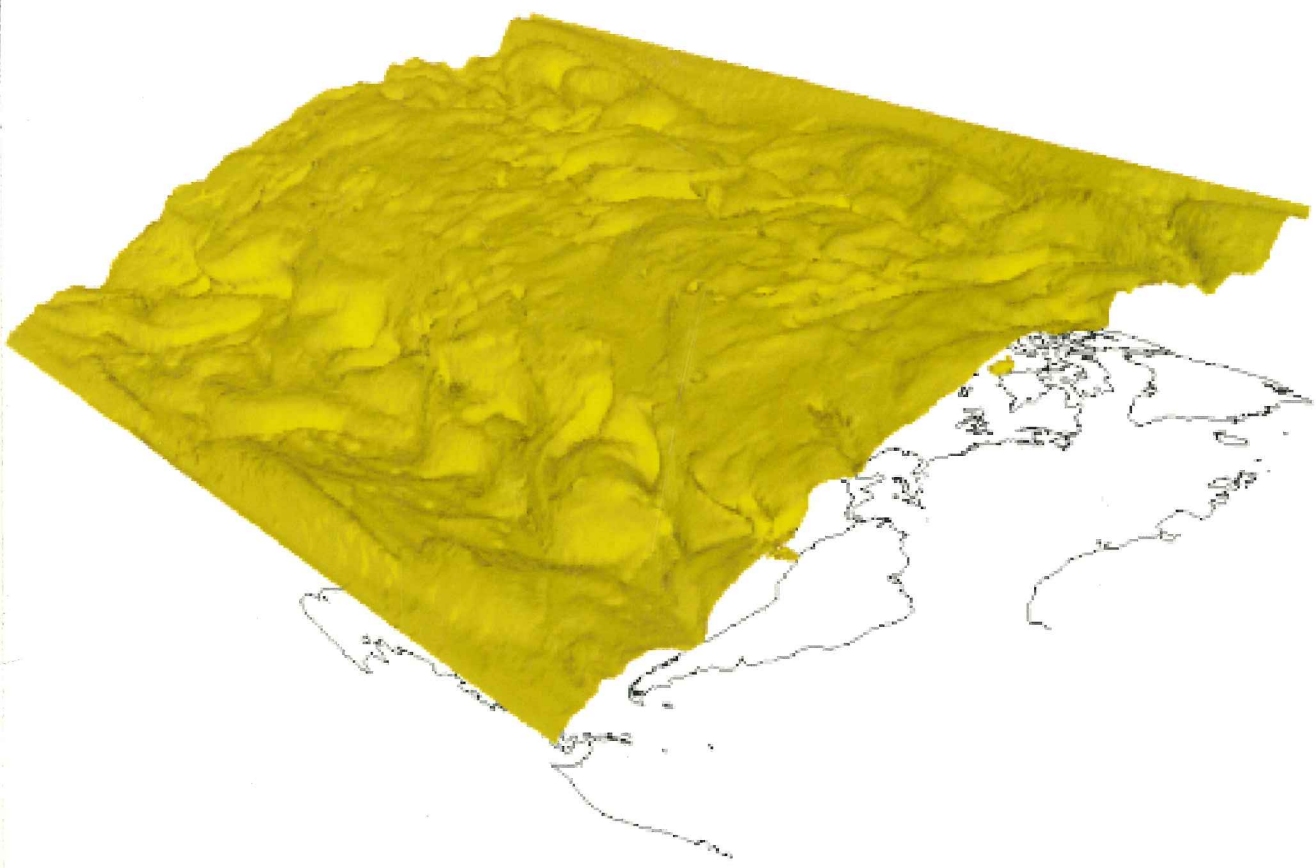
KNMI-publikatie 188

KNMI

SODA Workshop on Chemical Data Assimilation Proceedings

9 - 10 December 1998

KNMI, De Bilt, the Netherlands



De Bilt februari 1999

Koninklijk Nederlands Meteorologisch Instituut

SODA Workshop on Chemical Data Assimilation Proceedings

9 - 10 December 1998
KNMI, De Bilt, the Netherlands

Satellite Ozone Data Assimilation (SODA)

European Union project, 1996 - 1999

4th Framework program

Theme 3: Space techniques applid to environmental monitoring and research

Area 3.1: Methodological research and pilot projects

3.1.1: Methodological research

Cover page:

The figure shows an isosurface of ozone mass mixing ratio (2.5E-7), located in the lower stratosphere, from the sixth day of 4D-VAR assimilation of TOVS total ozone data in the ECMWF weather forecast model (25/5 1998, 15h). The undulation of the ozone closely follows the changes of the windfield.

Organising committee:

Henk Eskes
Hennie Kelder
Ad Stoffelen
Fons Baede

Contact:

Ad Stoffelen
WM/SD
KNMI
de Bilt, the Netherlands
E-mail: Stoffelen@knmi.nl
tel: +31 30 22 06 585
fax: +31 30 22 10 407

The workshop on *Chemical Data Assimilation* was financially supported by the KNMI and the European Union

KNMI publication series, no. 188

Contents

Introduction to chemical data assimilation

Without data assimilation techniques advanced weather forecasting as we know it today would be unthinkable. Data assimilation analyses observations of different origin, taken at different places and times, of several quantities such as temperature, pressure and wind, and combines all the information into a quality-controlled unified description of the dynamics of the atmosphere.

In the field of atmospheric chemistry the use of data assimilation is new. To our opinion this is surprising, since the well-established benefits of data assimilation apply equally well to this field. The technique transforms a host of observations into a global synoptic description of the chemical atmosphere in a way consistent with our knowledge of the chemical reactions, the atmospheric dynamics and radiation balance. Through these equations, defining an atmospheric model, the assimilation of a selection of observed chemicals will lead to an improved understanding of unobserved species that react with the observed ones. Treating chemical species as completely independent ignores the balance between the chemical constituents and the sensitivity to temperature and sunlight, i.e. the diurnal cycle. Data assimilation does not impose such constraints and makes full use of the often indirect information stored in the observations.

Ozone is a key species of the atmosphere. Record low global ozone levels have been observed in the last few years with downward trends over much of the globe. Data assimilation can provide a valuable consistent ozone data set and helps to determine the relative roles of chemistry and atmospheric dynamics in the observed changes in ozone. The aim of the European Union Satellite Ozone Data Assimilation (SODA) project is to develop the technique of data assimilation to generate a data base of ozone, consistent with the dynamics of the atmosphere. Such data bases are of considerable value for the documentation of the climatology of ozone, as a priori data for the interpretation of the chemical data from research satellites, and for the development of policy. Chemical observations from several dedicated space-borne instruments such as the ENVISAT satellite and the Earth Observation System (EOS) will become available in the near future. An ozone data assimilation procedure provides a wealth of statistics on the quality of the ozone observations. This allows to identify problems with the instruments and data algorithms, and help set the user requirements for future satellite instruments. A high quality ozone forecast will also serve as a basis for reliable ultraviolet (UV) sunlight forecasts, an important issue for human health protection.

Numerical weather prediction models will also benefit from the assimilation of ozone. A detailed knowledge of the space and time dependence of the ozone density improves the description of radiation in the models. In turn, this results in a more accurate assimilation of top-of-the-atmosphere radiances from satellite observations, and, by consequence, in more reliable temperature and humidity information. The variations in ozone are mainly a result of the dynamics of the atmosphere. As a consequence the assimilation of accurate ozone observations using 4D-Var will lead to adjustments of the wind field in the lower stratosphere in particular. This indirect impact of ozone observations may prove to be relevant for the description of the higher atmospheric dynamics in general-circulation models.

The aim of this workshop was to highlight present-day developments in the field of chemical data assimilation. This includes the retrieval of satellite observations, chemical and dynamical aspects of the atmosphere, and the analysis of ozone and other trace gases in numerical weather prediction models and chemistry-transport models.

Workshop summary and recommendations

The workshop reviewed several aspects of chemical data assimilation, and the program was divided in five sessions: the observation system for atmospheric composition measurements, atmospheric chemistry and its parameterisation, the ozone structure generated by the atmosphere's dynamics, ozone data assimilation, and advanced chemical data assimilation systems. The summary and outlook below are following these themes.

- Observations

A multitude of observations on atmospheric composition is being made available to the research community. It takes a large effort to analyse these measurements in a consistent way. A consistent analysis is however necessary in order to better understand the atmospheric dynamical and chemical processes underlying these data. Data assimilation is the appropriate tool to integrate heterogeneous measurements into a consistent atmospheric state. The atmospheric model state, constrained by realistic dynamical and chemical equations, was shown to be a good reference to detect problems in the interpretation of the observations. On the other hand, atmospheric composition measurements provide feedback on the adequacy of these model equations. It is important that the measured properties are simulated as closely as possible in data assimilation. Spectral and spatial representations (effective averaging kernels) of the measurements are needed, rather than level 2 products with implicit or explicit assumptions on atmospheric properties (e.g. cloud).

Spectrally resolved observations, such as from the GOME instrument, usually have a time-consuming processing. This has detrimental effects on the usability of the full data set for a wide range of applications, certainly when errors in the processing arise and a reprocessing has to be considered. At the workshop some promising attempts for a faster processing of such data were presented. Obviously, processing speed is compromised with retrieval accuracy.

- Dynamics

  At the workshop several presentations discussed atmospheric dynamics in relation to ozone transport. High horizontal and vertical resolution, extending well into the stratosphere, is necessary for a detailed description of ozone structures. Atmospheric dynamics determine the spatial structure in ozone to a very high degree. Ozone structures generated by the models are filamentary and essentially anisotropic and seem realistic. Deficiencies that were found in the modelled ozone structure can be due to problems in the model advection or the chemistry parameterisation. Comparison to ozone measurements will be needed to further investigate the exact cause.

- Chemistry

  In this session the chemical equations and parameterisations thereof were discussed. A parameterisation of ozone sources and sinks is needed for global circulation models to guarantee a reasonable ozone distribution on the longer term. The Cariolle scheme prevents ozone drifts, but turns out to be model dependent. Validation by observations has shown that the tuning of the parameterisation needs further development. In particular, the interaction of the fast relaxation times, the model temperature and the specific ozone climatology in the upper stratosphere needs attention. A next step is to include heterogeneous chemistry in the models. More elaborate chemical equations have been validated for campaign periods and turn out to be consistent with observations.

- Ozone data assimilation

  Several centres reported on their efforts to assimilate ozone in global circulation models. Studies on measured and background ozone error covariances indicate that model background ozone values are generally of good quality. As a result, ozone measurements have a modest impact on the ozone analysis, while assimilated ozone fields provide a good analysis of the 3D ozone distribution. Experimentation has started with heuristically-defined anisotropic structure functions. Given the filamentary nature of the ozone fields, more effort in this direction may be rewarding. Data assimilation results in the correction of some systematic model deficiencies. However, systematic errors in the modelled ozone structure are an obstacle for the effective extraction of dynamical information from ozone

measurements. As such, it remains a challenge to extract relevant dynamical information from ozone measurements.

- Chemical data assimilation

  Chemical data assimilation can be of great complexity due to the degrees of freedom provided in the chemical reaction equations. Ways to constrain these degrees of freedom were reported during the workshop. Tropospheric data assimilation was described on a limited area, and several data assimilation methodologies were discussed. Insufficient measurements are available to analyse the full heterogeneous chemistry on a global scale. However, chemical parameterisations may be extended to better resolve the interaction of chemistry and dynamics. In other cases, the degrees of freedom in the chemical-dynamical equations may become smaller than the number of constituent variables. Then composition measurements may reveal problems in the chemical equations or transport, offering a way to improve the models. Moreover, measurements could be calibrated in this way.

To conclude, the papers in this proceedings show that the assimilation of chemical observations is a rapidly growing field of research. We anticipate that chemical data assimilation will play a key role in the analysis of chemical information from satellite instruments in the coming years.

We would like to thank all participants for their stimulating contributions to a successful workshop.

Ad Stoffelen, Henk Eskes and Hennie Kelder

Workshop programme

Wednesday December 9.

Convener: Hennie Kelder.

9:00 Opening by Fons Baede

Observations

9:10 Chris Readings, *The chemical measurements of Envisat.*

9:40 Richard Siddans, *Height-resolved ozone retrievals in the lower atmosphere from GOME.*

10:20 Roeland van Oss, *Ozone profile retrieval at KNMI.*

11:10 Slimane Bekki, *Simple methods for validating satellite observations of chemical species against ground-based and balloon-borne measurements.*

11:50 Fernand Karcher, *Fast radiative transfer model for the TOVS; Winds vectors from total ozone images derived from geostationary IR sounders.*

12:10 Renske Timmermans, *Validation of model ozone calculations using ozone profile measurements.*

Chemistry

12:20 Pascal Simon, *The simplified chemistry including heterogeneous ozone destruction processes.*

14:00 Jack McConnell, *Assimilation of ODIN ozone data in the CMC weather forecast model with chemistry: chemical modelling.*

Dynamics

14:20 Adrian Simmons, *Studies of increased stratospheric resolution at ECMWF.*

14:50 Pierre Gauthier, *The impact of vertical transport on ozone analyses based on total ozone measurements.*

Data assimilation - ozone

16:00 Lars Peter Riishøjgaard, *The GEOS data assimilation system for ozone observations.*

16:40 Simon Grainger, *Some results from the statistical interpolation of satellite ozone data.*

Thursday December 10.

Convener: Ad Stoffelen

Data assimilation - ozone (continued)

9:00 Penny Connew, *Chemical data assimilation using the UKMO Unified Model.*

9:40 Henk Eskes, *Assimilation of total ozone with the TM3 model.*

10:00 Aline Peuch, *An OSSE using TOVS total ozone data.*

10:20 Elias Hólm, *Multivariate ozone assimilation in 4D-VAR.*

11:20 Alan O'Neill, *The CGAM project to assimilate temperature and ozone data from research satellites by using a troposphere-stratosphere GCM.*

Chemical data assimilation

12:00 David Lary, *Chemical data assimilation.*

14:00 Boris Khattatov, *Assimilation of global tropospheric CO observations.*

14:40 Hendrik Elbern, *Four-dimensional chemical data assimilation for tropospheric ozone analysis.*

15:20 Jack McConnell, *Assimilation of ODIN ozone data in the CMC weather forecast model with chemistry: a feasibility study.*

16:10 Saroja Polavarapu, *A computationally efficient extended Kalman filter for chemistry-transport models.*

16:50 Arnold Heemink, *Data assimilation algorithms for large scale transport models.*

17:30 Workshop summary by Ad Stoffelen

Observations

THE CHEMICAL MEASUREMENTS ON ENVISAT

C.J. Readings and J. Langen

The Earth Sciences Division,
ESTEC, Noordwijk,
The Netherlands
Phone: +31 71 565 5673, Fax: +31 71 565 5675, e-mail: creading@estec.esa.nl

1. Introduction

ENVISAT is an advanced Earth Observing satellite, due to be launched in the year 2000, which has been designed to observe the atmosphere, the oceans, the land and the cryosphere over a five year period. The package of instruments to be flown on this satellite (see Figure 1) will provide an almost unique opportunity for the synergetic use of satellite data.

Three new atmospheric sounding instruments, designed primarily for atmospheric chemistry, including the observation of ozone, are included in the payload of ENVISAT. These comprise:

- GOMOS - Global Ozone Monitoring by Occultation of Stars
- MIPAS - Michelson Interferometer for Passive Atmospheric Sounding
- SCIAMACHY - SCanning Imaging Absorption spectrometer for AtMospheric CHartographY

Together these instruments exploit not only the ultraviolet and visible parts of the spectrum, but also the near and middle infrared (Figure 2), to observe trace gases. Figure 3 indicates some of the atmospheric constituents that should be observed by these three instruments.

ENVISAT will fly in a polar orbit at a mean altitude of about 800 kms with a repeat cycle of 35 days. The local equator crossing time in the descent mode will be ten o'clock in the morning. Flight operations will be controlled from ESOC, Darmstadt, Germany. As usual Esrin, Frascatti, Italy, will provide the interface for the users. More information will be found in *Envisat Mission: Opportunities for Science and Applications*, ESA SP-1218 and in <http://envisat.estec.esa.nl>; the ENVISAT web site. A full listing of all the products that it is planned to make available shortly after the end of the commissioning phase will be found in *Envisat Mission: Product Summary Overview*, ESA SP-1221.

The object of this paper is to outline the capabilities of these instruments and to demonstrate the relevance of ENVISAT to atmospheric chemistry. This reflects the increasing concerns over the impact of human activities on the environment as it is clear that mankind's activities are not only causing damage to the environment itself but that they may be associated with long term climatic changes. These concerns are reflected in the list of priority issues identified by the Inter-Governmental Panel on Climate Research (IPCC) which includes topics such as the sources, sinks and concentrations of greenhouse gases, the Earth's radiation balance, ecosystem dynamics and the role of aerosols.

In partial response to this the Member States of the European Space Agency (ESA) have not only approved the ENVISAT mission but, in addition, they have supported the formulation of long term strategy for Earth Observation which envisages a series of research and/or demonstration missions called the Earth Explorer Missions. More information on these missions, including their research objectives, will be found in *Earth Explorers: The Science and Research Elements of ESA's Living Planet Programme*, ESA SP-1227. *The Living Planet Programme* includes mission preparation and data exploitation as well as the missions themselves.

2. GOMOS

GOMOS is a limb viewing spectrometer operating in the ultraviolet/visible regions of the spectrum which tracks stars as they sink through the atmosphere. Its main objectives are to contribute to the monitoring of stratospheric ozone and to further understanding of the chemistry and dynamics of the stratosphere. In addition to ozone itself it will observe several associated trace species (including NO_2 and NO_3) plus aerosols, temperature and water vapour. In each case it will provide concentration profiles.

A key requirement associated with its monitoring role will be the maintenance of the accuracy of the GOMOS observations over the five years of the ENVISAT

mission. This must be coupled with good geographic coverage and a vertical resolution of about 2 km. These are an exacting set of requirements as, although instruments like SAGE (Solar Absorption Gas Experiment), which exploit solar occultation, can provide observations to the requisite accuracy, their geographic coverage is poor.

GOMOS seeks to strike a better compromise between these two requirements by exploiting an inherently self-calibrating technique, namely stellar occultation (Figure 4). From its polar orbit GOMOS will observe stars whose lines of sight are tangential to the Earth's limb. For each individual star the spectrum observed outside the atmosphere will be compared with the spectrum seen through the atmosphere as the star sinks through the horizon. The changes reflect the presence of ozone and other trace gases.

The high accuracy will be assured by the use of the occultation technique which has been established by SAGE for solar occultation. The benefits of stellar occultation as far as geographic coverage are concerned are illustrated by Figure 5. GOMOS covers the spectral region of 250 nm to 750 nm, which includes both the Huggins and Chappuis bands, at 1 nm resolution so for ozone accuracies of the order of a few percent should be possible Temperature and water vapour will be derived from observations made at around 760 nm and 930 nm. GOMOS will also be capable of observing atmospheric scintillation with the aid of two fast photometers. Table 1 summarises the main technical features of GOMOS.

3. MIPAS

MIPAS is a high resolution limb viewing Fourier transform interferometer operating in the infrared which has been designed to measure the concentration profiles of various atmospheric constituents on a global scale. It will observe atmospheric emissions from the Earth's horizon in the mid-infrared region providing global observations of photochemically inter-related trace gases in the stratosphere, the upper troposphere and in the middle atmosphere.

Specifically it is intended to provide data required to further understanding of stratospheric chemistry, polar winter chemistry, the chemistry of the upper atmosphere, exchange between the stratosphere and the troposphere, upper tropospheric chemistry, diurnal changes and atmospheric dynamics. It should also contribute to work on non-LTE (local thermal equilibrium) processes.

It will be possible to use MIPAS to observe about twenty trace gases including the complete family of nitrogen oxides (NO_y) and several chlorofluorocarbons (CFCs). In addition it will be possible to derive atmospheric temperature, as well as aerosol and ice cloud distributions, from its observations. Table 2 provides a full list of the variables to be observed by MIPAS.

MIPAS has a very flexible viewing geometry as the instrument will be able to view both along and across track (Figure 6). This means that all parts of the globe will be accessible to it. Furthermore, its observations will be independent of illumination conditions so MIPAS will couple complete global coverage with the ability to view day and night in all seasons. The vertical resolution of the observations will vary between 2 km and 8 km.

At the heart of the instrument is a dual port Michelson interferometer. Spectral coverage extends from about 4 μm to 15 μm, well into the mid-infrared where the spectral signatures of the NO_y and the CFCs are to be found. A spectral resolution of 0.035 cm^{-1} will be achieved throughout this range. A typical elevation scan will last about 75 seconds including 16 interferometer sweeps. Table 3 summarises the main technical characteristics of MIPAS.

4. SCIAMACHY

SCIAMACHY is a limb/nadir viewing spectrometer operating in the ultraviolet, visible and near infrared regions of the spectrum. Its primary scientific objective is the global observation of various trace gases in the troposphere and the stratosphere. Its data should advance understanding of a wide range of atmospheric phenomena which influence the chemistry of the atmosphere including, in the troposphere, biomass burning, industrial pollution, arctic haze, forest fires, dust storms etc. In the stratosphere these include ozone chemistry and volcanic events.

As SCIAMACHY can view in limb as well as in nadir (Figure 7) it will be possible to derive concentration profiles as well as column amounts from its observations. Its inherent flexibility is further enhanced by its ability to track both the Sun and the Moon which, as well as providing very important calibration information, makes high accuracy occultation measurements possible of ozone and other trace gases. Its data should make important contributions to the monitoring of trends in ozone amounts and distribution.

The wide part of the spectrum covered by SCIAMACHY coupled with its high spectral resolution means that it will be able to observe a large number of trace gases plus ice clouds and aerosol. This is illustrated in Figure 8 which compares the spectral coverage of the GOME (Global ozone measuring experiment), which is flying on the ERS-2 satellite, with that of SCIAMACHY. GOME is a derivative of

SCIAMACHY which only operates in the ultraviolet and visible regions of the spectrum and which can only view in nadir. The figure not only identifies many of the species that can be observed but also the relevant spectral regions.

One of the major strengths of both SCIAMACHY and GOME is the combination of wide, continuous spectral coverage with high spectral accuracy means that new powerful algorithms can be used to process the data. This increases both the number of species that can be derived (relative to SBUV or SAGE) and the accuracy of the derived products. The main technical features of SCIAMACHY are summarised in Table 4.

5. Concluding Remarks

The data from these three instruments are stored on the tape recorders on-board ENVISAT and relayed to ground either directly via Ka Band to Esrin in Italy or via a relay satellite to Kiruna, Sweden. This means that in principle all the data from the three instruments could be made available in near-real time i.e. within 3 hours of observation. However, this also depends on the ground system and, generally, only a sub-set of the products listed in *Envisat Mission: Product Summary Overview*, ESA SP-1221, will be available on this time scale.

It is also important to note that products available in near real time will not always be of the same quality as those distributed later. This reflects mainly the availability of better information (orbit, meteorology etc.) and the use of better calibration parameters. Also slightly more sophisticated algorithms are used for off-line processing. One important topic is re-processing as none of the three instruments has flown in space before so all the algorithms will require "tuning".

These products will provide many opportunities for synergetic studies a point well illustrated by the responses to the recent ENVISAT Announcement of Opportunity. 80% of the SCIAMACHY proposals also requested GOMOS data and vice versa; 70% of the atmospheric proposals involve the use of data from all three atmospheric chemistry instruments i.e. GOMOS, MIPAS and SCIAMACHY (85% of the MIPAS proposers also request GOMOS data).

Optical Performance Parameters	UV/VIS Channels – 250 to 675 nm at 1.2 nm resolution IR 1 Channel – 756 to 773 nm at 0.2 nm resolution IR 2 Channel – 926 to 952 nm at 0.2 nm resolution Photometer 1 – 650 to 700 nm (broadband) Photometer 2 – 470 to 520 nm (broadband)
Altitude Range	20 to 100 km
Vertical Resolution	1.7 km
Operation Cycle	Continuous over whole orbit
Data Rate	222 kb/s
Mass	163 kg
Power	146 W

Table 1: *The GOMOS Instrument Parameters*

Class of Observable	Parameters
Ozone	O_3
Temperature	CO_2
Source Gases	H_2O, CO, CH_4, N_2O, CFC11, CFC12, CFC22, CCl_4, CF_4, SF_6, C_2H_2, C_2H_6, HDO
Radicals	NO, NO_2, ClO
Reservoir and Sink Species	N_2O_5, $ClONO_2$, H_2O_2, HNO_4, HOCl, HNO_3
Other Trace Gases	COF_2, OCS, HCN, NH_3
Other Parameters	Aerosols, PSCs (polar stratospheric clouds), Cirrus

Table 2: *The parameters to be observed by MIPAS*

Instrument $NESR_0$	Varies between 50 nW cm^2 sr^{-1}/ cm^{-1} at 685 cm^{-1} to 4.2 nW cm^2 sr^{-1}/ cm^{-1} at 2410 cm^{-1}
Radiometric Accuracy	2x$NESR_0$ + 2% to 5% of source radiance (depending on wavelength)
Spectral Coverage	685 cm^{-1} to 2410 cm^{-1}
Spectral Resolution	< 0.035 cm^{-1}
Spectral Stability	<0.001 cm^{-1} (over 1 day)
Elevation Scan Angle	Between 5 km and 150 km (tangent height)
Azimuth Scan Range	80° to 110° and 160° to 195° (wrt flight direction)
Line of Sight Pointing	<1.8 km in tangent height
Operation	Continuously over the full orbit
Data Rate	533 kb/sec; for raw data 8 Mb/sec
Mass	320 kg
Power	195 W

Table 3: *The MIPAS Instrument Parameters*

	Channel	**Spectral Range**	Spectral Resolution
High Resolution Channels	**1**	240-314 nm	
	2	309-405 nm	0.24nm
	3	394-620 nm	0.26 nm
	4	604-805 nm	0.44 nm
	5	785-1050 nm	0.48 nm
	6	1000-1750 nm	0.54 nm
	7	1940-2040 nm	1.48 nm
	8	2265-2380 nm	0.22 nm 0.26 nm
Polarisation Measurement Devices (broadband)	**PMD 1 to 7**	310-2380 nm	67 to 137 nm (channel dependent)
Altitude Range		0 km to 100 km (varies with measurement mode)	
Vertical resolution		2.4 km to 3 km in limb mode (varies with measurement mode)	
Operation		Continuously over the full orbit	
Data Rate		400 kb/sec nominal: 1867 real time mode	
Mass		198 kg	
Power		122 W	

Table 4: *The SCIAMACHY Instrument Parameters*

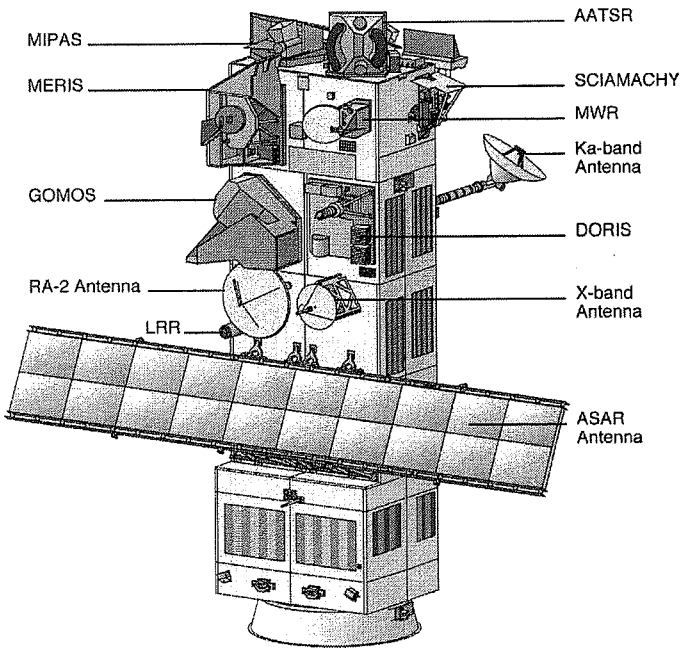


Figure 1: *ENVISAT and its payload*

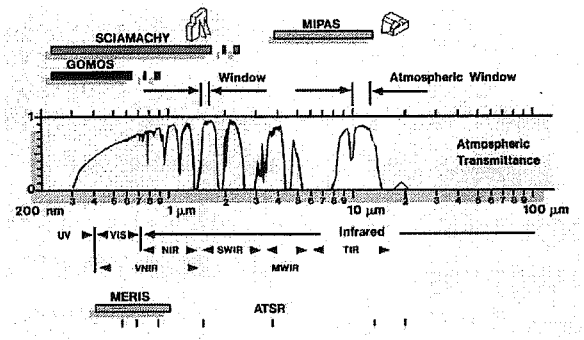


Figure 2: *The spectral coverage of the ENVISAT atmospheric sensors*

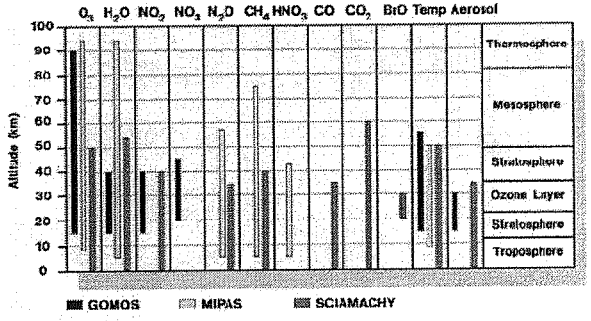


Figure 3: *The altitude ranges over which some atmospheric constituents are measured by GOMOS, MIPAS and SCIAMACHY*

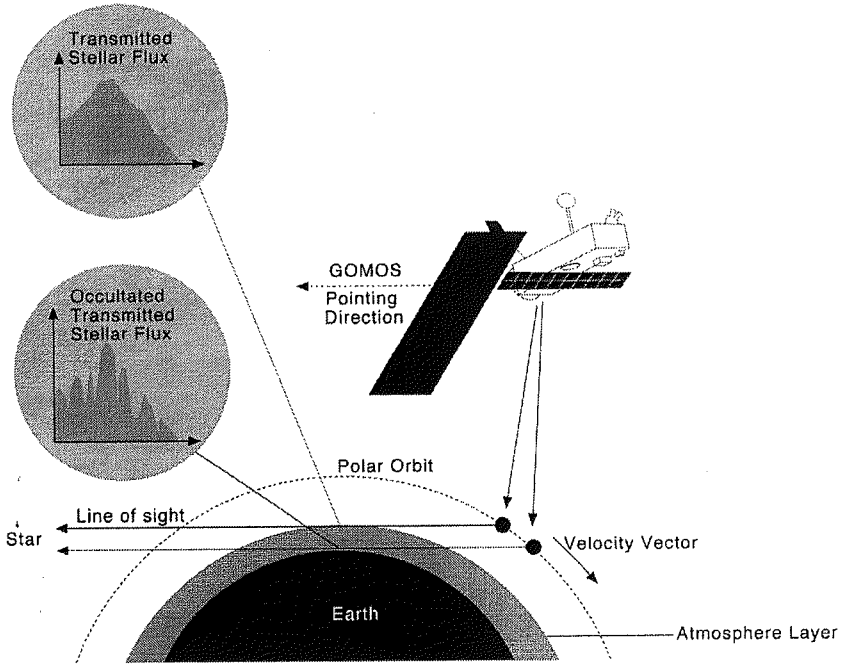


Figure 4: *The measurement principle underlying GOMOS*

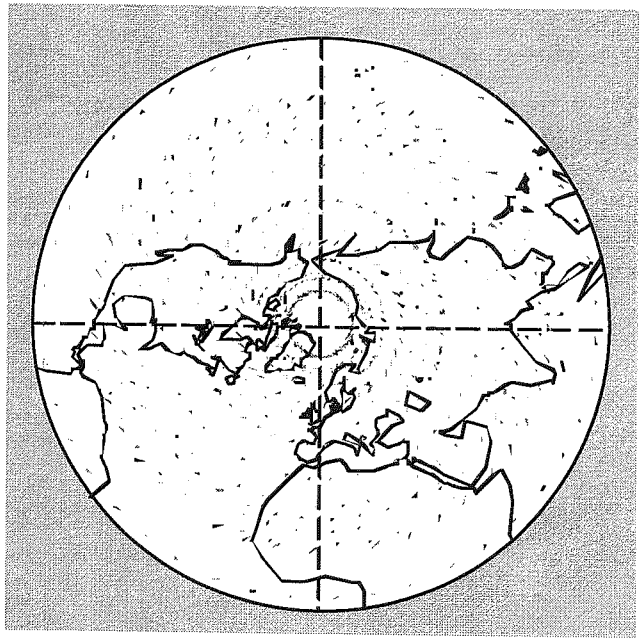


Figure 5: *Location of Ozone measurements in the Northern Hemisphere after three consecutive days (around 21 march) with specific star selection*

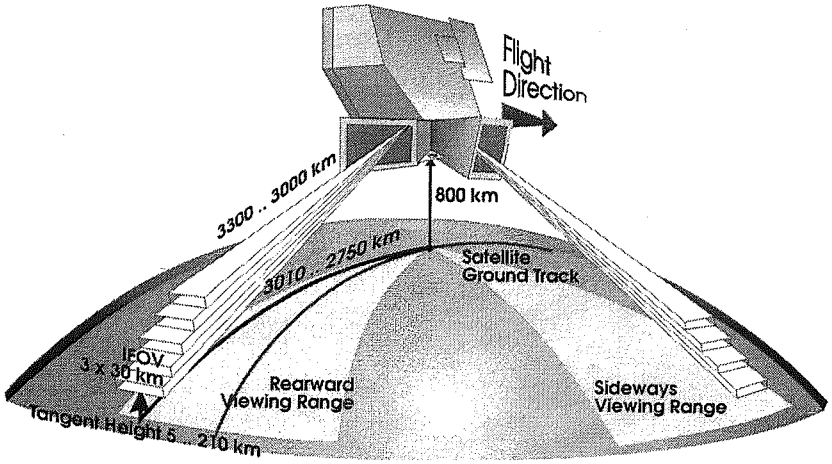


Figure 6: *The viewing geometry of MIPAS*

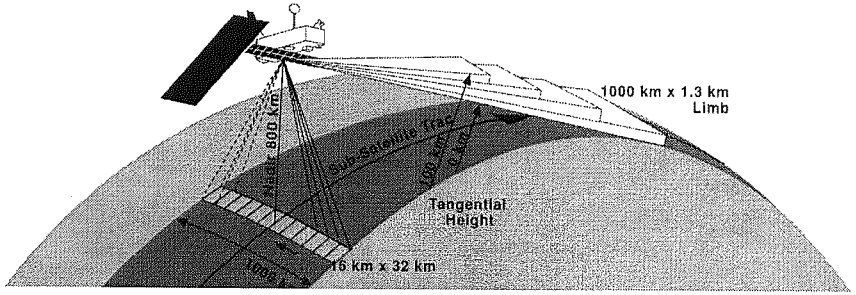


Figure 7: *The viewing geometry of SCIAMACHY*

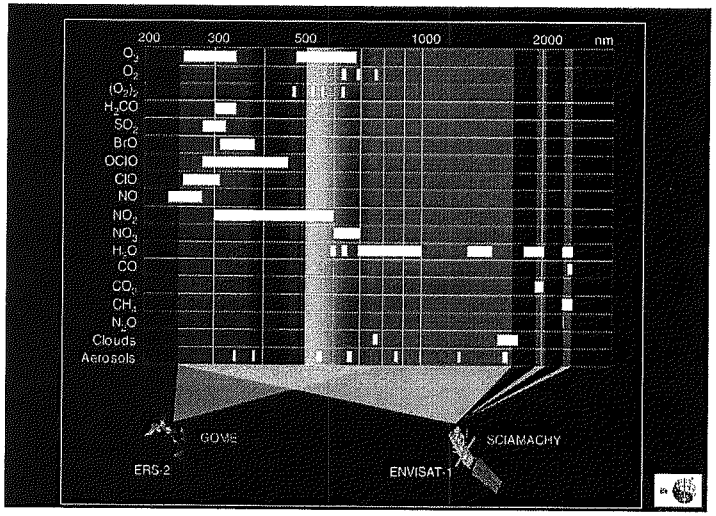


Figure 8: *A comparison of the wavelength coverage of GOME and SCIAMACHY (Acknowledgement: IFE, Bremen, Germany)*

HEIGHT-RESOLVED OZONE RETRIEVALS IN THE LOWER ATMOSPHERE FROM THE GLOBAL OZONE MONITORING EXPERIMENT, GOME

R. Siddans[1], B.J. Kerridge[1], W.J. Reburn[1], R. Munro[2]
[1]Rutherford Appleton Laboratory Chilton, Didcot, Oxon OX11 0QX, U.K.
[2]European Centre for Medium Range Weather Forecasting, Reading RG2 9AX, U.K.

ABSTRACT

A prototype retrieval scheme has been developed at the Rutherford Appleton Laboratory to retrieve height-resolved ozone profiles spanning the troposphere and stratosphere from GOME data, thereby providing the first direct measurement of tropospheric ozone from a single satellite instrument. A brief overview of the retrieval scheme is provided, followed by recent results from the analysis of a period in March 1997 during which sub-tropical air was observed to be irreversibly transported into the mid-latitude lower stratosphere. Work to optimise the algorithm and process the GOME dataset on a large scale data for general release is described.

1. Introduction

The Global Ozone Monitoring Experiment (GOME) was launched on ESA's ERS-2 satellite in April 1995. GOME is a downward viewing spectrometer which measures back-scattered solar radiation from 237-790nm at moderate spectral resolution (0.2 - 0.4nm). Due to contiguous spectral coverage and higher spectral resolution, GOME provides height-resolved ozone data of higher quality than previous UV back-scatter instruments. In particular ozone information extends down into the lower stratosphere and troposphere, a region not measured extensively by previous satellite sensors.

During the course of the SODA project, a retrieval scheme capable of generating ozone profiles on a global basis from GOME measurements was developed and optimised at the Rutherford Appleton Laboratory (RAL), thereby providing the first direct measurements of tropospheric ozone distributions from space (Munro 1998).

Section 2. contains a basic description of the prototype retrieval scheme, developed to demonstrate the potential of the GOME dataset for this application. Validation of this scheme is then described in section 3. and results from a period in March 1997 are presented in section 4.. Optimisation of the scheme to allow processing of a large sub-set of the GOME data is described in section 5., and the status of current data processing is detailed in section 6..

2. Overview of the retrieval scheme

Profile retrieval is carried out using Optimal Estimation (OE) (Rodgers 1990), which solves an otherwise under-constrained problem by combining information from measured spectra with that from *a priori* profiles. The primary input required is the GOME Level 1 Product, comprising measured backscattered spectra, and measurements of direct-sun spectra acquired in-flight via a diffuser. The wavelength calibration of the GOME spectra is refined by fitting to a high resolution solar spectrum (Chance 1996). *A priori* ozone profiles and uncertainties were taken from a monthly zonal mean climatology constructed from SAGE II data (McCormick 1989).

The GOMETRAN radiative transfer model (Rozanov 1997), developed by the University of Bremen, is interfaced to the retrieval scheme to provide synthetic reflectance spectra. It incorporates gaseous absorption, surface reflection and multiple scattering by air molecules, aerosol and clouds. In this study, a fixed single profile was adopted for aerosol scattering and cloud was not included in the radiative transfer.

Due to differences in viewing geometry and ozone absorption properties, a two-step approach is adopted, in which an ozone profile is first retrieved from the 237-307nm region (known as GOME Band 1A), which primarily contributes information on ozone down to, or just below, the stratospheric concentration peak. This profile is then used as *a priori* in a retrieval from 312-405nm (Band 2B). In addition to the ozone profile it is necessary to retrieve: (i) a scaling factor for filling in of Fraunhofer lines by the Ring effect; (ii) wavelength shifts of the back-scattered spectrum with respect to the reference ozone absorption cross-section and the direct-sun spectrum and (iii) total column amounts of BrO and NO2. As a precursor, an effective surface reflectance is retrieved from measurements near 340nm, where absorption by ozone and other gases is relatively small.

The current scheme retrieves the ozone mixing ratio on a fixed altitude grid: 0, 6, 12km, then 4km intervals up to 80km. These levels have been chosen

to slightly oversample the intrinsic resolution of the GOME measurements, while maintaining retrieval stability.

3. Validation

GOME retrievals have been compared to 34 co-located ozone sonde measurements made over Europe (including Northern Scandinavia) between July 1995 and February 1997. Means and standard deviations in the fractional differences in integrated sub-columns between the retrieval levels are shown in figure 1 for: (i) *a priori* profiles (SAGE II climatology); (ii) Band 1A (step 1) retrieved profiles and (iii) Band 2B (step 2) retrieved profiles. The Band 1A retrievals generally have much smaller biases and standard deviations than the *a priori* profiles, and Band 2B results provide very valuable extra information in the lower stratosphere and troposphere. The standard deviation in fractional difference between the final (Band 2B) retrievals and the ozone-sonde measurements is seen to be <10% in the stratosphere and <40% in the troposphere, with biases <5% above 6km and -30% below that altitude, where a negative bias is expected as GOME can only detect ozone above the cloud top and no screening was performed for this inter-comparison.

The profiles have also been validated by comparison to other satellite sensors (Stoffelen 1998). The zonal mean ozone field constructed from Band 1A (first step) retrievals from nine orbits processed on 11 January 1996 were compared with with those obtained from SSBUV and MLS. Agreement is generally within 10% between 1 and 40 hPa, the common altitude range within which the three sensors provide information.

In addition, the retrieved ozone fields have been compared to UKMO tropospheric chemical transport models (Siddans 1998) and, in the context of the SODA project, the ECMWF model (Stoffelen 1998).

4. Results from March 1997

At the suggestion of workers at the University of Aberystwyth, retrievals were carried out to determine the ozone distribution over the E. USA / N. Altantic / W.Europe region (from 20°N to 70°N and 80°W to 50°E), on several days in March 1997. During this period, air of sub-tropical origin had been observed in the mid-latitude lower stratosphere by a number of ozone-sondes (O'Connor 1998), resulting in a characteristic minimum in the ozone concentration profile at approximately 15km.

The potential for GOME to observe such structure in the lower stratosphere was clearly demonstrated by results from the period. Figure 2 shows a comparison between GOME and co-located ozone-sonde. The relatively coarse vertical resolution of the GOME retrieval is apparent, but the ozone minimum at 15km is detected in the Band 2B retrieval. In this case, the contribution of the Band 2B measurements to the retrieval in the lower atmosphere is readily apparent - the step 1 retrieval, which makes use of measurements in Band 1A only, produces a reasonable profile down to 16km but fails to capture the structure below.

A height vrs. latitude cross section of the orbit containing this profile is provided in figure 3.

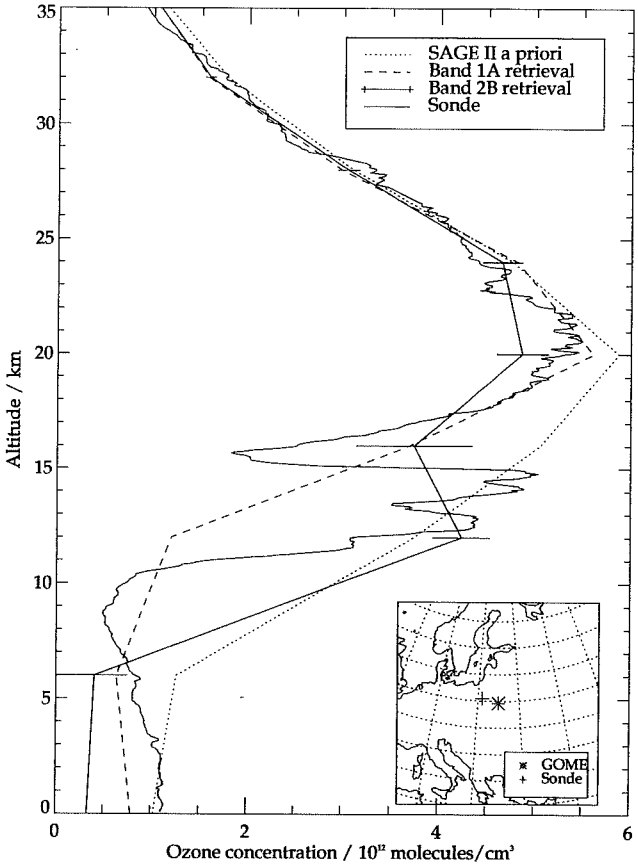


Figure 2. Comparison of GOME ozone retrieval with an ozone-sonde measurement from Legionowo (52.4°N,21.0°E) at 11:30 on 7 March 1997. The solid line with error-bars is the final (Band 2B) GOME retrieval. The dotted line shows the a priori profile used and the dashed line shows the output of the first retrieval step (using Band 1A measurements only).

5. Optimisation of retrieval scheme efficiency

In order that large scale processing of GOME data be feasible the prototype scheme (operating at 1 profile per hour c.f. measurement time of 1.5 seconds) required optimisation. Processing speed is limited by multiple-scattering, high-resolution, radiative transfer (or *forward*) calculations required to synthesise GOME spectra during the retrieval process. Approximations to these radiative transfer calculations were considered:

Band 1A obtains information, down to just below the stratospheric ozone peak, from the steep wavelength dependence of ozone optical depth in the Hartley band. In this region, information of adequate quality can be obtained by matching measured spectra at the 1% level (in reflectance, defined as the ratio of back-scattered radiance to direct solar irradiance). In

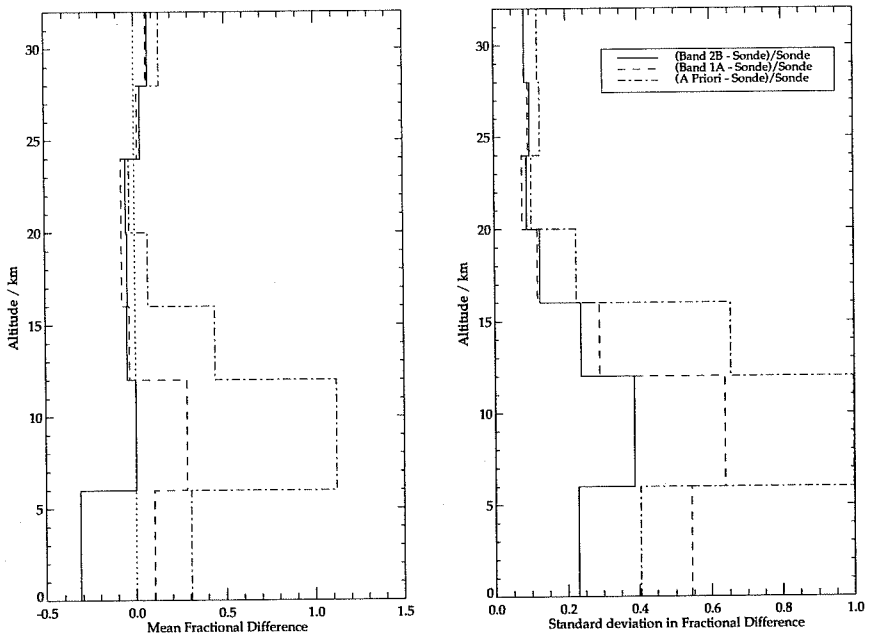


Figure 1. Inter-comparison of retrieved, a priori and sonde integrated sub-columns.

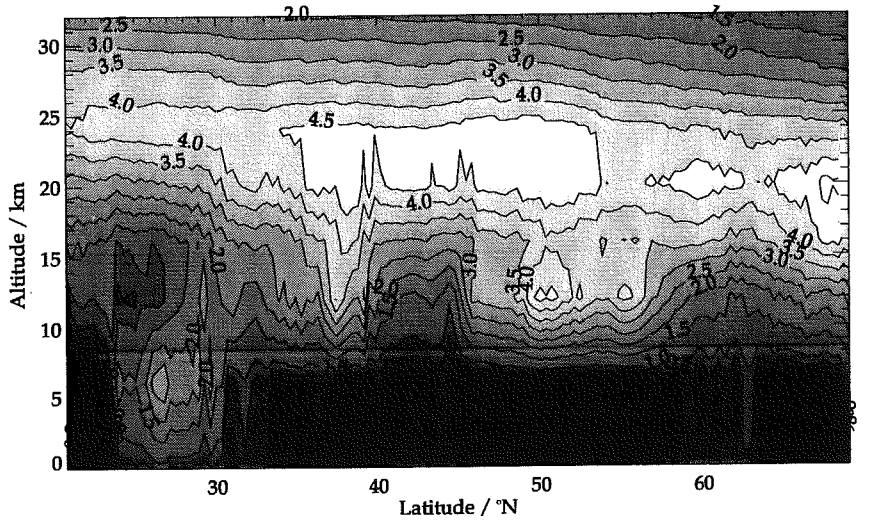


Figure 3. Height vrs. latitude distribution retrieved from GOME orbit commencing 09:20 on 7 March 1997. Contours of ozone number density in 10^{12} molecules/cm^2.

order that profile information be extended down into the lower stratosphere and troposphere by the Band 2B retrieval , however, measurements in the Huggins bands must be matched to at least 0.1% (ESA 1991) to exploit the temperature-dependent structure of the band. Forward model accuracy must therefore be significantly better than 1% and 0.1%, respectively, for retrievals from Bands 1A and 2B to be of similar quality to those of the prototype algorithm.

A scheme has been developed which meets the Band 1A accuracy requirement by making use of exact, GOMETRAN calculations at 24 wavelengths. These are interpolated in absorption and scattering optical depth space to obtain spectra at the required resolution, by making use of the analytic weighting functions provided by GOMETRAN. The scheme relies upon the relatively weak temperature dependence of the band, as a consequence of which the relationship

between total optical depth and reflectance is only a weak function of wavelength. The scheme is accurate at the 0.2% level and rapid enough to allow real-time processing of the Band 1A data within the available computational resource (5 Dec Alpha 433au machines).

The temperature dependence in the Huggins bands, coupled with the more stringent accuracy requirement, precludes the use of a similar scheme for the Band 2B retrieval. Instead, work is proceeding on a look-up table approach to parametrise the effects of multiple scattering in the band. Preliminary work has achieved a similar level of accuracy to the Band 1A scheme (0.2%), which is not adequate for this band. (Retrieved ozone from this approximate scheme deviated by up to 100% in the lower stratosphere from reference calculations performed by the full scheme).

The current scheme makes use of the Band 1A optimised scheme but retains the exact Band 2B scheme. Single profile processing time is now 10 minutes on a single machine. It is anticipated that the look-up table approach wll be refined to provide adequate retrievals from Band 2B, which would allow processing at a rate approaching real-time on the machines currently available.

6. Conclusions and further work

A prototype processing scheme has been developed at RAL which provides height-resolved ozone profiles spanning the troposphere and stratosphere, providing the first direct measurement of tropospheric ozone from space and representing a significant advance (in terms of temporal / spatial coverage) over previous measurements in the lower stratosphere by occultation techniques. The potential of GOME to resolve ozone structure associated with upper troposphere/ lower stratosphere dynamics has been demonstrated for an event in March 1997.

The optimised scheme allows large scale processing of GOME data at the rate of approximately 10 minutes per profile per CPU. The scheme is capable of real time processing of Band 1A data to provide information down to just below the stratospheric concentration peak, however the recovery of information at lower altitudes requires Band 2B measurements to be fitted to a level of accuracy not yet achieved with the approximate radiative transfer model developed thus far. Work is continuing on this model with the aim of enabling full retrievals at a real-time data rate.

The scheme, as it stands, is currently being used to process 1 day in 10 of the GOME data globally for one year starting on 24th March 1998. Sampling was selected to make use of the narrow-swath mode GOME data and the commencement of a special view mode of the ATRS-2 instrument (also on board ERS-2), designed to ensure full coverage of the GOME field of view. A scheme to make use of this ATSR-2 data to obtain accurate cloud coverage and altitude information for interpretation of the GOME retrievals is currently under development at RAL.

It is intended that the profile data, together with ATRS-2 derived cloud information will be made available via the British Atmospheric Data Centre (`http://www.badc.rl.ac.uk/`).

ACKNOWLEDGEMENTS

We would like to acknowledge the UK Natural Environment Research Council and the EC (via the SODA project) for providing the funding to process and validate the GOME data. Permission to use GOME data prior to its public release was granted by ESA and GOME Data Products were provided by DLR. The GOMETRAN radiative transfer model and the high resolution solar reference spectrum used in this study were supplied by Prof. J.Burrows (University of Bremen) and Dr. K.Chance (Harvard-Smithsonian Astrophysical Observatory), respectively. UK Met.Office temperature was acquired through the BADC at RAL and ozone-sonde data was obtained from NILU.

REFERENCES

Chance, K., and R. Spurr, "Ring effect studies, GOME Geophysical Validation Campaign". Workshop Proceedings., ESA WPP-108, pp.69-74, 1996.

ESA, "A study of advanced techniques for monitoring atmospheric trace gases", ESA Contract 8915/90/NL/BI, Final Report (1991).

Fortuin, J. P. F., and U. Langematz, "An Update on the Global Ozone Climatology and on Concurrent Ozone and Temperature Trends", Atmospheric Sensing and modelling, R. P. Santer (Ed.) pp. 207-216, Proc. SPIE 2311 Washington, DC., 1994.

McCormick, M.P., J.M. Zawodny, R.E. Veiga, J.C. Larsen, and P.H. Wang, "An Overview Of SAGE I And II Ozone Measurements" Planet. Space Sci., 37, 12, 1567-1586; 1989.

Munro, R., R. Siddans, W.J.Reburn and B.J.Kerridge, "Direct Measurement of Tropospheric Ozone from Space" Nature, v392(6672), pp.168-171.

O'Connor, F.M. and G. Vaughan, "Observations of sub-tropical air in the European midlatitude lower stratosphere", Q.J.R.Meteorol.Soc., in press.

Rodgers, C. D., "Characterisation and error analysis of profiles retrieved from remote sounding measurements", Journal of Geophysical Research, 95(D5), pp.5587-5595, 1990.

Rozanov, V., D. Diebel, R. Spurr, and J. Burrows, Gometran: A radiative transfer model for the satellite project gome, J. Geophys. Res., v102(D18), pp. 21809-21823.

R.Siddans, B.J.Kerridge, W.J.Reburn and R. Munro, "Height-resolved ozone retrievals in the troposphere and lower stratosphere from GOME", Earth Observation Quarterly, v58, pp.11-13.

Stoffelen, A. (ed.) "Satellite Ozone Data Assimilation", 1st Annual European Union project report.

OZONE PROFILE RETRIEVAL WITH GOME UV/VIS SPECTRA

Roeland van Oss[1], Ronald van der A[1]

Jochen Landgraf[2], Otto Hasekamp[2]

[1]KNMI, PO Box 201, 3730 AE De Bilt, The Netherlands
[2]SRON, Sorbonnelaan 2, 3584 CA Utrecht, The Netherlands

ABSTRACT

The Global Ozone Monitoring Experiment (GOME) on board ESA's ERS-2 satellite measures since April 1995 reflected sunlight from the atmosphere in the UV and visual range at the moderate spectral resolution of 0.2-0.4 nm. The UV part of the spectrum contains information on the height distribution of ozone due to the strong wavelength dependence of the ozone absorption. At KNMI a retrieval algorithm is under development that extracts this information. The algorithm retrieves ozone profiles with a vertical resolution of about 5 km in the stratosphere to 10 km in the troposphere. Recent efforts focus mainly on the optimalisation of the balance between algorithm speed versus accuracy. In this respect the bottleneck is the radiative transport model that simulates the measured spectrum for a given atmospheric state. It has been found that retrievals based on the fast two-stream method produce ozone profiles that compare well in the stratosphere with profiles retrieved using the accurate DISORT radiation transport model.

Key words: Ozone profiles; Retrieval; GOME.

1. INTRODUCTION

Since the late seventies satellites routinely monitor the global ozone distribution. The early TOMS/SBUV instruments measured the backscattered radiation for several frequencies, resulting in an accurate estimate of the total column densities of ozone. Apart from this limited information is obtained about the vertical ozone distribution in the stratosphere. However, a growing need evolved for more accurate profile measurements in stratosphere and in troposphere. Knowledge on tropospheric ozone is important for different reasons. First, because ozone belongs to the greenhouse gases and because of the toxic nature of ozone for humans and vegetation. In addition, the ozone concentrations in the troposphere are a measure for the oxidising capacity of the atmosphere and therefore important for understanding the chemical processes in the troposphere. Chemical modelling will benefit from global measurements of tropospheric ozone. In 1995 GOME (Global Ozone Monitoring Experiment) has been launched to measure in nadir direction the backscattered sunlight from the atmosphere in the range from 240-790 nm. The GOME spectrometer is a unique instrument to retrieve height-resolved ozone densities in the stratosphere as well as in the troposphere. The improvement in comparison to former satellite instruments is the high spectral resolution of the observations in the ozone absorption bands (the Huggins, Hartley and Chappuis bands). The steep rise of ozone absorption from 350 to 270 nm offers the possibility to infer height-resolved information on the ozone concentration from backscattered sunlight. The reason for this is that highly absorbed photons at the short- wavelength side penetrate the atmosphere only shallowly, whilst weakly absorbed photons on the long-wavelength side travel through the full height of the atmosphere. Therefore, backscattered short-wave photons carry only information on the upper layers of the atmosphere and photons with increasing wavelength reveal information on lower layers. Combining the measured reflectances in this wavelength range then gives the desired height-resolved atmospheric properties. The full spectrum between 270 and 350 nm can be used to infer the ozone height profile in a global fit (or retrieval) using the ozone concentrations at a number of specific heights as unknowns and the measured reflectances as known parameters. The retrieval necessitates a radiative transfer calculation to obtain reflectances given the ozone profile and other atmospheric properties.

Below the details of the retrieval method are outlined. Special attention is drawn to the radiative transfer model with respect to its accuracy and computational speed. A set of observations has been chosen for testing the retrieval of ozone.

2. ALGORITHM DESCRIPTION

The earthshine radiances measured by GOME are combined with the sun irradiance measurements to produce the reflectances $r(\lambda)$ that contain information on the atmospheric composition:

$$r(\lambda) = \frac{\pi I_{\text{earth}}(\lambda)}{\cos\theta_{\text{sun}} F_{\text{sun}}(\lambda)}, \qquad (1)$$

with $I_{\text{earth}}(\lambda)$ the earthshine radiance, $F_{\text{sun}}(\lambda)$ the sun irradiance and θ_{sun} the solar zenith angle. A forward model $\mathcal{F}$ is introduced that relates the reflectances at different wavelengths, contained in the vector $\mathbf{r}$, to the atmospheric ozone profile:

$$\mathbf{r} = \mathcal{F}(\mathbf{x}, \mathbf{b}), \tag{2}$$

where $\mathbf{x}$ is a vector of ozone concentrations at a number of altitudes, or, alternatively, ozone column densities between a series of pressure levels. The vector $\mathbf{b}$ consists of atmospheric and instrumental parameters, other than the ozone profile, that also determine the reflectances. These parameters are known on forehand with a satisfactory precision, like the viewing and solar angles, or are solved for along with the ozone profile, like, for instance, the ground albedo.

The inversion of Equation 2 is performed by first linearising the forward model, solving the linear equation, using this solution as a new linearisation point, solving again, and thus iteratively converge to the final solution of Equation 2.

Ignoring for the moment the dependence on $\mathbf{b}$, the linearisation yields:

$$\mathbf{r} = \mathcal{F}(\mathbf{x_0}) + \frac{\partial \mathcal{F}}{\partial \mathbf{x}}(\mathbf{x} - \mathbf{x_0}), \tag{3}$$

or, defining the weighting function:

$$\mathcal{K} = \frac{\partial \mathcal{F}}{\partial \mathbf{x}}, \tag{4}$$

$$\mathbf{r} - \mathbf{r_0} = \mathcal{K}(\mathbf{x} - \mathbf{x_0}), \tag{5}$$

where $\mathbf{x_0}$ denotes an initial profile sufficiently close to the solution and $\mathbf{r_0}$ the reflectances for this initial profile.

The ozone profile retrieval algorithm performs an inversion of Equation 5 by finding the most probable ozone profile $\mathbf{x}$ given the measured reflectances $\mathbf{r}_{\text{meas}}$ and measurement errors $\mathbf{dr}_{\text{meas}}$. The optimal estimation method (OEM) (Rodgers 1976, Rodgers 1990) has been used to perform the inversion. This method allows the inclusion of apriori information on the profile retrieval and deals with the ill-posed nature of the inverse problem by stabilising the solution against error amplification. OEM gives the maximum of the probability distibution of the profile, given the measurements: $\mathcal{P}(\mathbf{x}|\mathbf{r})$. This probability distribution is proportional to the product of the apriori probability distibution $\mathcal{P}(\mathbf{x})$ and the probability distibution of the measurement given the profile: $\mathcal{P}(\mathbf{r}|\mathbf{x})$. Assuming Gaussian probability distibutions an analytic solution for the maximum exists which is known as the optimal estimate. In practise, OEM gives a result close to the apriori values for those parameters whose effect on the measurement is small compared to the measurement error (Case I). If the measurement, on the other hand, critically constrains the parameters (Case II) the apriori has little influence on the retrieved values. The retrieval error is always less then or equal to the apriori error and serves as a measure for the information added by the measurement. For Case I parameters the retrieval error is close to the apriori error, whilst for Case II it will be significantly reduced. For the problem at hand the tropospheric ozone concentrations represent the first case because they only have a small effect on the measured reflectances. The ozone concentrations at higher altitudes are Case II parameters since they have more effect on the reflectances and the retrieved concentrations there are therefore less affected by the apriori. The balance between apriori information and information from the measurement is determined by the measurement error and the apriori error. It is therefore of utmost importance to estimate these errors correctly.

For the apriori ozone profile a zonal average climatology has been used from Fortuin and Kelder 1998 This is a zonal mean ozone climatology based on a 12 year observation period of ozone sonde measurements and satellites. For the apriori errors the natural variance of the ozone concentrations in the climatology has been used.

The final solution for the ozone profile is reached after following through a number of iteration steps in which the forward model is repeatedly linearised and optimal estimates are computed. This iterative proces is terminated when decrease of the residu between the measurements and the simulated reflectances is not significant anymore.

3. RADIATION TRANSPORT MODELS

The bottleneck for accuracy and computational speed of the retrieval is the forward model and it therefore deserves our close attention. The forward model produces, for a given wavelength, the earthshine radiance given the solar irradiance, the solar and line-of-sight angles and the absorption and scattering cross-sections througout the atmosphere. The earthshine radiance is the total of a single- and a multiple-scattered component. The single scattered solar radiance can be calculated analytically and dominates the total radiance for wavelengths smaller than 300 nm. Calculating the multiple scattered radiance poses the largest challenge in the forward model. Among the state-of-the-art radiation transport models that account for multiple scattering the Discrete Ordinates method (DISORT) (Chandrasekhar 1960, Stamnes & Swanson 1981) features prominently. This method applies a discretisation of the angular dependency of the radiance based on Gaussian quadratures. By adjusting the number of polar angles (or streams) used in the discretisation the accuracy of the model can be tuned for various circumstances. In cases with complex scattering phase functions, for instance when dealing with clouds or aerosol layers, at least eight streams should be used. For situations where molecular scattering controlls the upwelling radiance eight streams is generally adequate. Inhomogeneous atmospheres have to be divided into a sufficient number of adjacent homogeneous layers for which DISORT gives the radiances at the boundaries leaving certain constant undefined. Subsequently, the radiances at the boundaries are matched which fixes the values of the afore mentioned constants. With DISORT radiances at the top-of-the-atmosphere (TOA) can be calculated which are accurate enough for our purposes. However the code demands too much processing time on present

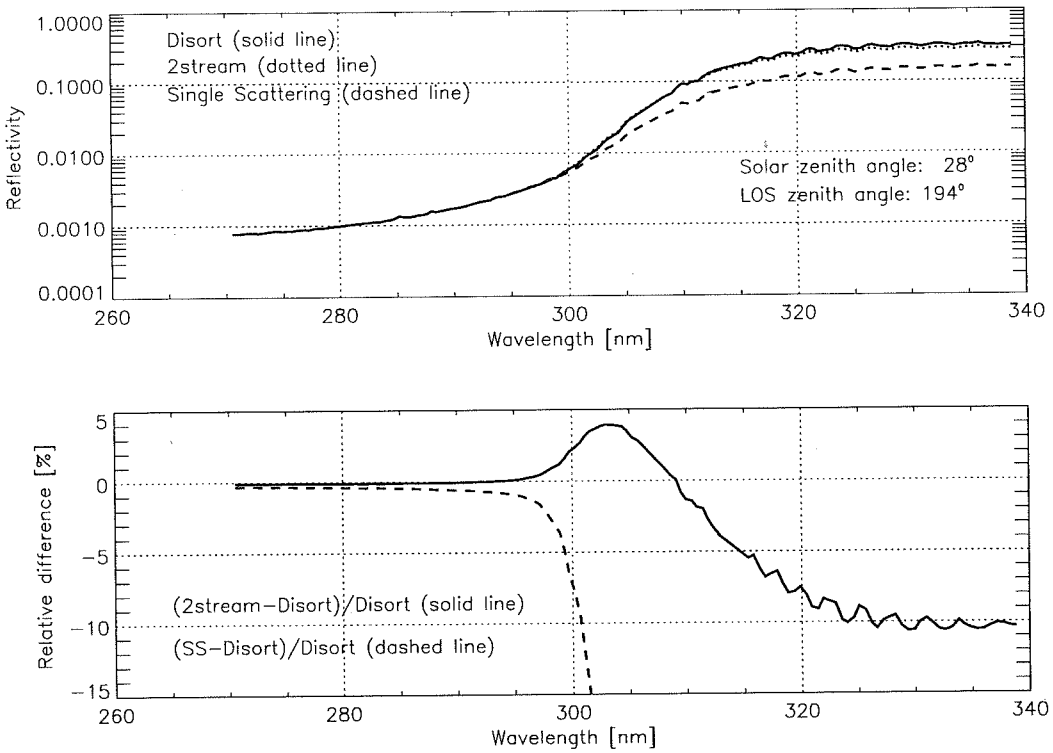


Figure 1. A comparison between the reflectances calculated with DISORT *(8 streams), the two-stream model and the single scattering approximation is shown. The bottom panel shows the relative differences between the relflectances calculated with the three models. The solar and viewing zenith angles are as given in the figure. (The large deviation with the single scattering reflectance above 300 nm is not relevant here and therefore allowed to fall outside the panel.)*

day computational facilities for the operational retrieval of ozone profiles from GOME spectra. Therefore, a need exists for fast algorithms that produce radiances which are still accurate enough for ozone profile retrieval. As a promising candidate we have used an analytic two-stream approximation (the hybrid delta-Eddington model, Meador and Weaver 1980) in our retrieval code to assess its performance and compared it with retrievals using the DISORT model. The present version of the algorithm incorporates the radiative transfer shell MODTRAN 3.7 (Berk et al. 1989) which contains the two-stream and DISORT multiple scattering codes as options. The two-stream model does not give radiances directly, but instead produces downward and upward fluxes at the boundaries of each atmospheric layer. These fluxes are used to estimate the source function in the viewing direction in each layer. The source function is then integrated along the line of sight from the ground to TOA. This procedure is also used in DISORT to calculate the radiance at a user specified direction from the solution at the discrete angles (Stamnes & Swanson 1981).

4. RESULTS

In Figure 1 a comparison between the reflectances calculated with DISORT (8 streams) and the two-stream model is shown for a specific choice of solar and (θ_{sun}) viewing (θ_{view}) angles (near nadir). Also, the singly scattered radiance is shown. The bottom panel shows the relative differences between the relflectances calculated with the three models. Evidently, the two-stream model approximates the reflectances better than 1% below 295 nm. The single scattering contribution to the total radiance is larger than 98% in this spectral range. For the spectral range [300 - 340 nm] the two-stream model gives radiances that deviate up to 11% and the single scattering contribution drops below 50 %. For larger solar zenith angles (lower sun) the two-stream approximation grows worse: for $\theta_{\text{sun}} = 73^0$ the deviation reaches 25% at 320 nm.

To quantify the loss in accuracy when two-stream is used in the retrieval, ozone profile retrievals are carried out for 13 GOME spectra using two-stream and DISORT as the forward model. The spectra belong to ground pixels that cover a representative set of solar zenith angles and ground scenes. The pixels from the GOME orbit (date: 23 Feb 1998; no: 80223061) extend from the North Pole southward to a latitude of -40^0 and cross Tibet and the western part of the Indian Ocean. The solar zenith angle varies between 90^0 at the pole to 26^0 close to the equator. Table 1 shows the results of the comparison averaged for three altitude ranges: troposphere, lower and upper stratosphere. The relative deviations are shown jointly with the relative deviations between the DISORT retrieval and the climatology used as apriori.

From Table 1 it can be inferred that in the upper stratosphere the deviations between the results with

Table 1. Relative rms deviations for three altitude ranges between retrieved ozone concentrations [cm^{-3}] using DISORT *and two-stream. These values have to be compared with the relative deviation between the* DISORT *result and the apriori values from the climatology displayed between brackets. The first column gives the latitude of the ground pixel, the second column the solar zenith angle. The relative deviation has been defined as* $100\%(c_{2\text{stream}} - c_{\text{DISORT}})/c_{\text{DISORT}}$, *with* c_{method} *the retrieved ozone concentration using method two-stream or* DISORT.

Lat ^{0}	Sza ^{0}	[0-16 km] %	[16-28 km] %	[28-48 km] %
76.23	86.50	14.0 (8.1)	31.6 (83.4)	2.1 (18.2)
68.89	78.87	11.3 (23.7)	2.5 (39.1)	0.4 (16.8)
60.21	70.66	8.2 (16.5)	2.9 (26.3)	0.3 (12.7)
51.25	62.56	12.9 (18.9)	2.7 (13.8)	0.3 (12.6)
44.26	56.45	63.1 (11.9)	9.3 (16.0)	0.6 (15.1)
21.65	38.49	3.7 (8.0)	2.6 (5.0)	1.6 (19.2)
12.40	32.63	10.1 (32.4)	25.5 (58.9)	0.9 (23.7)
3.86	28.69	3.2 (21.7)	18.3 (31.5)	0.6 (21.8)
-5.40	26.79	14.8 (37.8)	7.6 (11.2)	0.2 (21.2)
-13.94	27.64	19.2 (53.5)	12.2 (23.1)	0.6 (17.6)
-23.17	31.15	8.6 (33.7)	2.4 (5.2)	0.2 (17.6)
-31.66	36.09	15.8 (27.4)	3.2 (6.0)	0.3 (18.2)
-40.11	42.08	19.1 (49.6)	4.1 (6.5)	0.5 (26.1)

two-stream and DISORT are small compared to the deviation between apriori and retrieval. This can be attributed to the small error made in the radiances at wavelengths smaller than 300 nm when calculated with the two-stream method. The radiances at these wavelengths are sensitive to the ozone concentrations above $\sim$ 25 km. The tropospheric values, however, are considerably different for the two methods and are, at the least, a sizeable fraction of the apriori update. For the lower stratosphere we can conclude that the two-stream method supplies extra information to the apriori, but can give rise to a difference with the DISORT retrieval of up to 10 - 30 %.

Finally we note that the difference in CPU time between the retrieval with two-stream and DISORT is striking: the retrieval of the fourteen pixels on a fast computing facility took 15 minutes with the two-stream option and 20 hours with DISORT.

5. CONCLUSIONS

Ozone profile retrievals have been perfomed using an accurate radiative transfer model (DISORT) and an analytic model (two-stream). Reflectances calculated for typical circumstances with both radiative transfer models show a good agreement for wavelengths below 300 nm. For longer wavelength the relative deviation reaches 10%. The results of the ozone profile retrievals show a small deviation of around 1% in the upper stratosphere. For the lower stratosphere the retrieval improves on the apriori, but in the troposphere the retrieval with two-stream does not give reliable results. The retrievals with two-stream decreased the CPU time by a factor of 80 compared to the DISORT retrievals.

ACKNOWLEDGMENTS

This ongoing work is partly funded by the Netherlands Remote Sensing Board (BCRS).

REFERENCES

van der A, R. J., and H. M. Kelder 1998, Earth Observational Quarterly no.58, March, 1998.

van der A, R. J., R. F. van Oss, H. M. Kelder 1998, Ozone profile retrieval from GOME data, *Satellite Remote Sensing of Clouds and the Atmosphere II, SPIE Vol. 3495,* edited by J.E. Russell.

Berk, M. A., L. S. Bernstein, D. C. Robertson 1989, MODTRAN: A moderate resolution model for LOWTRAN7, *Tech. Rep. GL-TR-89-0122,* Geophysics Laboratory, Hanscom AFB, Mass.

Burrows, J.P., M. Buchwitz, V. Rozanov, M. Weber, A. Richter, A. Ladstätter-Weißenmayer, M. Eisinger 1997, The global ozone monitoring instrument (GOME): mission, instrument concept, and first scientific results. Proc. 3rd ERS Symp. on Space at the service of our Environment, ESA SP-414, 3 Vols., pg 585.

Chandrasekhar, S. 1960, Radiative Transfer, Dover

Fortuin, J. P. F., H. M. Kelder 1998, J. Geophys. Res., in press

Meador, W.E., W.R. Weaver 1980, J. Atmos. Sci. 37, 630

Munro, R., R. Siddans, W. J. Reburn, B. J. Kerridge 1998, Nature, 392,

Rodgers, C. D. 1976, Rev. Geophys. and Space Phys., 14, 609

Rodgers, C. D. 1990, J. Geophys. Res. 95, 5587

Stamnes, K., Swanson, R.A. 1981, J. Atmos. Sci., 38, 387

WIND VECTORS FROM TOTAL OZONE IMAGES DERIVED FROM GEOSTATIONARY IR SOUNDERS

F. Karcher, Y. J. Orsolini [1], J.-P. Meyer, S. Muller, P. Armand, D. Blaison
Météo-France, Centre National de Recherches Météorologiques (CNRM)
Toulouse, France
and
H. Kelder, M. Allaart, and V. Pultau
KNMI, De Bilt, The Netherlands

1. INTRODUCTION

Among the novel features of Meteosat Second Generation is the capability to observe the upwelling radiation in the 9.7 micron region by the so-called ozone channel. Information on the ozone distribution in the atmosphere at a high temporal frequency and high spatial resolution, if available with MSG, should provide a new insight on the motions in lower stratosphere where ozone has high concentrations and behaves as a passive tracer under most circumstances. This statement underlies the studies undertaken for the ozone assimilation in Numerical Weather Prediction. Besides, it is worth to examine the possibility of extracting directly the information on air motion from ozone observations, at scales smaller than those currently accessible to assimilation, in a way similar to what was achieved by the use of infrared, visible and water vapour channels of Meteosat, by automatic tracking of clouds or water vapour structures.

We present here the main results of the study performed in the framework of an EUMETSAT contract and described in more details by Karcher et al. (1998).

2. METHODOLOGY

The atmospheric transmission integrated over the 9.7 micron channel vanishes only in cloudy situations and consequently the radiance observed in this channel always contains a contribution from the surface or from the clouds. These observing conditions are different from the water vapour channels and do not permit the mere use of the ozone channel brightness temperature for pattern recognition in two consecutive images. A prior determination of the ozone columns is performed. High horizontal resolution is needed and the CNRM TOVS total ozone algorithm was selected for its ability to use raw radiances.
For the ozone wind application, the requirements on accuracy may be relaxed but the sensitivity of the total ozone deduced to any perturbing condition need to be carefully studied.

The feasibility study of « ozone winds » was performed using data from the geostationary sounder with 18 infra-red channels of the NOAA GOES-8 satellite. Time and horizontal resolutions (respectively 1 hour and 10 km) are close to those of the SEVIRI instrument of MSG (15 minutes and 3 km at the sub-satellite point).

3. MODELLING OF A SELECTED STUDY CASE

A study case (February 23 and 24, 1996) was selected using an animation of the TOVS total ozone fields at the pixel horizontal resolution and projected onto the GOES-8 sounder observing window. Simultaneous series of total ozone images were produced using the GOES-8 sounder radiances on one hand and a using a fine scale modelling on the other hand.

The model is the isentropic Prather advection model used in Orsolini (1995). Ozone is passively advected on 25 isentropes, independently, ranging from 335K to 650K, that is, approximately from 250 to 15 hPa. The horizontal resolution of the model grid corresponds to T213, i.e. about 1/2 degree. The three-dimensional ozone field at high horizontal and vertical resolution is afterwards reconstructed and the total ozone field is derived. The period simulated spans 11 days from February 13 to 24, 1996. Winds are derived from the European Center for Medium Range Weather Forecast global meteorological analyses at resolution T106, archived on 11 pressure levels between 500 and 10 hPa.

Figure 1 shows the GOES-8 total ozone and model total ozone. A noteworthy feature is the filament stretching east to west along the 40° N latitude circle. Ozone increases markedly from 290 DU to the south of the filament up to near 400 DU in the core of the filament, decays sharply to values as low as 320 DU before rising up to above 400 DU further north. The meridional extent of the broad tongue of subtropical air, characterised by low total ozone, is very nicely captured by the model. The model total ozone shows a main structure oriented east to west, but secondary structures are also present, sideways of the main filament, which are not seen in the GOES-8 total ozone.

The three-dimensional model results allow to determine the height range over which this filament is present. The latitude-pressure cross section of ozone partial pressure at 0900 GMT and at a longitude of 100°W on Figure 2 reveals a relatively deep, coherent but slightly tilted feature near 40°N over a pressure range of 250 to 80 hPa.

[1] Now at Norwegian Institute for Air Research (NILU), Kjeller, Norway.

The ozone features described have horizontal scales ranging from 100 to 1000 km and appear with a contrast of some 100 DU. Smaller structures are also searched for but are more difficult to extract due to the measurement noise of the total ozone product.

4. THE PERTURBATION OF THE ALGORITHM BY THE PRESENCE OF CLOUDS

The TOVS total ozone algorithm was set up for clear air conditions. Clouds affect the brightness temperatures measured in most of the TOVS channels. The behaviour of the total ozone algorithm was studied in cloudy conditions by modelling the radiative transfer for 24 mid latitude profiles, adding a cloud layer with variable top pressure from 1000 to 100 hPa, by steps of 100 hPa. Clouds were modelled as blackbodies. From the synthetic cloudy radiances, total ozone values were derived and compared to the real ozone columns.

With cloud top pressure greater than 500 hPa, the effect of clouds is less than 5 DU. Above the 500 hPa level, the effect increases fast as a consequence of the decreasing contrast between foreground and background temperature which appears at the denominator of the expression providing the ozone channel transmission.

The effect of the clouds was also studied at each step of the algorithm and particularly for the correction added to the window channel temperature in order to obtain the background temperature (brightness temperature in the ozone channel of the upwelling radiance measured at 400 hPa). While the theoretical correction determined by the radiative transfer calculation tends towards zero when the altitude of the cloud increases, the correction determined by the clear sky algorithm increases above zero (see figure 3). The zero theoretical correction for a high cloud merely indicates that the cloud has the same brightness temperature in all channels. To improve the algorithm correction, a factor proportional to the temperature difference between window channel and the 400 hPa temperature is applied to it. Figure 4 shows the results of applying the correction reduction in a comparison with simultaneous GOME observations.

5. CHARACTERISATION OF OZONE FEATURES

A total ozone structure can be observed in a given window (sub-array of an image) as soon as the total ozone value is not constant in this window. Taking into account the noise of the total ozone determination, the structure can be observed whenever the total ozone variation (maximum minus minimum value) is greater than the mean noise.

A practical means to extract the structures out of the noise is to calculate displacement vectors, as is done usually for cloud motion vectors. The size of the correlation window determines the scale of the structure. Correlation windows of dimensions 85, 150 and 300 km were studied. Many displacement vectors are tried inside a search window whose size is dependent on the range of possible winds to be extracted. The displacement providing the highest correlation in the correlation window is compared to the analysed winds at various levels. We defined three classes of structures moving like the wind at one only of the levels 500, 250 and 100 hPa. The method misses the winds equal at all levels but strictly selects the structures at a given level.

For each of the nine classes of structures defined (3 levels, 3 scales) averages and standard deviations of some parameters could be derived : number of wind vectors detected, total ozone and its amplitude and noise level within the correlation window, window channel brightness temperature, analysed wind.

Structures of the level 500 hPa are not ozone structures. Structures at the other levels can be considered as ozone structures because the tropopause was below the level 250 hPa or the structures were detected in clear sky. It is observed that the number of structures decreases by a factor of 3 and 2 from level 250 hPa to 100 hPa for the scales 85 and 150 km whereas it increases by a factor 2 for the structures at the 300 km scale. So the scale of the structures increases with altitude. Such results are still to be confirmed on different study cases and their methodology will be tested on total ozone fields derived from fine scale modelling.

6. CONCLUSIONS

The ozone winds feasability study pointed out many interesting questions to be further investigated. Main achievements are :

1. An algorithm has been developed to process the radiances of the GOES sounder.
2. The ozone total content of the atmosphere above 400 hPa is a suitable parameter to detect and track ozone features at sizes 80 to 600 km.
3. Using a high resolution stratospheric advection model, total ozone features originating in the erosion of the stratospheric wintertime polar vortex were found at all latitudes, with signatures ranging from 10 to 50 DU.
4. The algorithm was adapted to the SEVIRI conditions.
5. Improvements of the infrared total ozone algorithms (SEVIRI, TOVS) in cloudy conditions are predicted.

7. REFERENCES :

F. Karcher, Y.J. Orsolini, J.-P. Meyer, S. Muller, P. Armand, D. Blaison, H. Kelder, M. Allaart, and V. Pultau, Study on the exploitation of the ozone channel on Meteosat Second Generation for the extraction of wind information, *final report of contract EUM/CO/96/439/MPe*, May 1998.

Orsolini, Y. J., On the formation of ozone laminae at the edge of the Arctic polar vortex. *Quart. J. Roy. Met. Soc.*, **121**, 1923-1941, 1995.

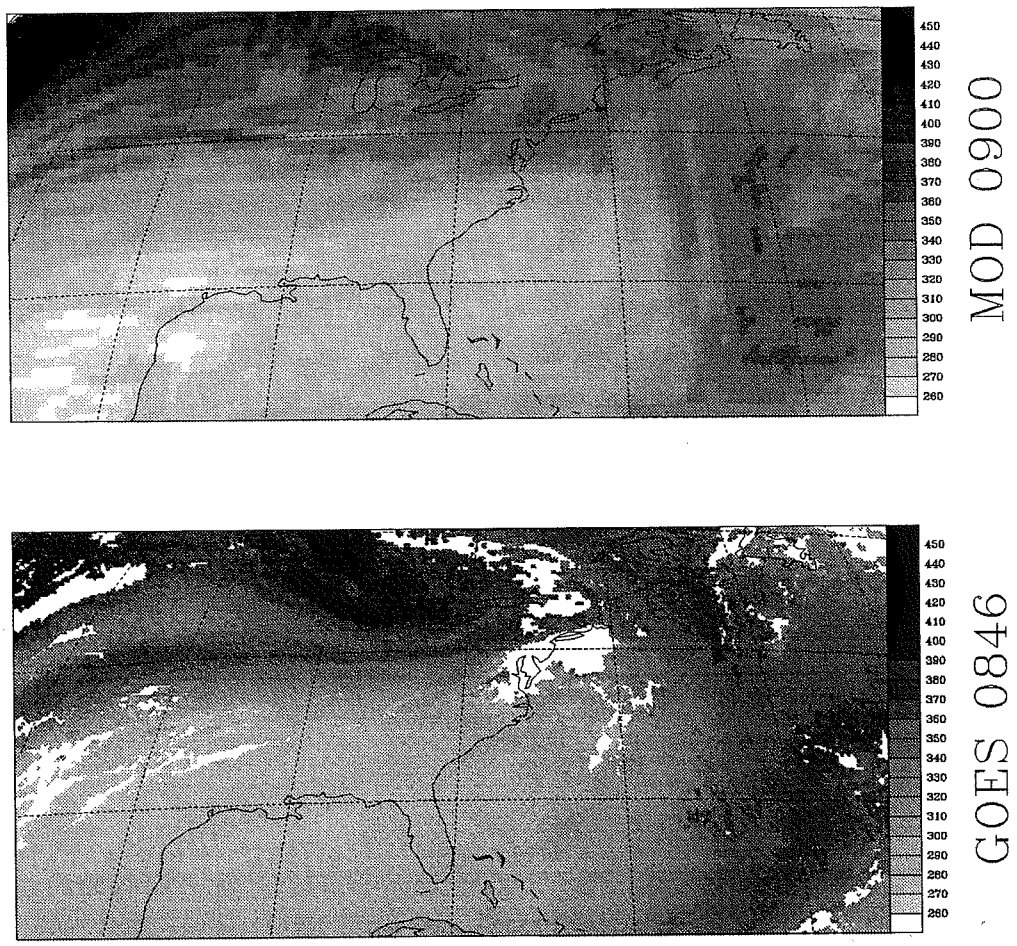


Fig. 1a et 1b : Total ozone fields, produced by the fine scale advection model for the 96/02/24 at 09 TU and observed by GOES-8 for the same day at 08h 46 TU.

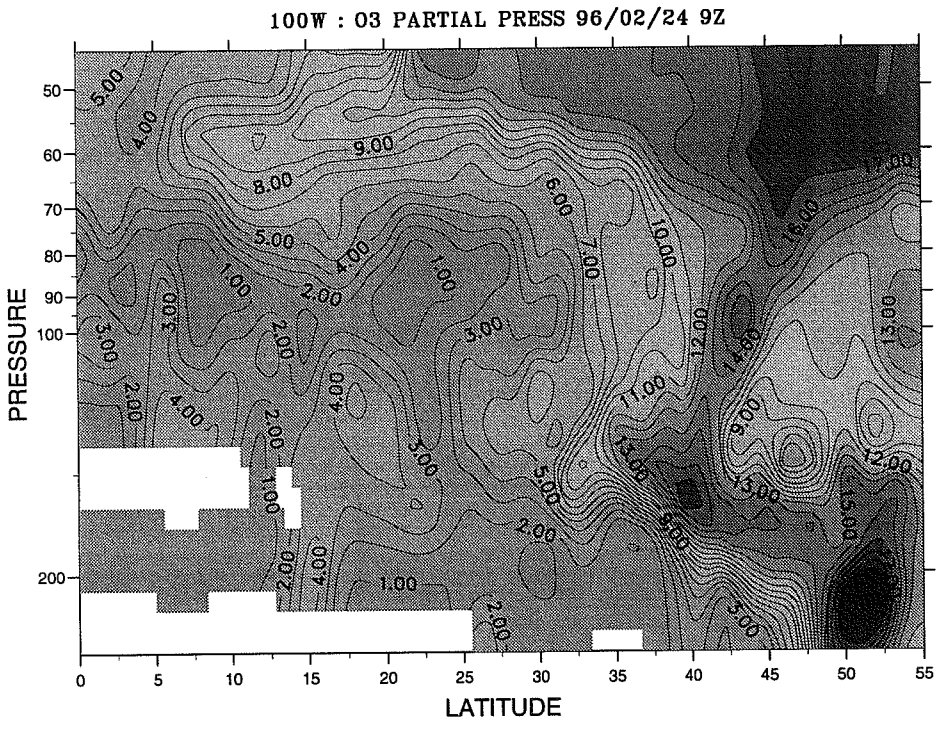


Fig. 2: Latitude-pressure cross section of ozone partial pressure at 100 W for FEB 24 (0900 GMT).

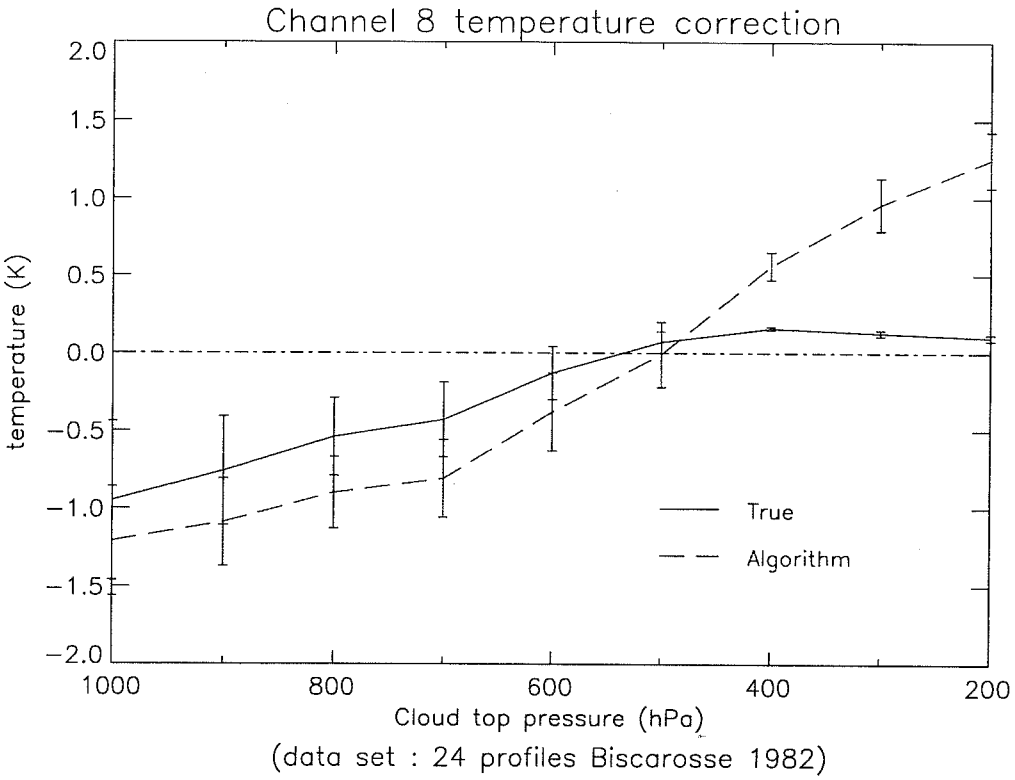


Figure 3 : Corrective term c for the background temperature T8 + c . Theoretical value given by the radiative transfer calculation and value given by the algorithm before correction.

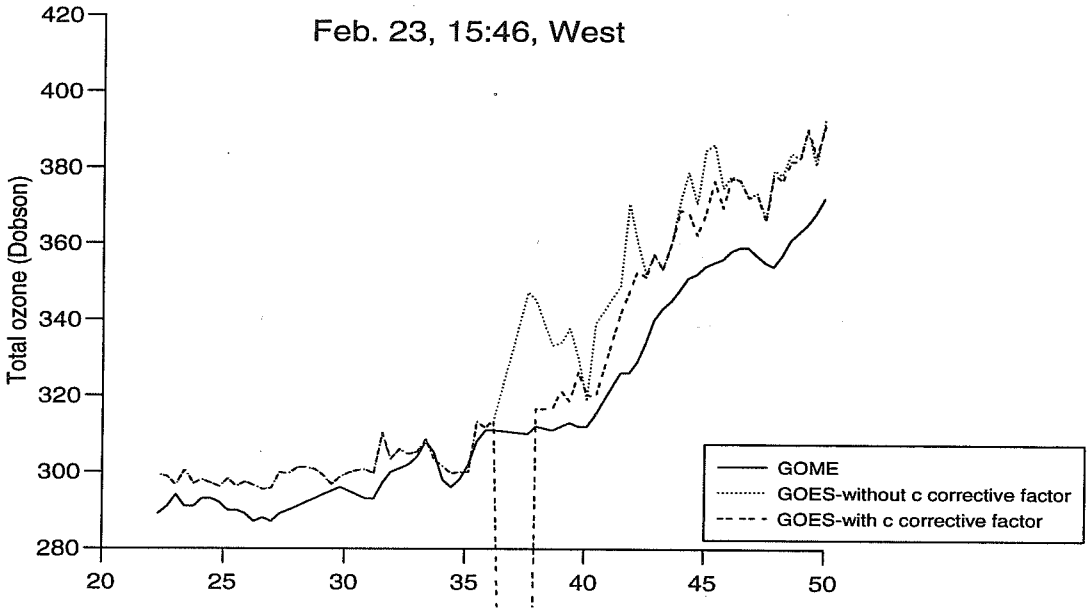


Fig. 4 : Total ozone values along the orbit GOME/ERS-2 versus latitude produced by GOME , GOES-8 (with and without the corrective term c)

FAST RADIATIVE TRANSFER FOR THE TOVS, A SUMMARY OF SODA RELATED TASKS

P. Brunel, Météo-France,
Centre de Météorologie Spatiale (CMS), Lannion, France
and
F. Karcher, Météo-France,
Centre National de Recherches Météorologiques (CNRM)
Toulouse, France

1. INTRODUCTION

To directly assimilate the radiances measured by the TOVS/HIRS for numerical weather prediction, a fast radiative transfer model has been developed since many years at the European Centre for Medium-range Weather Forecasts (Eyre, 1991). Using first guess atmospheric profiles, the fast model (RTTOV) calculates first guess radiances which are fitted to the observed radiances. The addition of the ozone mixing ratio as a variable describing the atmospheric state is achieved within the SODA project and leads to an update of the RTTOV which has to include the calculation of the radiances of the 9.7 micron « ozone » channel (Saunders et al., 1998). Taken from this publication, only the work related to the ozone channel is presented here.

2. THE RTTOV FAST RADIATIVE TRANSFER MODEL

The fast radiative transfer model constructs the outgoing radiance to space in the TOVS channels from the transmittance profile of the atmospheric gases discretized over some 40 layers. The optical depth of each layer in each channel and hence its transmittance can be accurately calculated from the composition, pressure and temperature of each layer as a spectral integration over the TOVS channel of the sum of the contributions of the absorbing lines. For computationnal efficiency, the optical depth is regressed from a set of 9 parameters describing each layer and shown in Table 1. The coefficients of the regression equation are calculated on a data set for which both the predictors (9 parameters) and the optical thickness are accurately calculated using a line by line code. This code is HARTCODE (for RTTOV version described in Saunders et al, 1998) and GENLN2 (for RTTOV version 5 used now). The learning data set is made up of 32 selected atmospheric profiles observed at 5 different nadir angles.

3. IMPROVEMENTS OF THE FAST MODEL

First the list of the 9 predictors was changed according to Table 1 where the subscript j refers to the layer number. The fact that the predictors of a given layer also depend on parameters of the overlaying layers is related to the fact that an accurate radiance calculation is to be performed with a vertical integration prior to the spectral integration.

A second improvement is obtained by considering that the layer and channel transmittances are not merely the product of the transmittances of each of the atmospheric gaseous constituants. Again the main reason for the difference is the spectral integration performed in the calculation of « channel » transmittances. As a consequence, the layer and channel transmittances are now calculated as the product of 3 new factors, each of them statistically predicted :

$$\tau_{i,j} = \tau_{i,j\ \ mix} \times \frac{\tau_{i,j\ \ mix+wv}}{\tau_{i,j\ \ mix}} \times \frac{\tau_{i,j\ \ mix+wv+oz}}{\tau_{i,j\ \ mix+wv}}$$

In the above formula, $\tau_{i,j\ \ mix+wv+oz}$ is the transmission of layer j in channel i taking into account the absorption by « mixed gases » (subset of atmospheric constituents with constant mixing ratios), by water vapour and by ozone.

Finally, for the fast calculation of the factor $\frac{\tau_{i,j\ \ mix+wv+oz}}{\tau_{i,j\ \ mix+wv}}$, it was necessary to built a new learning set of 34 profiles containing temperature, water vapour and ozone concentrations.

4. COMPARISON WITH THE CNRM OZONE PROFILE DATABASE

The accuracy of the fast model was tested on an independent data set of profiles and corresponding radiances calculated with the FASCOD-3P line by line model. This data set was constructed by CNRM for the purpose of calibrating a total ozone algorithm using the TOVS/HIRS radiances and was designed to best take into account the variability of ozone and temperature. Temperature, water vapour and ozone were collected for a series of 22 radiosonde stations spread over the latitude range 82.5 N to 90 S and representing a whole annual cycle. The upper part of the profiles are extracted from the results of a 2D photochemical model by Teyssèdre (1994). The data base is called « SODA » in Saunders et al., 1998.

Differences observed in radiances calculated with the CNRM profiles using the fast model RTTOV and FASCOD-3P (see standard deviations in Figure 1) can be interpreted both as :

1. differences in the data sets used (CNRM versus 32 profiles used for the regressions),

2. differences in the line by line codes (FASCOD-3P and HARTCODE) used to calculate the radiances and transmissions,

the residuals between true and predicted transmittance leading to less than 0.3 K brightness temperature.

The first type of difference appears in window, ozone and water vapour channels (7 to 12) where the variability in atmospheric and surface conditions are not exactly the same in the profile sets. The second type appears in channel 17 near 4 microns where HARTCODE had no N_2 absorption. Figure 1 also shows the improvements obtained by the upgrade of RTTOV described in the previous paragraph. Most interresting is the reduction of near 1K rms obtained in the ozone channel.

Further improvements of the fast model including use of the GENLN2 line by line radiative transfer model led to the version 5 of RTTOV presently used at ECMWF and CMS. Figure 2 shows that the standard deviation of the differences on the radiances calculated on the CNRM data set is less than 0.7 K for all HIRS channels.

RTTOV includes a linear tangent model providing the partial derivatives of the synthetic radiances with respect to the profile parameters. To that respect, is is worth to observe that atmospheric ozone also influences the « temperature channels » 1 to 8 though some 25 to 50 times less than the « ozone » channel.

5. CONCLUSION

Fast modelling of the ozone channel radiance as a function of the ozone mixing ratio profile is now performed with the RTTOV and enables the direct assimilation of the TOVS ozone channel.

The comparison of outputs with the CNRM data base of profiles and line by line radiances helped explaining part of the differences between first guess and observed radiances. In addition to forecast errors, the RMS of the differences may contain various types of contributions from instrumental and radiative transfer errors which need to be well understood.

6. REFERENCES :

Eyre, J. R., 1991. A fast radiative transfer model for satellite sounding systems. *ECMWF Research Dept. Tech. Memo.* 176

R. Saunders, M. Matricardi and P. Brunel, 1998. An Improved Fast Radiative Transfer Model for Assimilation of Satellite Radiance Observations, *submitted to Q. J. Roy. Met. Soc.*

H. Teyssèdre, 1994. Observations satellitaires et modélisation de l'évolution à long terme de l'ozone stratosphérique et influence des activités humaines. *Thèse de doctorat de l'université Paul Sabatier, Toulouse.*

	uniformly-mixed gases	old water vapour/ ozone	HIRS new water vapour/ ozone	AMSU new water vapour
X_{1j}	$\delta T_j \sec\theta$	δT_j	δT_j	δT_j
X_{2j}	$\delta T_j^2 \sec\theta$	$\overline{p\ \delta T_j}$	$\overline{p\ \delta T_j}$	$\overline{p\ \delta T_j}$
X_{3j}	$\overline{\delta T_j} \sec\theta$	δq_j	δq_j	δq_j
X_{4j}	$\overline{p\ \delta T_j} \sec\theta$	$\overline{p\ \delta q_j}$	$\overline{p\ \delta q_j}$	$\overline{p\ \delta q_j}$
X_{5j}	$(\sec\theta - 1)$	$\delta T_j (\sec\theta\ u_j)^{1/2}$	$\delta T_j (\sec\theta\ u_j)^{1/2}$	$\delta T_j (u_j)^{1/2}$
X_{6j}	$(\sec\theta - 1)^2$	$\delta T_j^2 (\sec\theta\ u_j)^{1/2}$	$\delta T_j^2 (\sec\theta\ u_j)^{1/2}$	$\delta T_j^2 (u_j)^{1/2}$
X_{7j}	$\overline{\delta T_j} (\sec\theta - 1)$	$\delta q_j (\sec\theta\ u_j)^{1/2}$	$\delta q_j (\sec\theta\ u_j)^{1/2}$	$\delta q_j (u_j)^{1/2}$
X_{8j}	$\overline{p\ \delta T_j} (\sec\theta - 1)$	$\delta q_j^2 (\sec\theta\ u_j)^{1/2}$	$(\sec\theta - 1)(\sec\theta\ u_j)^{1/2}$	0
X_{9j}	$\delta T_j (\sec\theta - 1)$	$\delta T_j\ \delta q_j (\sec\theta\ u_j)^{1/2}$	$(\sec\theta - 1)^2 (\sec\theta\ u_j)^{1/2}$	0
Y_j	*1*	$(\sec\theta\ u_j)^{1/2}$	$(\sec\theta\ u_j)^{1/2}$	$\sec\theta (u_j)^{1/2}$

Table 1: Model predictors for mixed gases, and old and new versions for water vapour/ozone.

$$\overline{\delta T_j} = \frac{1}{p_j} \sum_{i=1}^{j} \delta T_i (p_i - p_{i-1}) \qquad u_j = \tfrac{1}{2}(q_j + q_{j-1})(p_j - p_{j-1})$$

$$\overline{p\ \delta T_j} = \frac{2}{p_j^2} \sum_{i=1}^{j} p_i \delta T_i (p_i - p_{i-1}) \qquad \overline{p\ \delta q_j} = \frac{2}{p_j^2} \sum_{i=1}^{j} p_i \delta q_i (p_i - p_{i-1})$$

$$\delta T_j = \tfrac{1}{2}\left(T_j - T_j^{ref} + T_{j-1} - T_{j-1}^{ref}\right) \qquad \delta q_j = \tfrac{1}{2}\left(q_j - q_j^{ref} + q_{j-1} - q_{j-1}^{ref}\right)$$

T_j and q_j are the temperature and specific humidity or ozone volume mixing ratio profiles. T_j^{ref} and q_j^{ref} are corresponding reference profiles (- the mean of a set of global water vapour and ozone profiles obtained from NESDIS has been used for this study).

Figure 1 : Comparison of old and new RT model

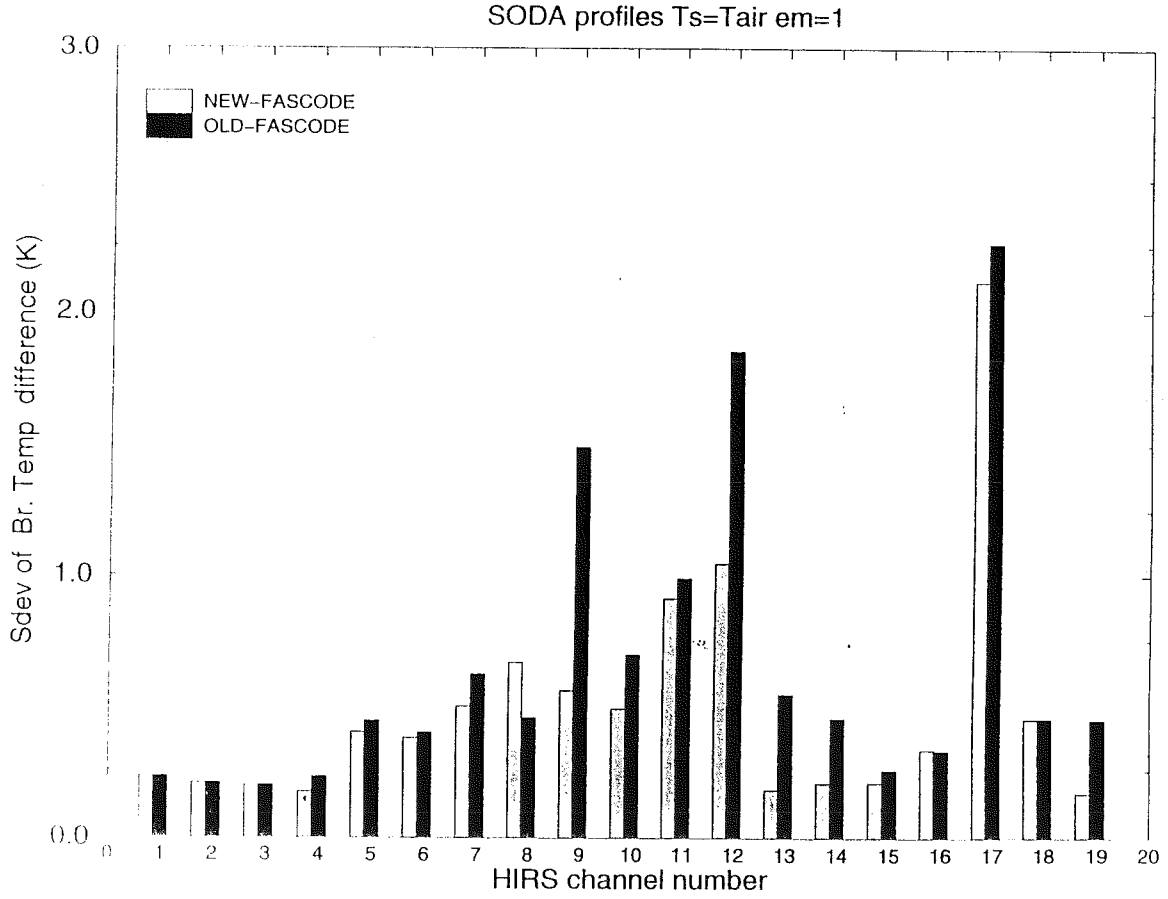
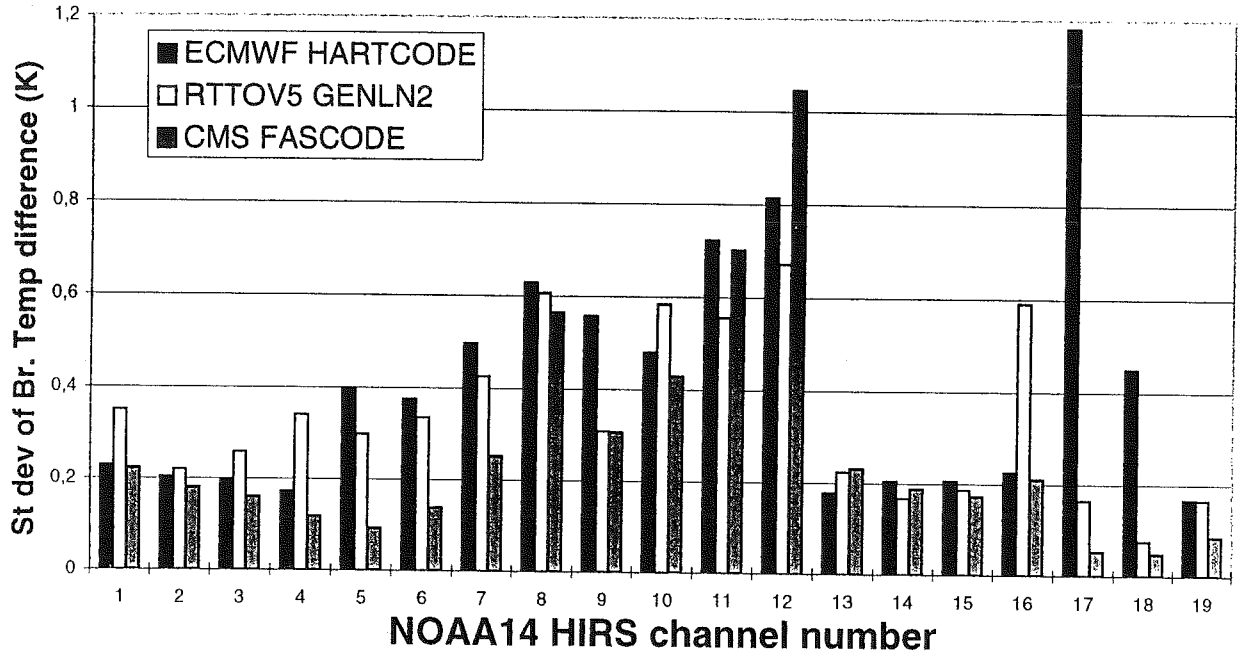


Figure 2 : **Comparison of Fast RT models with FASCODE on 528 SODA profiles**



VALIDATION OF MODEL OZONE CALCULATIONS USING OZONE PROFILE MEASUREMENTS, AS PART OF THE EU-SODA PROJECT

R.M.A. Timmermans

Atmospheric Chemistry Division, KNMI, Postbus 201, 3730 AE De Bilt, The Netherlands

ABSTRACT

At the European Centre for Medium-range Weather Forecasts (ECMWF) an 80 day ozone simulation is performed with the ECMWF weather prediction model. Ozone is included as prognostic variable and an ozone chemistry parameterisation is used to parameterise the sources and sinks of ozone. In this research the model performance is evaluated by comparing the model ozone calculations with ozone sonde measurements from the EASOE (European Arctic Stratospheric Ozone Experiment) campaign. The comparison shows a good description of the dynamical details by the model. However, the model overestimates the ozone concentration around the ozone maximum at high latitudes (and underestimates the ozone concentration around the ozone maximum in the tropics).

1. INTRODUCTION

The first task of the SODA project was to introduce ozone as a prognostic variable and include simplified ozone chemistry in the participating numerical weather prediction models and off-line transport models. One of these models is the ECMWF (European Centre for Medium-range Weather Forecasts) weather forecast model. To perform an initial test of the ozone transport in the ECMWF model, ozone was simulated for 80 days (December 1991- February 1992) coinciding with the ozone measuring campaign EASOE (European Arctic Stratospheric Ozone Experiment).

The aim of this research is to validate the model ozone calculations from the relaxation run using ozone profile measurements by ozone sondes from the EASOE campaign.

2. MODEL

The 80 day simulation at ECMWF was performed with the ECMWF weather forecast model, version T106/L43, which has 43 levels reaching into the mesosphere (0.05 hPa). In this simulation an ozone chemistry parameterisation was included but ozone was otherwise left to evolve freely. This experiment is referred to as a "relaxation" run. The dynamics of the model were kept close to reality by nudging towards ECMWF and UKMO temperature and wind fields. The sources and sinks of ozone are parameterised using the Cariolle parameterisation (Cariolle et al., 1993,1986). Heterogeneous chemistry is not included in these calculations. The ozone initial condition was achieved by spin-up from the zonally symmetric climatology of Fortuin and Langematz (Fortuin et al., 1994), normalised by TOMS total ozone to have correct total ozone on the starting date.

3. MEASUREMENTS

For the validation of the model ozone profiles, use is made of ozone profile measurements from the EASOE campaign (Pyle et al., 1994) which took place in the winter of 1991/92. The ozone sondes were launched from 22 European and 6 Canadian stations at latitudes ranging between 38.0 N and 82.5 N. ECC (Electrochemical Concentration Cell) sondes were used at all stations except for the three European stations with long records (Hohenpeissenberg, Payerne and Uccle) which used Brewer-Mast sondes, and Lindenberg where the old East German OSE-4 sonde was used until 31 December 1991 at which time there was a switch to ECC sondes.

4. RESULTS AND DISCUSSION

Figure 1 shows the ozone profiles at Goose Bay (53.3 N, 60.4 E) on 16 and 19 January 1992 from both model and observations. The structure in the ozone profile around 100 hPa on 16 January and the rapid increase of ozone at this same level, which is seen between the ozone soundings of 16 and 19 January, are well captured by the model. At this 100 hPa level, in the lower stratosphere, ozone is mainly controlled by transport processes and not so much by chemistry. The good agreement between model and observations is therefore an indication for the realistic model dynamics in this part of the atmosphere.

Figure 2 shows the mean bias (over day 31-80) between the model calculations and the sonde measurements in 4 different latitude bands. The bias has a

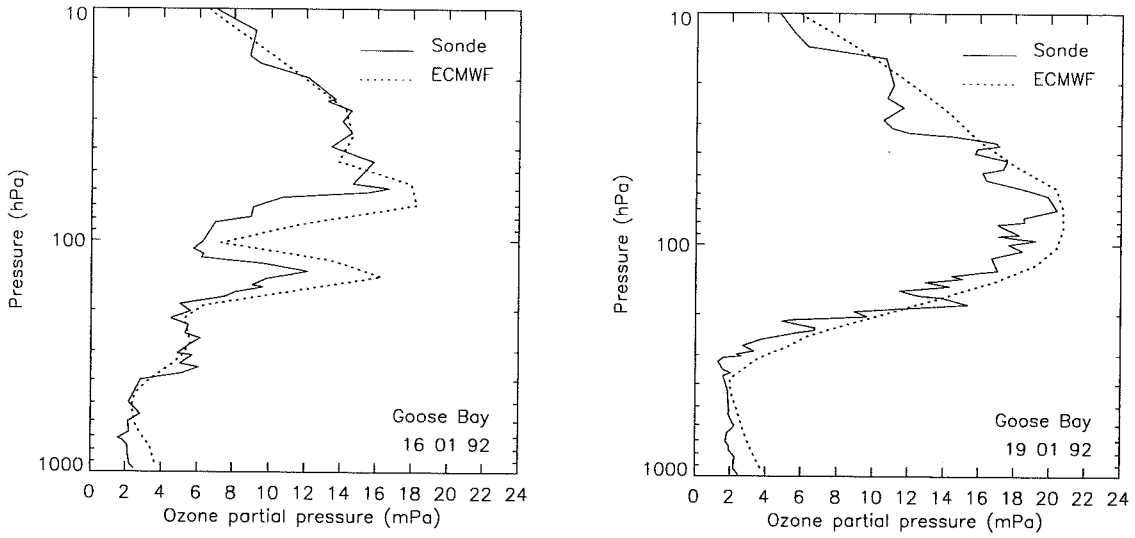


Figure 1. Ozone profile at Goose Bay (53.3 N, 60.4 W) on 16 and 19 January 1992. Shown are ozone sonde measurements (solid line) and model profiles (dashed line).

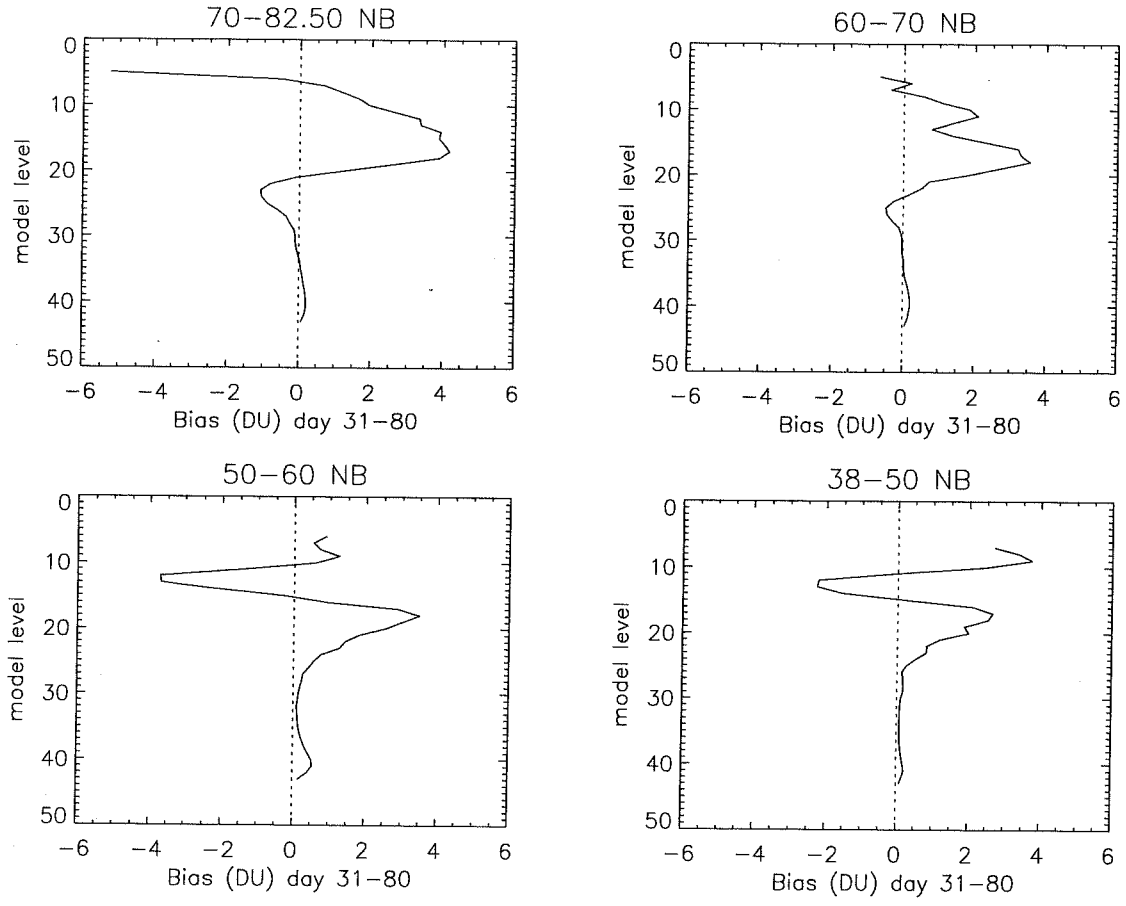


Figure 2. Latitude dependence of mean bias (ECMWF-sonde) in DU per model layer, for days 31-80 of simulation.

clear latitude dependence. In the two higher latitude bands the model overestimates the ozone in the region of the ozone maximum (layers 10-20). In the two other latitude bands a dipole structure can be seen which can indicate that the height of the ozone maximum in the model is lower than for the sondes.

These results are supported by a comparison of the model ozone with SBUV/2 satellite ozone data (SODA 1st annual report, 1997) which shows a positive bias between model and SBUV/2 at high latitudes and a negative bias in the tropics.

The cause of the bias is probably a combination of several factors, which need to be investigated. One of the major problems is that the chemistry parameterisation is designed for a climate model with very different dynamics than the ECMWF model. This leads to incorrect equilibrium coefficients in the ozone chemistry parameterisation. The large ozone destruction in the tropics seems to be a major error source. This ozone destruction is mainly caused by a difference between the temperature of the ECMWF model and the equilibrium temperature of the Cariolle parameterisation in combination with a high temperature sensitivity of the chemistry. Furthermore there is no surface deposition included in the parameterisation.

5. CONCLUSION

Model forecasts capture large ozone features (secondary ozone maxima) in the lower stratosphere surprisingly well for many of the soundings. There is however a bias between the model and the observations: the model ozone around the ozone maximum is too high at high latitudes (and too low in the tropics). The cause of the bias is probably a combination of several factors, most of them related to the Cariolle ozone chemistry parameterisation in the model. Note: In this experiment ozone was left to evolve freely, the bias will be smaller when data assimilation of ozone measurements is included in the model.

REFERENCES

Cariolle, D. 1983, The ozone budget in the stratosphere: Results of an one-dimensional photochemical model,Planet. Space Sci., Vol. 31, No. 9, pp. 1033-1052.

Cariolle, D., Dqu,M. 1986, Southern hemisphere medium-scale waves and total ozone disturbances in a spectral general circulation model, J. Geophys. Res., Vol. 91, No. D10, pp. 10,825-10,846.

Eskes, H.J., Piters, A.J.M., Levelt, P.F., Allaart, M.A.F., Kelder, H. 1998, Variational data assimilation: How to extract more information from GOME total ozone data", Earth Observation Quarterly - GOME special, vol. 58, p. 35-38.

Fortuin, J. P. F., Langematz, U. 1994, An update on the global ozone climatology and on concurrent ozone and temperature trends, Atmospheric Sensing and Modelling, R. P. Santer (Ed.), pp. 207-216, Proc. SPIE 2311.

Pyle, J.A., Harris, N.R.P., Farman, J.C., Arnold, F., Braathen, G., Cox, R.A., Faucon, P., Jones, R.L., Megie, G., O'Neill, A., Platt, U., Pommereau, J-P., Schmidt, U., Stordal, F. 1994, An overview of the EASOE campaign, Geophys. Res. Letts., Vol. 21, No. 13, pp. 1191-1194.

SODA 1995, Studies of Ozone Distributions based on Assimilated satellite measurements, project proposal to the European Union.

SODA 1997, Satellite Ozone Data Assimilation, 1st Annual European Union project report.

Chemistry

ASSIMILATION OF ODIN OZONE DATA IN THE CMC WEATHER FORECAST MODEL WITH CHEMISTRY: PRELIMINARY RESULTS

J. W. Kaminski[1], J. C. McConnell[1], G. Brunet[2], P. Gautier[2], S. Pellerin[2]

[1]Department of Earth and Atmospheric Science, York University, Toronto, Canada, M3J 1P3
[2]Meteorological Research Branch, Atmospheric Environment Service, Dorval, Canada, H9P 1J3.

ABSTRACT

A stratospheric chemistry module with gas phase and aerosol chemistry has been incorporated into the research version of the Canadian Meteorological Centre's (CMC) global forecast model which we call SEF+C. The number of levels has been increased and the top raised to about 1 mb. This model allows for the forecasting of ozone and other species over a height regime from the middle troposphere to the stratopause. The model is being used to assess the feasibility of assimilating retrieved ozone and other species profiles from the ODIN satellite using SEF+C. We anticipate that incorporation of the measured ozone profiles using the CMC's 3DVAR assimilation system will allow the information from the local satellite measurement to be transported to other locations thus assisting in the generation of a more realistic 4D ozone field. Herein, we briefly describe the chemical additions to the SEF model, the improvements to the physics software and the simplified simulation of retrieval from the ODIN (OSIRIS and SMR) satellite instruments. A nature run has been done for a 31 day period in June/July, 1995 preceded with a 40 day chemical "warm-up": for the nature run the objective analysis (OA) has been updated every 24 hours. Using the synthetic ozone data from the SEF+C nature run as input, with attendant retrieval errors, several assimilation scenarios using the CMC 3DVAR assimilation system have been done. These mimic data input from ODIN instruments (1) continuously as during a measurement intensive, (2) 1 day of measurements, followed by 2 days with no measurements, representing standard operating conditions, (3) daytime measurements only, typical of measurement conditions for OSIRIS during an intensive measuring period. These preliminary tests have been done using ozone as a passive tracer in the SEF model and the results are very encouraging. However, there appear to be problems with the dynamics in SEF in the stratosphere and ozone being lofted too high. Part of the problem is related to an initial poor choice of levels added to the SEF. A 2 km vertical resolution forecast run made recently gives better ozone properties from SEF+C. Further work is planned with the assimilation of ozone using SEF+C, i.e. the ozone will be assimilated directly into SEF+C: the chemical model will provide suitable constraints above (roughly) 10 mb where chemical time constants for ozone are shorter than transport time constants so that the evolution of the ozone field will be more appropriate than at present.

Key words: ODIN; stratosphere; chemistry; ozone; OSSE.

1. INTRODUCTION

Chemical data assimilation of satellite measurements is one means by which they can benefit from value-added processing, i.e. it may be possible to transform an asynoptic and sparse 4D field into a more dense and relatively more useful 4D field by using a dynamical model to translate the measured information to other points within the domain while constrained by the error field. However, for certain species in particular regions, such as ozone above 10 mb, the measured species may be under photochemical rather than dynamical control. Thus by adding a validated chemical model as an additional constraint it may be possible to control errors.

With this in mind, the research version of the Canadian Meteorological Centre's (CMC) forecast model, SEF (Ritchie & Beaudoin 1994) is being used to provide a basis for a chemical-dynamical data assimilation for ozone and other species. The preliminary version is being applied using CMC's 3DVAR assimilation system to test the viability of ingesting ozone profiles, in particular, those expected to be retrieved from the ODIN satellite mission. We first briefly describe the ODIN mission and experiments, then the model used for the nature run, followed by the 3 different OSSEs attempted and finally the results obtained so far.

2. THE ODIN MISSION

The ODIN satellite is a joint astronomy/aeronomy mission led by the Swedish Space Corporation that is due to be launched in the fall of 1999. It is a collaborative effort

between Sweden, Canada, Finland and France. There are two bore-sighted instruments aboard which will view the limb in the direction of the orbit. The prime instrument to be used by both the astronomy and aeronomy teams is the sub-millimetre radiometer, SMR. Canada's contribution is an optical spectrograph and IR imaging system, OSIRIS, which will be used just for aeronomical studies. ODIN aeronomy science objectives focus on springtime stratospheric polar ozone loss with an emphasis on the boreal vortex. But the vertical extent of the region that can be sensed by the instruments will permit studies of coupling between the stratosphere and mesosphere. Investigations of the formation of Polar Mesospheric Clouds (noctiluscent clouds) and PSCs will also be undertaken.

The mission time will be split equally between astronomy and aeronomy but will not be distributed evenly throughout the year. Because of the seasonal nature of the atmospheric objectives the aeronomy program will have several sequences devoted 100% to atmospheric measurements. In particular, there will be measurement intensives during the boreal and austral spring times when PSC processing and associated ozone loss is important. At most other times of the year there will be one day of aeronomy measurements followed by two days of astronomy measurements.

The SMR will measure thermal emission from the limb in 1 GHz bands near 119, 494, 557, and 575 GHz with 1 MHz resolution. Thus it will be possible to obtain data 24 hours per day. Inversion of the radiances will provide vertical profiles of an interesting suite of minor species such as ClO, CO, NO, NO_2, N_2O, HO_2, HO_2O_2, H_2O, HDO, H_2O^{18}, HNO_3, O_3^{48}, O_3^{50}: O_2, temperature and pressure will also be inferred. However, not all of this data will be available at all times due to the limited number of channels.

By applying the differential optical absorption spectroscopy (DOAS) technique to the spectra of the scattered sunlight from the limb in the 300-750 nm region and appropriate inversion techniques it will be possible to extract profile information on ozone, NO_2, BrO profiles: temperature and pressure and information on aerosols will also be available. Clearly, the OSIRIS data will only be for sunlit periods.

The limb will be scanned by nodding the whole satellite up and down. This will take approximately 2 minutes and this will provide a minimum of about 720 profiles of ozone per day from the SMR instrument if focussed continuously on ozone measurements. The UV-vis spectrograph will be able to measure ozone profiles using scattered light during daylight periods and the IR imager will be able to extend these measurements in height and time. Detailed forward modelling indicates that ozone retrievals should be accurate to $\sim$ 5-15%, depending on the altitude, between 15 and 50 km and this information has been incorporated into the OSSEs.

The Odin orbit will be a nearly circular 600 km sun synchronous orbit with an inclination of 97.8° and will cross the equator at local times of 6:00 (ascending) and 18:00 (descending). It is a 3-axis spacecraft stabilized with momentum wheels and gyros with a pointing accuracy of $\pm$ 15 arcsec while staring and $\pm$ 1.2 arc while scanning. The expected lifetime of the mission is two years.

3. THE CHEMICAL MODEL AND NATURE RUN

The basic dynamical model used is the Canadian Meteorological Centre's (CMC) SEF (spectral element finite) weather forecast model (e.g., Ritchie & Beaudoin 1994) with the top extended from 10 mb to 1 mb for the results shown here. The version used in these calculations uses a modified sigma coordinate and the model levels used for the nature run are those shown in Figure 1 of Gautier et al. (1999). The chosen levels proved rather coarse. However, in spite of this the final results are illuminating.

In order to use the model in the stratosphere there have been several improvements most of which are described in Kaminski et al. (1999). An improved semi-Lagrangian transport scheme which is monotonic and positive definite was been added and tested (Ritchie 1991, and private communication, 1998).

Above about 10 mb the ozone distribution is controlled more directly by chemistry than dynamics and in order to capture this behaviour it is important to have a chemical model. Thus a comprehensive chemical scheme (de Grandpré et al. 1997) with 36 prognostic species and $\sim$ 130 reactions has been added which is also described in Kaminski et al. (1999). For these tests discussed herein, the chemistry is on line but not interactive. The rate constant data and photolysis cross sections are taken from DeMore et al. (1994) and references therein except where noted. The hydrolysis reactions of N_2O_5 and $BrONO_2$ have been included as well as a background field of stratospheric aerosols.

The method of solution that we adopt for this suite of equations has been tested in a box model and discussed by Sandilands & McConnell (1997) and its implementation in the Canadian Middle Atmosphere Model (Beagley et al. 1997) has been described by de Grandpré et al. (1997). The photolysis values are computed by a table lookup method.

Operator splitting is used to calculate the new chemical fields. Currently every species is transported, even those that are in local photochemical steady state. This is perhaps a little inefficient but we feel that it is more than compensated by the simpler logistics of handling each transported species in the same manner.

The meteorological conditions applied are a combination of CMC OA data up to 700 mb while above this level UKMO OA data (Swinbank & O'Neil 1994) has been used. The model was first run for 40 days of chemical spin-up using four 10-day forecasts followed by a 31 day integration from 26th June to 27th July, 1995, with the meteorological conditions updated every day. The chemical species mixing ratios are not altered during the ingestion of the OA, i.e. it is solely a meteorological restart. Chemical initial conditions for the long lived species have been taken from the CMAM model or from a 2D model as necessary.

Chemistry is applied from 500 mb to the top of the model and for the nature run the SEF+C has been run at T63 (128x64). The resolution adopted has been largely limited by the number of chemical species solved for.

The two-week zonal and temporal average of the ozone mixing ratios, in parts per million by mass (ppmm), from the nature run are presented in Figure 1a while Figure 2a shows the standard deviation (SD) for the same time period. For the following discussion we have used this latter quantity as one measure of temporal and spatial variability of the solution. We have also used a more comprehensive empirical orthogonal function (EOF) analysis for these results which reveals similar behaviour (G. Brunet, private communication, 1998).

The ozone mixing ratios shown in Figure 1a are quite reasonable below 10 mb. The latitudinal location of the ozone peak at the 10 mb level is roughly correct location for the Sun in the northern hemisphere. The SDs are $\sim$ few percent. However, above 10 mb, because of the poor vertical resolution the peak in the ozone mixing ratio is spread between 1 and 10 mb whereas it should peak about 10 mb (cf., de Grandpré et al. 1997). This can be seen in Figure 3 which shows the results from a 12 day forecast of SEF+C with $\sim$ 2 km resolution above about 20 km. The ozone peak is placed at the correct height and the general structure is correct. This resolution will be used for future work.

4. THE OBSERVATIONAL SYSTEM SIMULATION EXPERIMENTS

For these observational system simulation experiments (OSSE) we have used the SEF+C to provide the ozone mixing ratios (omr) for the nature run and thus act as 'atmospheric truth'. The omr information is extracted from the nature run by 'flying' the ODIN satellite over the 4D field and allowing it to scan the model limb. The retrieval is simplified in this case and is done by taking the vertical profile of omr's from the horizontal location of the mid-point of the scan. The error profile and confidence limits adopted are based on the geometrical factors associated with the observations as well as an assessment of retrieval errors from the the OSIRIS and SMR retrievals (Ian McDade and Donal Murtagh, private communication, 1998). Finally, the synthetic ozone satellite data is assimilated into SEF every 6 hours, run at T119, using 3DVAR (Gautier et al. 1999). For these first experiments ozone was assimilated as a passive tracer. In future work the assimilation will be into SEF+C.

We report on three different OSSEs that mimic the possible observing constraints of the ODIN mission. Our first OSSE, OSSE A, was based on profile information being available for 24 hours each day for the assimilation period of one month. This is equivalent to having the data from the SMR and OSIRIS available during the measurement intensives. OSSE B is representative of the non-intensive measurement periods when the SMR and OSIRIS ozone data will only be available every third day, the other two days being directed towards astronomical observations.

As has been noted above the SMR may not always be assigned to measure ozone. However, the OSIRIS will still be measuring ozone but only during the day lit periods. With the assumption of a measurement intensive, i.e., data every day, this constitutes the scenario for OSSE C where ozone profiles will only be associated during the day lit period. Since the assimilation is during July there will be no sunlight over the austral pole and this will be reflected in the assimilated data.

Figure 1b shows the two week zonally averaged profile for OSSE A where it can be seen that the assimilation of the profiles is constraining the passive tracer to resemble the main features of the nature ozone distribution. One of the features that is quite similar is the location and magnitude of the peak mixing ratio. Also a comparison of Figure 2b with Figure 2a shows that the SD distribution is quite similar so that individual zonal slices should be quite similar. We return to this later.

The zonally averaged ozone for OSSE B is presented in Figure 1c. Comparison with Figure 1b shows that the ingestion of data every third day produces, not unexpectedly, results inferior to that of OSSE A. However, the main features in these average results are captured, although the peak region is now more diffuse. The biggest differences appear in the SD contours shown in Figure 2c. The region of maximum variability about 5 mb and 30°N is smaller and the variability around 20 mb has increased by a factor of 2 to 4 depending on the latitude. However, we note that the magnitude of the variation is still quite small, $\sim$ 3-5%.

The results for OSSE C zonally and temporally averaged ozone mixing ratios are shown in Figure 1d and the associated SDs in Figure 2d. This assimilation appears quite to that for OSSE A, even to the pattern of the SD exhibited.

We have noted above that the SD plots are, to some degree, a measure of the spatial and temporal variability and the more similar they are for the various OSSEs and the nature run the better the agreement is likely to be. We also noted that the SDs were only $\sim$ a few percent. Thus in Figure 4 we show a horizontal snapshot of ozone mixing ratios in parts per million by volume (ppmv) at 22 mb. Bearing in mind that the assimilation of the observations is at higher resolution that that of the nature run (T119 vs T63) we see that there is a great deal of similarity in the structures in the snapshots. In a qualitative sense there is more similarity between the OSSE A and OSSE C than for OSSE B. Presumably the major differences between the nature run and OSSE A are attributable to the higher dynamical resolution of the OSSE.

5. CONCLUSIONS AND FUTURE WORK

These preliminary OSSE experiments have fulfilled two main purposes. One, they have tested the CMC's 3DVAR system for ingesting ozone profiles using synthetic satellite data. Second, they have shown that the ozone data from ozone either from measurement intensives or from

the 'standard' operating conditions either using the SMR or OSIRIS instruments will be able to be usefully assimilated which will be an invaluable tool for extending the impact of the measurements.

Regarding future work: One of our first goals will be to redo the nature run with improved vertical resolution and repeat the OSSEs discussed above. Additionally, it would be useful to repeat the assimilation with the column ozone derived from the nature run.

Simultaneously we wish to proceed to our original goal of using the SEF+C for the assimilation process and also address the possibility of assimilating other species that will be measured by the ODIN satellite such as N_2O, CH_4, HNO_3, H_2O etc.

ACKNOWLEDGMENTS

The authors wish to thank Wade Sandilands, Lori Neary, Hal Ritchie, Christiane Beaudoin and Louis Garand in getting SEF+C operational and Ian McDade and Donal Murtagh for estimates of the errors associated with the ODIN ozone retrievals. JCMcC and JK wish to thanks NSERC CSP grant for ODIN and MOPITT for support for this work.

REFERENCES

Beagley, S. R., J. de Grandpré, J. N. Koshyk, N. A. McFarlane, and T. G. Shepherd, Radiative-dynamical climatology of the first-generation Canadian Middle Atmosphere Model, *Atmos.-Ocean, 35,* 293–331, 1997.

Gautier, P., G. Brunet, S. Pellerin and P.-A. Michelangeli, The impact of vertical transport on ozone analyses based on total ozone measurements, Workshop on Chemical Data Assimilation, KNMI, de Bilt, 9-10 December, 1999, This issue.

de Grandpré. J., J. W. Sandilands, J. C. McConnell, S. R. Beagley, P. Croteau, M. Y. Danilin, Canadian Middle Atmosphere Climate Model: Preliminary Chemistry Results, *Atmos. Ocean, 35,* 385–431, 1997.

DeMore, W. B., S. P. Sander, D. M. Golden, R. F. Hampson, M. J. Kurylo, C. J. Howard, A. R. Ravishankara, C. E. Kolb, M. J. Molina, Chemical Kinetics and Photochemical Data for use in Stratospheric Modeling, JPL Publication 94-26, Pasadena, CA, 1994.

Kaminski, J. W., et al. On-line chemistry in the CMC weather forecast model: a feasibility study, to be submitted to the Can. J. Phys., Feb., 1999.

Ritchie, H., 1991: Application of the semi-Lagrangian method to a multilevel spectral primitive equations model. *Quart. J.R. Met. Soc., 117,* 91-106, 1991.

Ritchie, H. R. and C. Beaudoin, Approximations and sensitivity experiments with a baroclinic semi-Lagrangian spectral model, *Mon. Weath. Rev., 122,* 2391–2399, 1994.

Sandilands, J. W., J. C. McConnell, Evaluation of a Reduced-Jacobian Numerical Chemical Solver, *J. Geophys. Res, 102,* 19,703–19,087, 1997.

Swinbank, R. and A. O'Neil, A stratospheric-tropospheric data assimilation system, *Mon. Weath. Rev., 122,* 686-702, 1994.

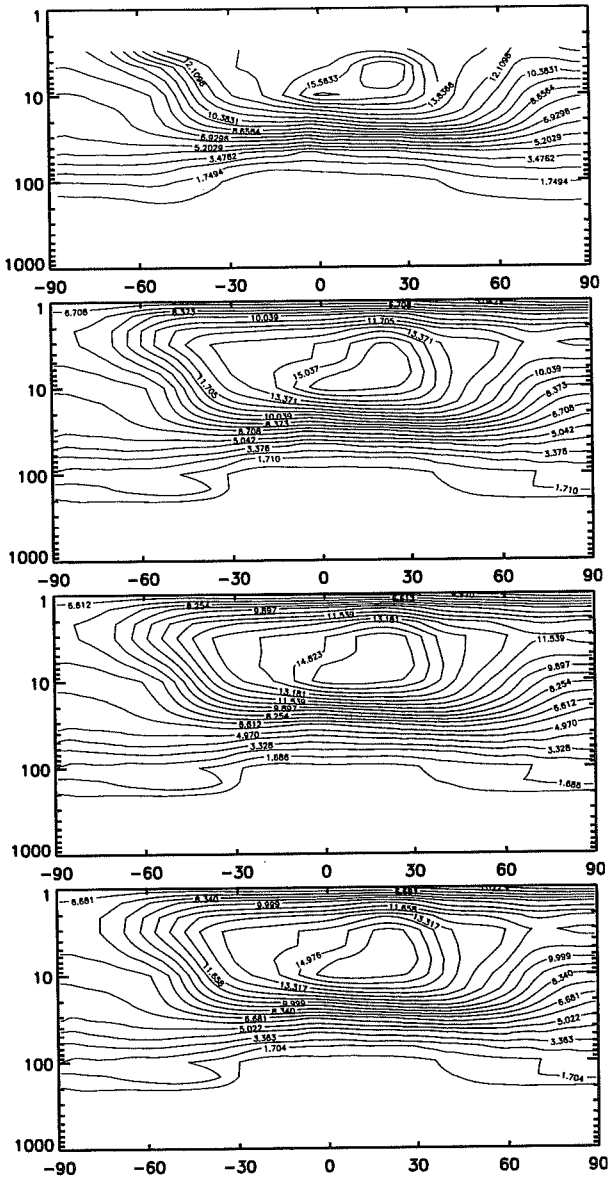


Figure 1. Zonally and temporally averaged ozone mixing ratios for the last 2 weeks of the assimilation period (ppmm). (a) The nature run, (b) OSSE A, (c) OSSE B and (d) OSSE C. See text for details.

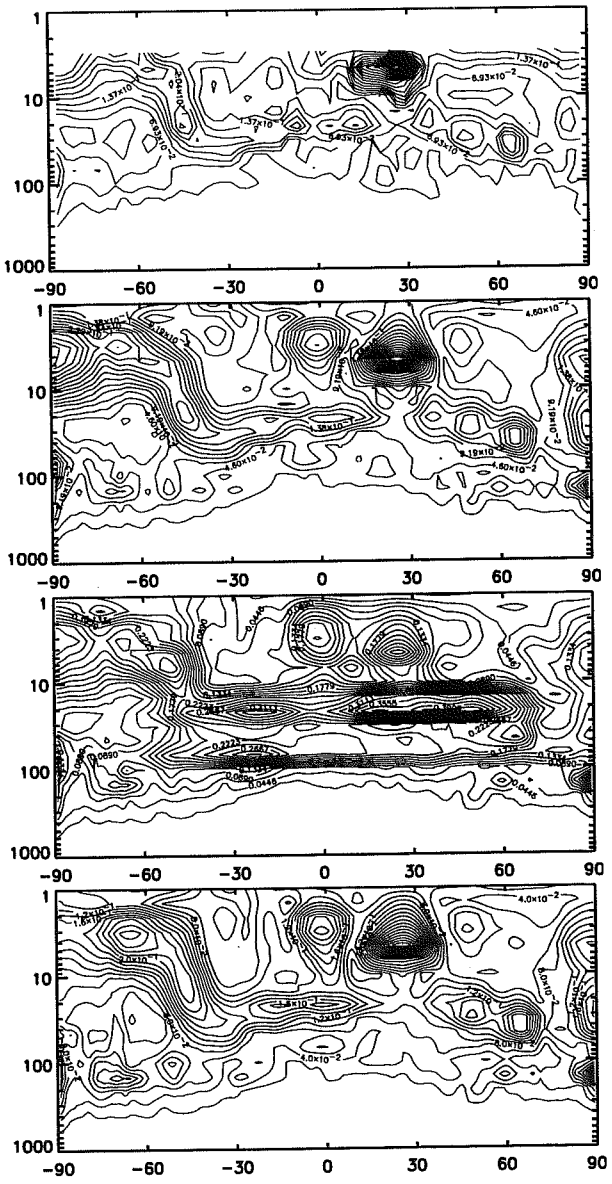


Figure 2. Standard deviations associated with the ozone profiles shown in Figure 1 for the last 2 weeks of the assimilation period (ppmm). (a) The nature run, (b) OSSE A, (c) OSSE B and (d) OSSE C. See text for details.

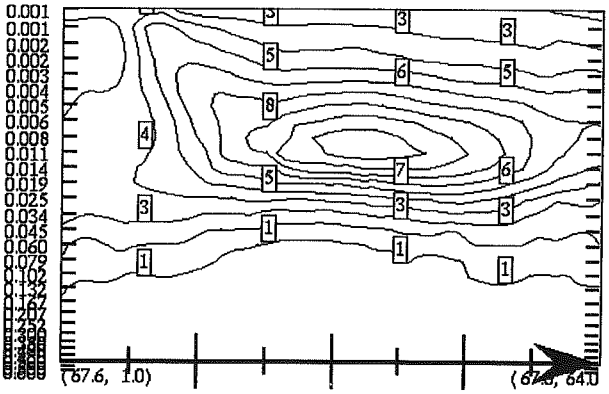


Figure 3. Zonal snap shot of ozone mixing ratios after a 12 day forecast run using SEF+C at 2 km stratospheric resolution. Ozone mixing ratios in ppmv.

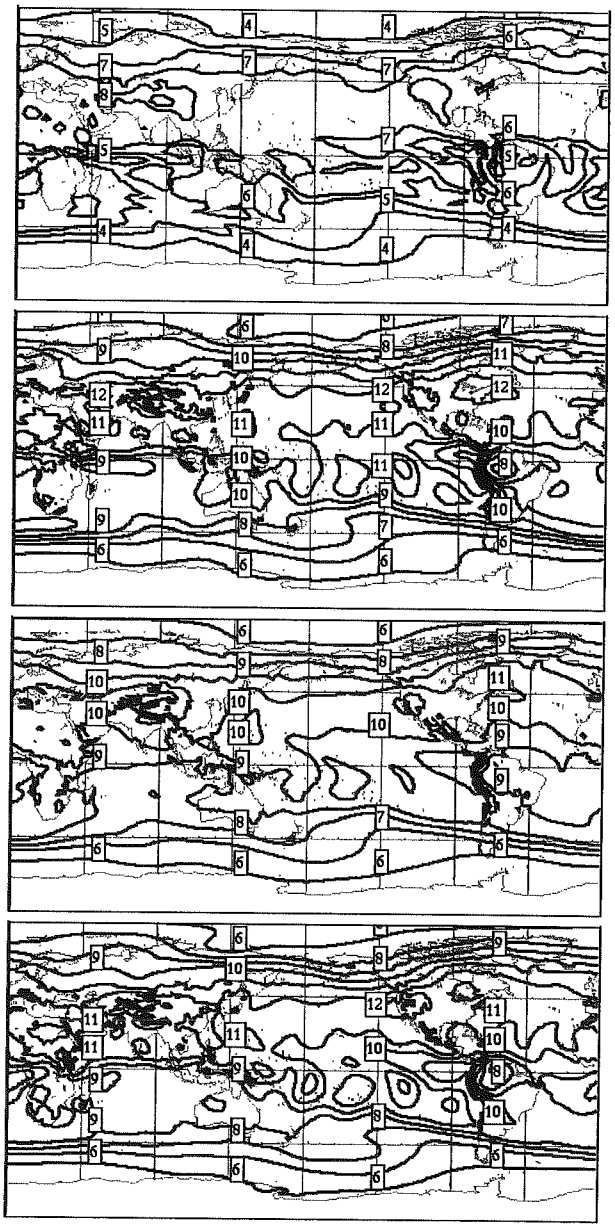


Figure 4. Horizontal slice of ozone mixing ratio in ppmv for the last data of the assimilation period at 20 mb for (a) the nature run, (b) OSSE A, (c) OSSE B and (d) OSSE C.

Dynamics

INCREASED STRATOSOPHERIC RESOLUTION IN THE ECMWF FORECASTING SYSTEM

Agathe Untch, Adrian Simmons, Mariano Hortal, Christian Jakob and colleagues

European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, Berks, RG2 9AX, UK
(email: adrian.simmons@ecmwf.int; tel: (44)-118-949-9700; fax: (44)-118-986-9450)

Abstract

ECMWF has been developing versions of its model and data assimilation system with finer and more extensive vertical resolution in both the stratosphere and the planetary boundary layer. In particular, 50- and 60-level versions (differing primarily in their boundary-layer resolution) have been extensively tested in data-assimilation, forecasting and climate-simulation experiments. The 50-level version is currently undergoing near-real-time trials with a view to operational implementation early in 1999, and replacement by the 60-level version is expected later in the year. Some stratospheric aspects of the results of the experimentation are presented here, showing in particular much improved forecasts in the lower stratosphere and much improved simulations of stratospheric humidity, in part due to inclusion of a simple representation of moistening due to methane oxidation. The improved stratospheric resolution is being utilized in development of the system for ozone data assimilation and forecasting.

1. Introduction

The operational forecasting model of ECMWF uses a hybrid vertical coordinate that reduces smoothly from a terrain-following coordinate in the lower troposphere to a pressure coordinate in the lower stratosphere (Simmons and Burridge, 1981; Simmons and Strüfing, 1983). Since September 1991, a 31-level resolution has been used operationally (Ritchie et al., 1995), with levels distributed as shown in the left-hand portion of Fig. 1. The top four levels are located at pressures of exactly 10, 30, 50 and 70hPa, and the pressures at the next two levels depend only slightly on surface pressure, being within 0.1hPa of 90hPa and 1hPa of 110hPa.

A 50-level version of the model, illustrated in the right-hand portion of Fig. 1, is currently undergoing pre-operational trials. The distribution of levels is identical to that of the 31-level version below 150hPa, and levels between 60 and 5hPa are close to equally-distributed in height with a spacing of 1.5km. The spacing increases above the 5hPa level and the top level is at 0.1hPa.

Earlier experimental work was carried out using a 43-level resolution, with a top level at 0.05hPa and a layer-spacing which increased much more rapidly with increasing height in the lower and middle stratosphere. This version performed satisfactorily in initial data assimilation and forecasting tests, but was found in extended-range simulations to form a quite unrealistically strong and persistent westerly jet in the tropical stratosphere. It was this that led to the development of the 50-level version of the forecasting system in which the more-uniform spacing of levels in the lower and middle stratosphere was sufficient to avoid the persistent strong westerlies. In the meantime, work on ozone modelling and assimilation (discussed by Hólm and Timmermans at this workshop) has continued using the 43-level resolution.

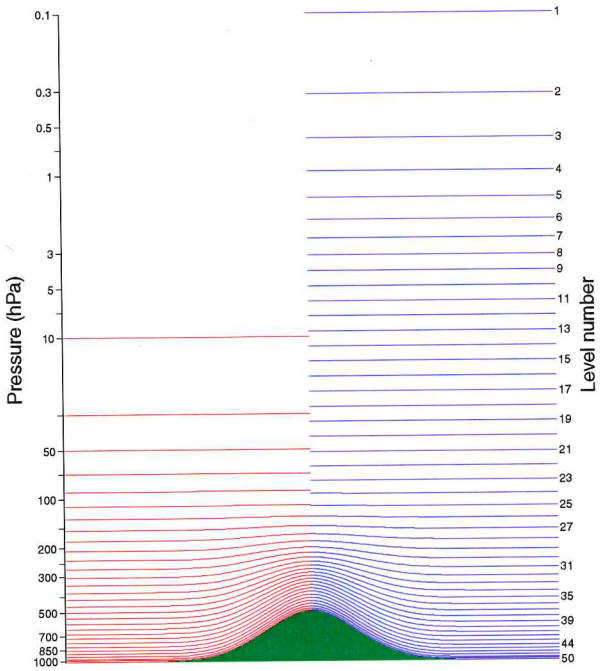


Fig. 1. The distribution of the full model levels at which wind, temperature and humidity are represented, for 31-level (left) and 50-level (right) vertical resolutions, plotted for surface pressures which vary from 1013.25 to 500 hPa.

Independently, a study of increased vertical resolution in the planetary boundary layer and lower troposphere has yielded promising results (Teixeira, 1999). This has led to construction of a 60-level version with stratospheric resolution similar to that of the 50-level version (but with a minor rearrangement of levels), and with the main resolution increase in the planetary boundary layer and immediately above. This will form the basis for further development, as discussed in the concluding section of this paper.

2. Revisions to the forecasting system

A number of problems had to be addressed in developing the versions of the forecasting system with improved stratospheric resolution, and these versions have benefited also from some other developments of the forecasting system. Several of the beneficial changes have already been introduced into the operational 31-level version of the system. These include revisions of the radiative parametrization, the two-time-level semi-

Lagrangian advection scheme, the calculation of saturation specific humidity at low temperatures, the analysis of humidity and the climatological distribution of ozone.

Several other changes have been included in the experimental versions. Rayleigh friction has been introduced at the uppermost few model levels to ensure a broadly realistic simulation of the mean circulation close to the stratopause. Changes have been made to parameters in the semi-implicit scheme to remove noise arising from weak computational instabilities. The non-linear normal-mode initialization used in the 31-level version at several points within the incremental four-dimensional variational data assimilation (4D-Var; Rabier et al., 1997) has been switched off to avoid large initialization changes at high levels arising from the large amplitudes of the normal modes at low pressures. In addition, a representation of the high-level moisture source due to methane oxidation has been introduced. The specification of this is set out in section 4.

It was also necessary to compute, for each new vertical resolution, new sets of balance operators and background error covariances for the data assimilation (Bouttier et al., 1997). These are derived in general from sets of differences between two- and one-day forecasts verifying at the same time. The first set of 50-level forecasts used for these calculations was based on initial conditions formed by merging 31-level ECMWF analyses with UKMO stratospheric analyses (Swinbank and O'Neill, 1994), with upper-level errors reduced to counter effects of incompatibilities between the ECMWF model and the UKMO analyses. These background statistics were used for two periods of data assimilation and forecasts. A revised set of statistics was then computed from these 50-level forecasts.

3. General performance of 50-level system

Data assimilation has been carried out continuously since 15 May 1998 with one for or other form of the 50-level system, and 50-level assimilations have also been run for December 1997 and most of January 1998.

Results show clearly that the enhanced stratospheric resolution provides substantially better analyses and forecasts at stratospheric levels up to 10hPa, where comparison can be made with results from the standard 31-level system. An example of fits of 30hPa temperatures and winds to radiosonde measurements is presented in Fig. 2. The fits are computed as an average over 165 days of experimentation run with the original background statistics; in experiments carried out over a subset of these cases, slightly better stratospheric fits are found for the 50-level system using the revised background statistics. The 50-level analyses (day 0) and forecasts (throughout the range to day 10) can be seen in Fig. 2 to be clearly better than the 31-level analyses and forecasts as judged by the fit to radiosonde data. Similar plots for other areas and other stratospheric levels, and for forecasts verified against analyses, confirm the generally superior performance of the 50-level system. The only notable exception is in the tropical stratosphere, where the 50-level forecasts exhibit a larger growth of bias in temperature, although temperature analyses (and wind analyses and forecasts) are nevertheless better from the 50-level than from the 31-level system.

Synoptic assessment of the stratospheric forecasts for the northern hemisphere winter period confirms the better performance of the 50-level system. Fig.3 presents an example. The upper panel shows the analysis at 10hPa from the 50-level system for a day on which there was a marked elongation of the polar vortex. The corresponding analysis from the 31-level system is largely similar. The seven-day forecast verifying on this day from the 50-level system captured quite accurately the elongation of the vortex, whereas the 31-level system predicted an erroneous strong vortex centred over northern Canada and a weaker circulation over Siberia. This is a pronounced case, but examination of the whole winter period shows that the stratospheric forecasts from the 50-level version (or indeed, the rather similar forecasts from the 60-level version) are almost invariably better synoptically than those from the 31-level version in depicting features such as the strength and position of the Aleutian high or the shape and orientation of the vortex. Dramatic synoptic improvements are also being seen in the current near-real-time trial of the 50-level system.

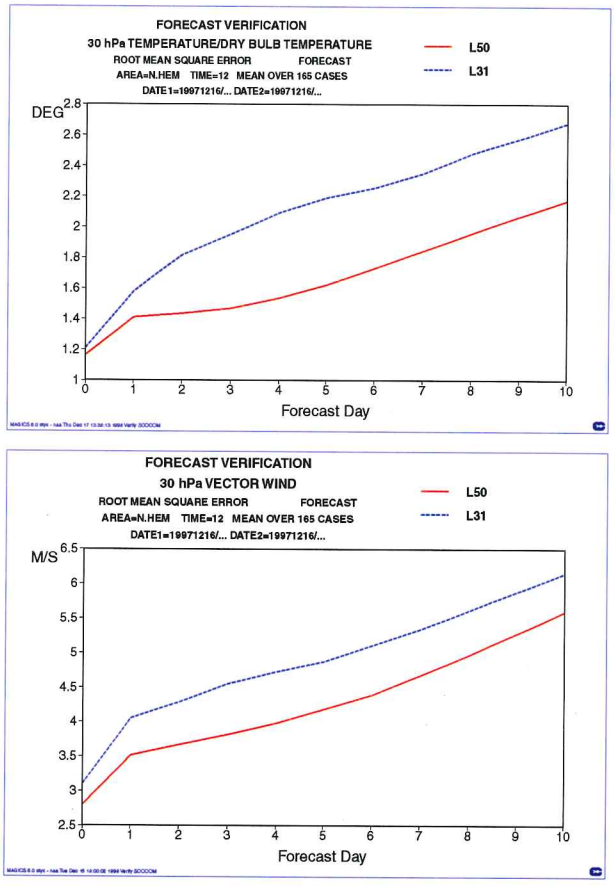


Fig.2. Root-mean-square errors of 30hPa temperature and vector-wind analyses (day 0) and forecasts (days 1 to 10) verified against radiosonde measurements over the extratropical northern hemisphere, averaged over a set of 165 cases run with 50-level (red, solid) and 31-level (blue, dotted) vertical resolutions.

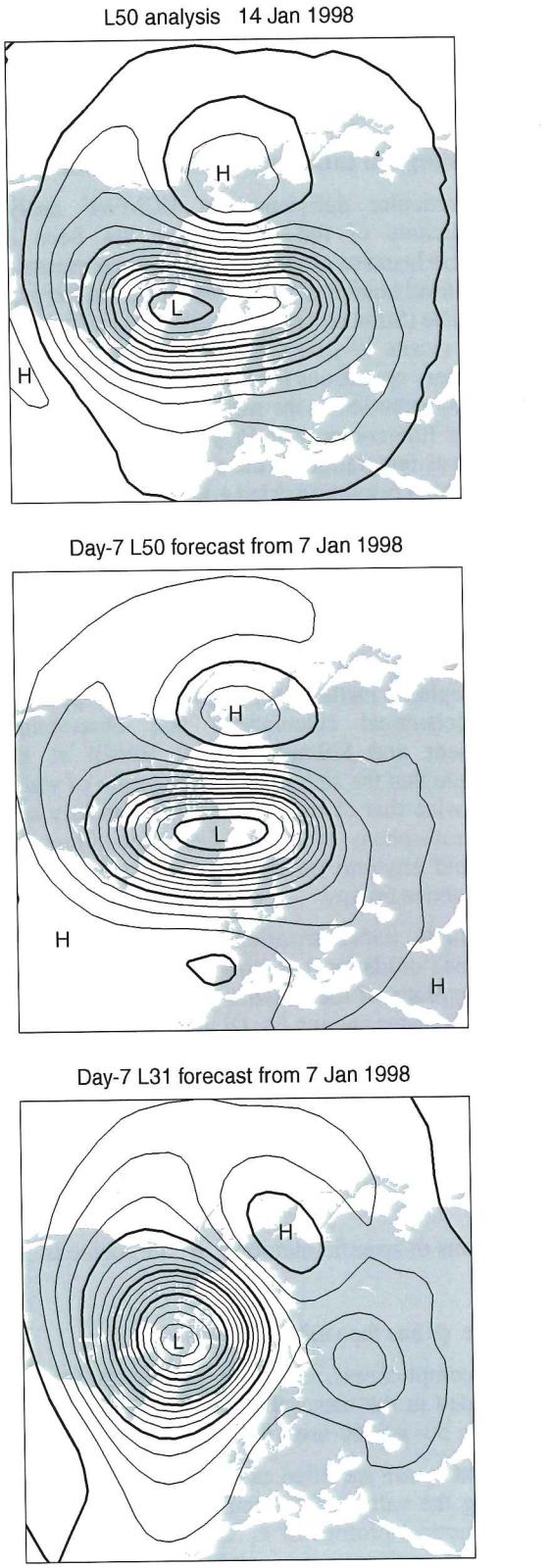


Fig. 3. The 10hPa height analysis for 12UTC 14 January 1998 from the 50-level version (upper), and seven-day forecasts verifying on this date from the 50-level (middle) and 31-level (lower) versions of the forecasting system. The contour interval is 160m.

A comparison of the zonal-mean zonal flow for October 1998 from the 50-level analyses with that from the UKMO stratospheric assimilation system is presented in Fig. 4. Differences in the troposphere should be disregarded, as the averaging was performed on model coordinate surfaces (terrain-following in the troposphere) for the ECMWF analyses and on pressure surfaces for the UKMO analyses. There is a reassuring agreement between the two sets of stratospheric analyses in the extratropics, where the only difference of note occurs close to the stratopause in the southern hemisphere. Differences in the tropics are, however, more marked. The easterly maximum located at about 40hPa is more than $5ms^{-1}$ stronger in the mean ECMWF analysis, and this analysis has a westerly maximum stronger than $25ms^{-1}$ above 10hPa, in a region where the UKMO analysis has easterly flow. It is not surprising that differences are largest in this region, both because of the difficulty in modelling the quasi-biennial oscillation (QBO) and because of the paucity of radiosonde wind observations and limited extent to which the wind analyses in the tropics can be controlled by assimilation of satellite radiance measurements. There is evidently a need for further study of the realism of the tropical stratospheric wind analyses from the new system, particularly to examine performance over the complete cycle of the QBO.

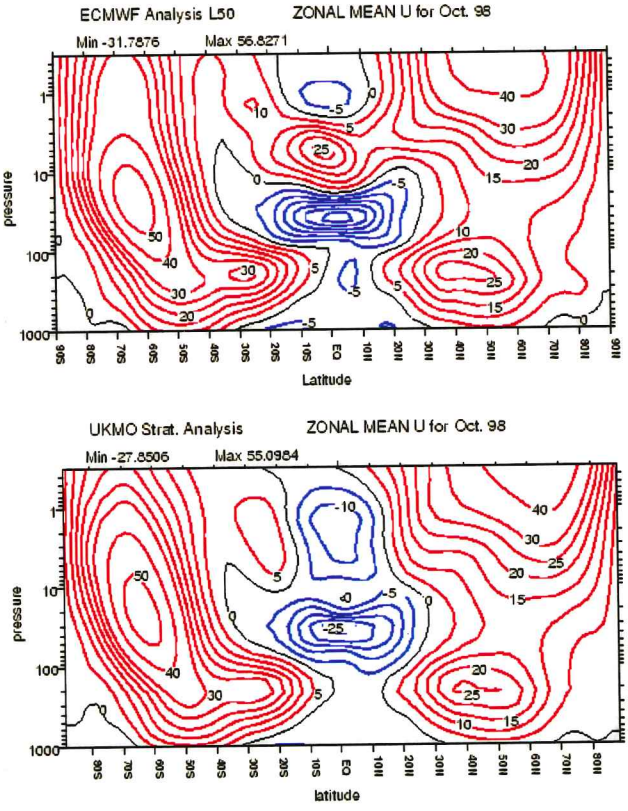


Fig. 4. Meridional cross-sections of zonal-mean zonal wind averaged for the month of October 1998 from experimental 50-level ECMWF analyses (upper) and from the stratospheric analysis system of UKMO (lower). Red contours denote westerlies, blue contours easterlies, and the contour interval is $5ms^{-1}$. Tropospheric differences should be disregarded.

Fig. 5 presents objective verification of 500hPa height forecasts from the 50- and 31-levels systems. It is based on the extensive sets of forecasts available with the original background statistics; the average is over the 154 forecasts that can be verified against analyses produced using the same model version as used to provide initial conditions. The 50-level version gives better mean verification scores throughout the medium range. Although differences may appear small to the general reader, they are seen to be quite significant when interpreted in the context of the long-term evolution of forecast skill, which has seen an increase by the order of one day per ten years in the forecast range at which the anomaly correlation falls below 60%.

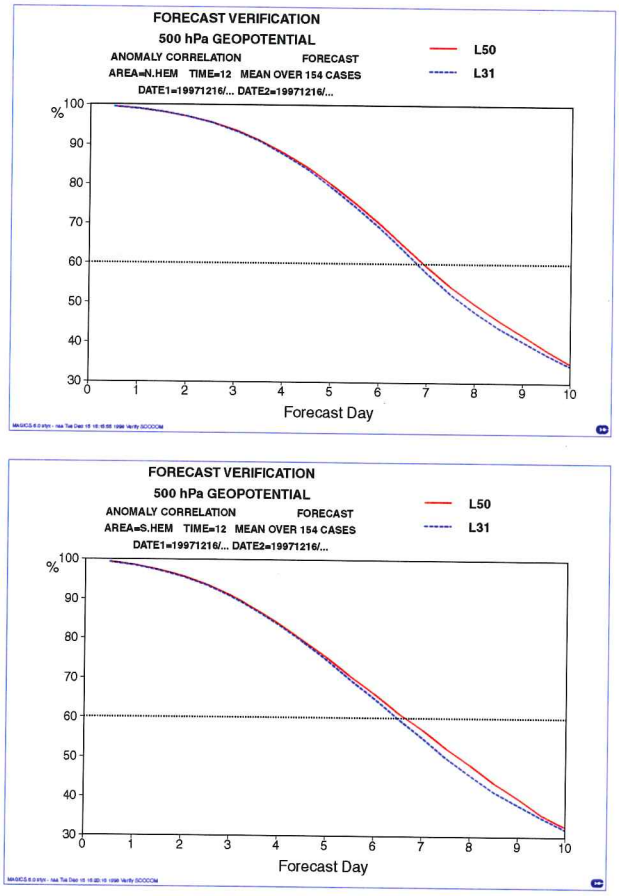


Fig. 5. Anomaly correlations of 500hPa height averaged over 154 cases run with 50-level (red, solid) and 31-level (blue, dotted) vertical resolutions, for the extratropical northern (upper) and southern (lower) hemispheres.

Repeating a subset of the forecasts included in Fig. 5 with the revised background statistics gives very similar hemispheric verification scores for the 50-level system, but some marked regional differences. In particular, an improvement found over Europe in the original 50-level experiments largely disappears with the revised background statistics, whereas the converse is the case for North America. Further study of the specification of the background statistics is evidently needed and is planned (see later). For the time being, the 50-level system using revised background statistics that is undergoing near-real-time parallel trials is expected to yield substantial improvements in operational stratospheric products and some small improvement in tropospheric products, with further improvements expected from additional work to be carried out after operational implementation.

4. A simple treatment of methane oxidation

A particular deficiency of ECMWF analyses and simulations of the stratosphere has been excessive dryness because of the absence of a representation of the high-level source of water vapour due to the oxidation of methane (Simmons et al., 1999). This has been confirmed by a recent 19-year simulation using observed surface boundary conditions for the years 1979-1997, carried out as a contribution to the second phase of the Atmospheric Model Intercomparison Project (AMIP; Gates, 1992). A 50-level integration of the model carried out using T63 horizontal resolution dried gradually over the first decade of integration, leading to upper stratospheric specific humidities around 1.7-1.8mg/kg that were maintained over the second half of the integration period. This is not unreasonable in the absence of methane oxidation as the asymptotic level corresponds to a mixing ratio of around 2.8ppmv, giving a value of 6.2ppmv when twice the tropospheric methane mixing ratio of 1.7ppmv is added. Photochemical calculations and observations (e.g. Brasseur and Solomon, 1984; Bithell et al., 1994) indicate that the sum of the mixing ratio of water vapour and twice that of methane is approximately constant in the stratosphere (away from the region of precipitation in the cold Antarctic polar night), with a value close to or a little above 6ppmv.

A simple parametrization of the moistening due to methane oxidation has thus been developed. The basic assumptions of the scheme are that the volume mixing ratio of water vapour $[H_2O]$ increases at rate $2k_1[CH_4]$ and that there is a steady balance between the mixing ratios of methane and water vapour:

$$2[CH_4] + [H_2O] = 6ppmv$$

The rate of increase of water vapour (in ppmv) is then

$$k_1(6 - [H_2O])$$

In terms of specific humidity, q, the source is

$$k_1(Q - q)$$

where Q has the value 3.75mg/kg.

For completeness, an extra photolysis term $-k_2 q$ is included in the mesosphere, although it has little effect for the 50- and 60-level resolutions discussed here.

k_1 and k_2 are specified as functions of pressure, with k_1 taking the value (100 days)$^{-1}$ at pressures below 1hPa. The vertical profiles of k_1 and k_2 are chosen such that the dependence on altitude of the combined photochemical lifetime, $(k_1 + k_2)^{-1}$, shown in Fig. 6, is similar to that presented by Brasseur and Solomon(1984). The slow time scale of the process in the stratosphere enables latitudinal and temporal variations in relaxation rate to be neglected.

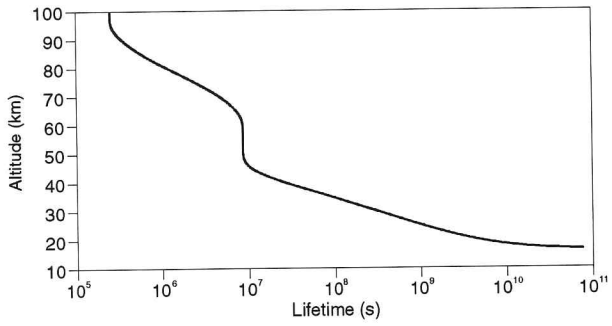


Fig. 6. Combined photochemical lifetime $(k_1 + k_2)^{-1}$

5. Stratospheric humidity in simulations and analyses

A new AMIP simulation has been run including the parametrization of methane oxidation. The 60-level vertical resolution was used for this run. After an initial adjustment period of several years, the stratospheric humidity in this simulation exhibits a fairly regular annual cycle. Latitude/pressure sections showing 15-year means for January, April, July and October are presented in Fig. 7. These sections may be compared, for example, with the sections compiled from observations from the HALOE and MLS instruments on the UARS satellite by Randel et al.(1998). Such comparison indicates a generally satisfactory performance of the model, which successfully captures the dryness both of air entering the stratosphere in the tropics in the boreal winter and of air in the cold Antarctic lower stratosphere in austral winter and spring. The consequences of a general ascent of relatively dry air throughout the tropical stratosphere, and the net high-latitude wintertime descent of air moistened by methane oxidation can be clearly seen. The model is unable to represent a drying of the winter polar upper stratosphere seen in the observations as it does not have the resolution to represent correctly the descent of dry air from the mesosphere into this region. The data from the simulation was stored only on standard pressure levels, which makes it difficult to track precisely the upward transfer of the annual cycle in water vapour (the "tape-recorder" effect discussed by Mote et al., 1996), but the simulation is generally drier in the tropical middle stratosphere by some 10-20% compared with UARS data, which may be due to a difficulty in representing accurately over a long period of time the vertical advection of humidity by the very slow mean tropical ascent (Simmons et al., 1999).

A realistic simulation of stratospheric water vapour is dependent on the model not producing excessively cold temperatures, and thus excessive condensation, in the polar winter stratosphere. This has been a problem in earlier simulations produced by ECMWF models and by other models. Fig. 8 compares temperatures at 10hPa from the final years of the 60-level AMIP simulation with analysed temperatures from the corresponding years of the UKMO stratospheric analyses. The former are available only as monthly means; the latter are plotted as daily values. The simulation is evidently quite successful in representing both annual cycles and interhemispheric differences, and in particular does not produce too cold a winter pole. Winter polar temperatures are if anything slightly warmer than analysed (or observed, based on a limited comparison with radiosonde data at 50hPa from the South Pole). Underestimation of the relative warmth of the summer pole is rather more marked.

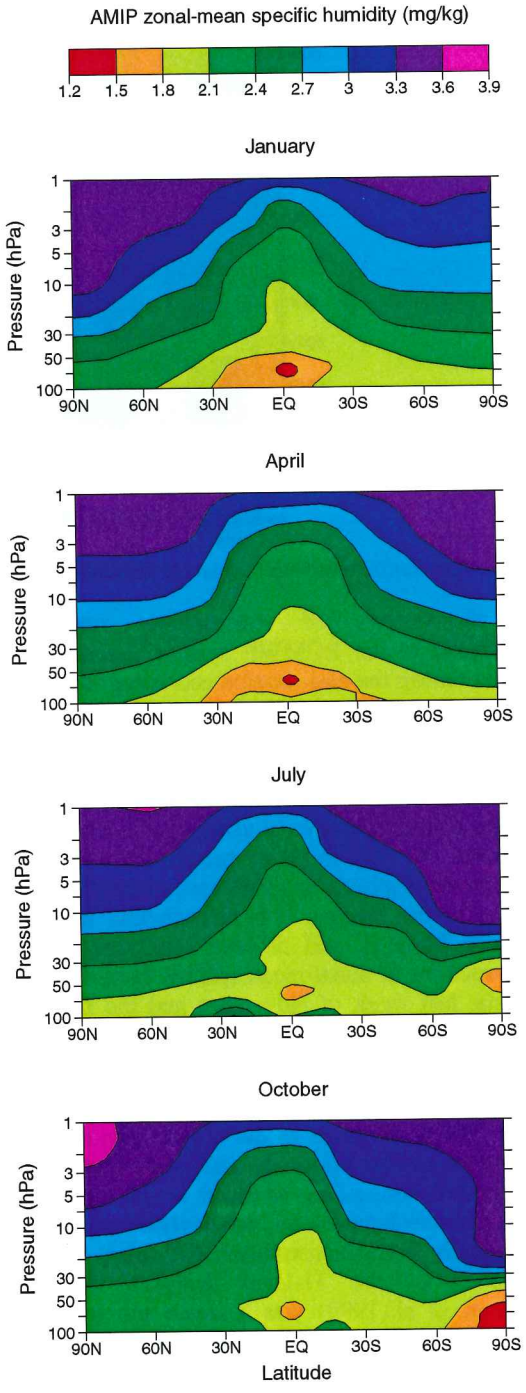


Fig. 7. Meridional cross-sections of zonal-mean specific humidity for January. April, July and October from a 60-level AMIP simulation. Each section is a 15-year average for 1983-1997.

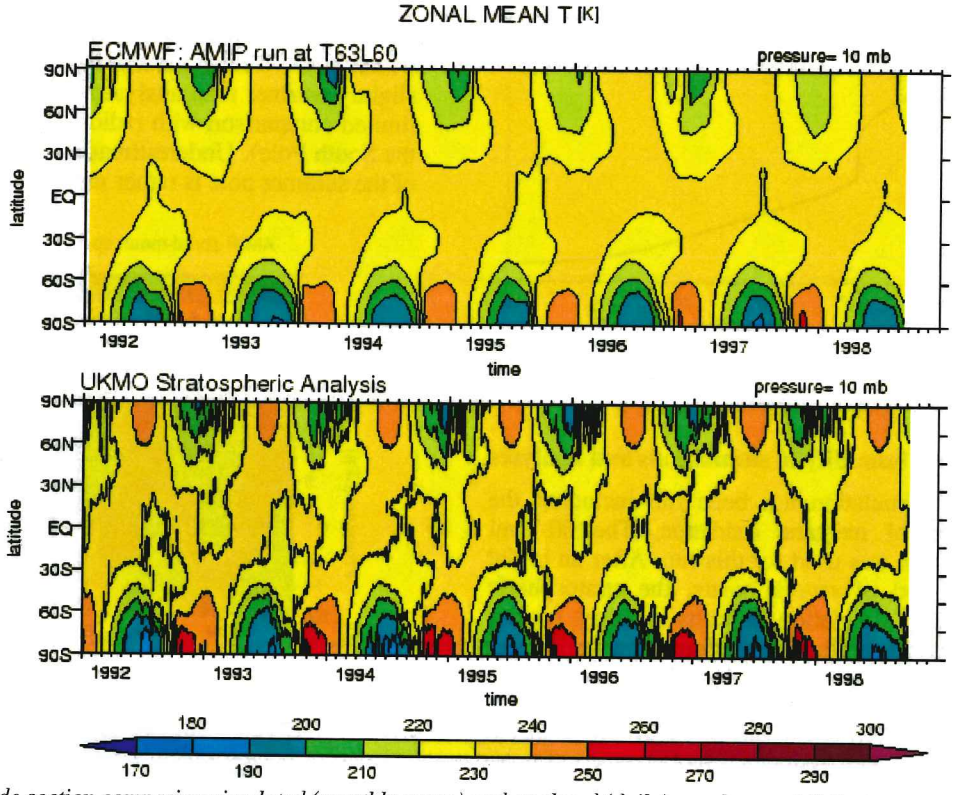


Fig. 8. Time/latitude section comparing simulated (monthly-mean) and analysed (daily) zonal-mean 10hPa temperatures.

No observations of stratospheric humidity are assimilated in the ECMWF system. Humidity simply evolves during the assimilation according to the model's dynamical and parametrized physical processes, with winds and temperatures (and tropospheric humidity) constrained by the analysis process. The parametrization of methane oxidation was activated on 17 June 1998 in the data assimilation cycles using the 50-level version of the model. As only about six months have elapsed since then, the stratospheric analyses have not yet become fully adapted to the change. It is nevertheless instructive to compare 31-level and 50-level analyses. Meridional cross-sections of zonal-mean specific humidity averaged over the last week of November and the first week of December are shown in Fig. 9.

The influence of the parametrization of methane oxidation on the 50-level system is seen predominantly above 10hPa, although as the northern winter proceeds we expect to see relatively moist air from the upper stratosphere descending to lower levels. Below 10hPa, the differences between the two analyses shown in Fig. 9 arise mostly from the resolution difference. A particular deficiency of the 31-level analyses identified by Simmons et al.(1999) was a much too rapid upward transfer of relatively moist or dry air introduced at the tropical tropopause. A slower, more realistic transfer was seen in a simulation with the 50-level version of the model, although upward transfer and attenuation in the deep tropics was still stronger than observed. Fig. 9 indicates that the improvement seen in the 50-level simulation carries over into the 50-level analyses. Relatively moist air was introduced into the stratospheric analyses in the boreal subtropics over the summer, and by December the maximum in humidity has reached 10hPa in the analyses from the 31-level system, and is still below 30hPa in those from the 50-level system. The maximum is located further north in the 31-level system, and much more moistening of the northern extratropics has taken place in this system. In addition, the 31-level system is drier than the 50-level system at 10hPa at high southern latitudes. This appears to be due partly to colder 10hPa temperatures (and more condensation) in winter and early spring in the 31-level system and partly to descent of moister air from the upper stratosphere in the 50-level system. By December, however, there is mean ascent at high latitudes (as occurs also in the AMIP simulation shown earlier), this acting to dry the middle stratosphere in this region and season.

A set of maps showing daily analyses of humidity over the southern hemisphere is presented in Fig. 10 for the period 30 November to 5 December. Fields are shown at the model level close to 51hPa for the 50-level system. The driest air is located in the decaying, perturbed westerly vortex, where dehydration by the parametrized condensation process occurred earlier when temperatures were sufficiently cold. The maps show evolving filamentary structure in the humidity field, and indications of mixing. Observations of fields which behave largely as passive tracers can be used effectively to extract information on the wind field in a 4D-Var assimilation system. This is a prime motivation for the work to include stratospheric ozone as a variable of the ECMWF system.

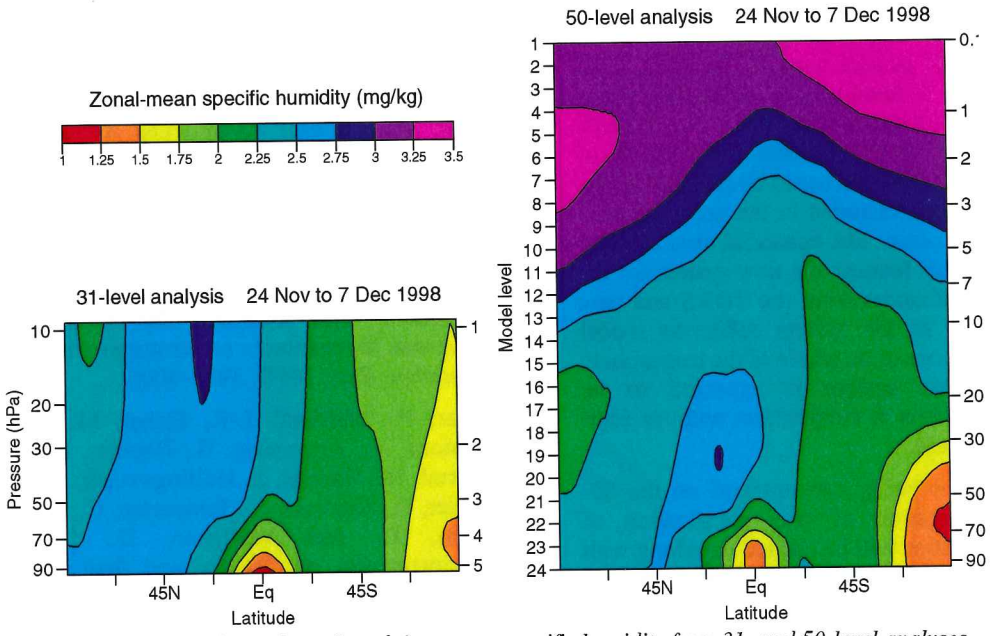


Fig. 9. Meridional cross-sections of zonal- and time-mean specific humidity from 31- and 50-level analyses

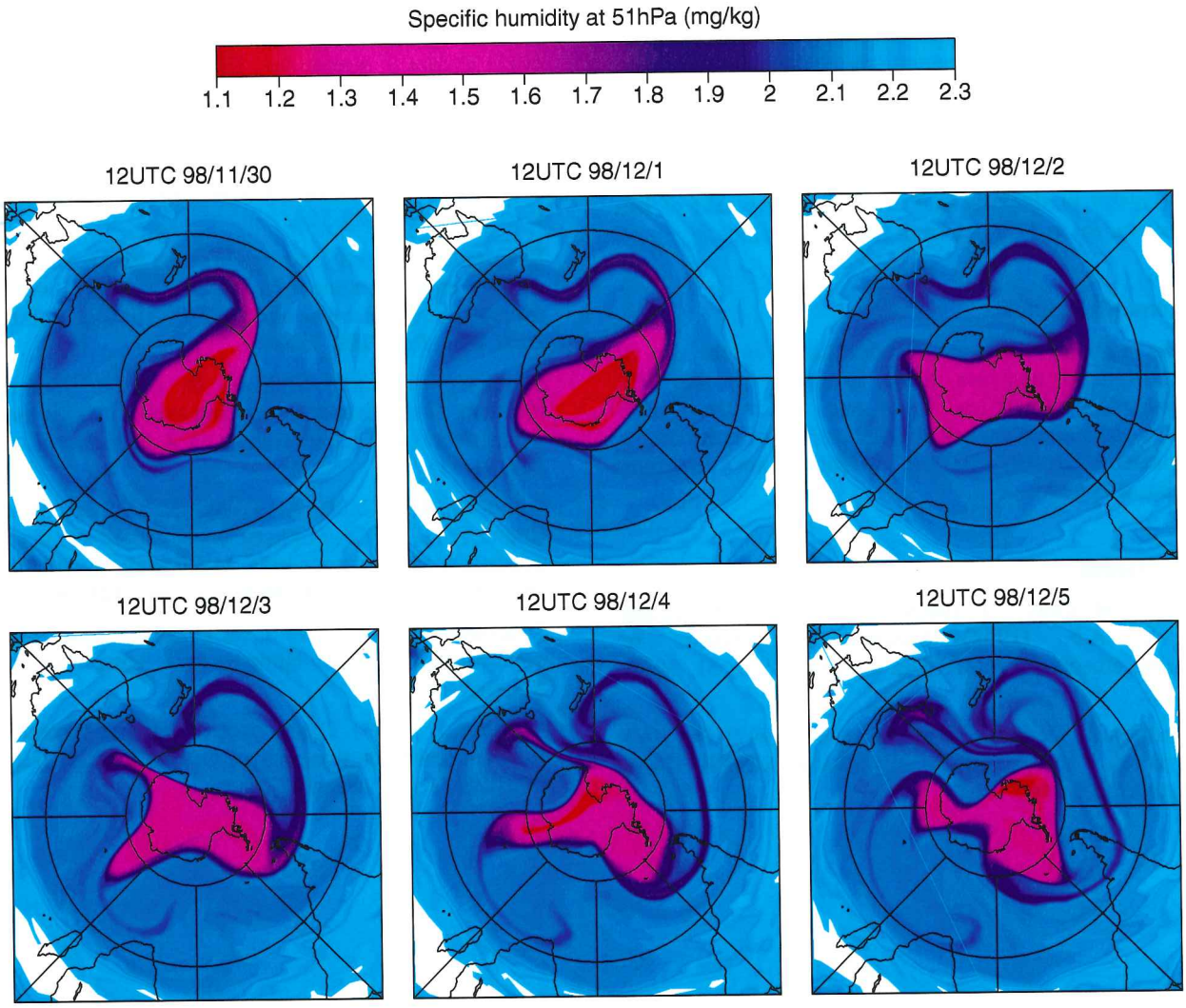


Fig. 10. Specific humidity at close to 51hPa from 50-level analyses for consecutive days from 30 November to 5 December 1998

6. Plans

The immediate goal of present work is the operational implementation of the 50-level version of the forecasting system early in 1999. Included with this will be the new fast radiative transfer model for the assimilation of radiance data from satellites discussed by Karcher at this workshop. This has been included in the near-real-time trials currently taking place. The enhanced stratospheric resolution is an essential feature of a new system for the assimilation of raw radiances from the TOVS and new ATOVS instruments, as the system relies on model background fields that cover the whole of the troposphere and stratosphere. This system is expected to be introduced into operations in conjunction with, or soon after, the 50-level system.

Further development is being concentrated on the 60-level system. Investigation of the specification of background error statistics will be carried out along with a revision of the specification of observation errors, which is known to have deficiencies (Järvinen, personal communication). The assimilation of ozone data will be adapted to use the 60-level resolution. The aim is for operational implementation later in 1999. A further aim is to use the 60-level system including ozone assimilation in a reanalysis of the observations available for the period from 1957 to the present day. Amongst other applications, the reanalyses and new operational analyses should provide the basic meteorological fields of good quality needed for further studies of stratospheric chemistry, particularly studies utilizing the wealth of data from recent and forthcoming satellite missions.

Acknowledgements

This work has been greatly facilitated by the use of stratospheric analyses made available by the UK Meteorological Office. Ideas and advice from Bob Harwood and colleagues at Edinburgh University on the representation of methane oxidation are gratefully acknowledged.

References

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

Bouttier, F., Derber, J., and Fisher, M. (1997) The 1997 revision of the J_b term in 3D/4D-Var. *ECMWF Tech. Memo.*, **238**, 54pp.

Brasseur, G., and Solomon, S. (1984) Aeronomy of the Middle Atmosphere. *D. Reidel Publishing Co., Dordrecht*,452pp.

Gates, W.L. (1992) AMIP: The Atmospheric Model Intercomparison Project. *Bull. Amer. Meteorol. Soc.*, **73**, 1962-1970.

Mote, P.W., Rosenlof, K.H., McIntyre, M.E., Carr, E.S., Gille, J.C., Holton, J.R., Kinnersley, J.S., Pumphrey, H.C., Russell, J.M. III, and Waters, J.W. (1996) An atmospheric tape recorder: The imprint of tropical tropopause temperatures on stratospheric water vapour. *J. Geophys. Res.*, **101D**, 3989-4006.

Rabier, F., Mahfouf, J.-F., Fisher, M., Järvinen, H., Simmons, A., Andersson, E., Bouttier, F., Courtier, P., Hamrud, M., Haseler, J., Hollingsworth, A., Isaksen, L., Klinker, E., Saarinen, S., Temperton, C., Thépaut, J.-N., Undén, P., and Vasiljevic, D. (1997) Recent experimentation on 4D-Var and first results from a Simplified Kalman Filter. *ECMWF Tech. Memo.*, **240**, 42pp.

Randel, W.J., Wu, F., Russell, J.M. III, Roche, A., and Waters, J.W. (1998) Seasonal cycles and QBO variations in stratospheric CH_4 and H_2O observed in UARS HALOE data. *J. Atmos. Sci.*, **55**, 163-185.

Ritchie, H., Temperton, C., Simmons, A., Hortal, M., Davies, T., Dent, D., and Hamrud, M. (1995) Implementation of the semi-Lagrangian method in a high resolution version of the ECMWF forecast model. *Mon. Wea. Rev.*, **123**, 489-514.

Simmons, A.J., and Burridge, D.M. (1981) An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates. *Mon. Wea. Rev.*, **109**, 758-766.

Simmons, A.J., and Strüfing, R. (1983) Numerical forecasts of stratospheric warming events using a model with a hybrid vertical coordinate. *Q. J. R. Meteorol. Soc.*, **109**, 81-111.

Simmons, A.J., Untch, A., Jakob, C., Kållberg, P., and Undén, P. (1999) Stratospheric water vapour and tropical tropopause temperatures in ECMWF analyses and multi-year simulations. *Q. J. R. Meteorol. Soc.*, **125**, *in press*.

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

Teixeira, J. (1999) The impact of increased boundary layer vertical resolution on the ECMWF forecast system. *ECMWF Tech. Memo.* 267.

THE IMPACT OF VERTICAL TRANSPORT ON OZONE ANALYSES BASED ON TOTAL OZONE MEASUREMENTS

Pierre Gauthier, Gilbert Brunet, Simon Pellerin and Paul-Antoine Michelangeli
Atmospheric Environment Service
2121 Trans-Canada Highway
Dorval, Québec, CANADA H9P 1J3

ABSTRACT

The Canadian 3D-var assimilation system has been used to produce univariate ozone analyses based on total ozone measurements retrieved from TOVS data. The assimilation experiments were driven by a modified version of the operational global spectral model used at the Canadian Meteorological Centre (CMC). Ozone is treated at first as a passive tracer and the background-error statistics were determined from a time-series of lagged forecasts (the NMC method) to provide variances of ozone defined as a function of latitude and pressure while the correlations are taken to be horizontally homogeneous and isotropic. The resulting total analyses compare well with those from NCEP but the vertical distribution shows some deficiencies after a week into the assimilation cycle. When only total ozone measurements are used, the vertical distribution in the analysis relies entirely on the structure functions introduced in the 3D-var. The background-error variances were changed to correspond to the climatological distribution of ozone so that the analysis was acting as a source of ozone which was at least climatologically correct. Ozone being transported incorrectly upwards, the analysis cannot correct the ozone field once it is displaced at levels where the variances become small. In this context, a good analysis could only be obtained with a good coverage of data providing some information about the vertical profiles. Recently, a complex chemistry was introduced in the model (Kaminski *et al.*, 1998). A preliminary analysis of the results indicates that the vertical transport is still displacing too high the ozone maximum. A global grid-point model has been used instead of the spectral model and the characteristics of the vertical distribution has been found to be quite different from that of the spectral model.

1. INTRODUCTION

Measurements of ozone profiles are very few and come from a sparse network of ozonesondes, Lidar measurements or other ground measurements. On the other hand, satellite measurements provide a large number of measurements, mostly of total ozone only. The production of an analysis of total ozone only requires that the transport of ozone be done correctly. In Levelt *et al.* (1996), a 2D advection of total ozone is done by prescribing the 2D wind field to correspond to those at the 250 hPa levels that were found to be optimal. It is better to advect the complete ozone field and this is what will be done here.

To what extent is the data assimilation process capable to reconstruct the full 3D distribution of ozone from total measurements only is one question to which some answers will be sought. This redistribution is highly dependent on the background-error covariances used in the analysis and also on the dynamics resolved by the model. The paper describes first in section 2 the assimilation system. Section 3 presents the total ozone analyses obtained while section 4 focusses on the vertical distribution of ozone in those analyses. Finally we conclude in section 5 by presenting the work being currently done at AES on the production and use of ozone analyses.

2. DESCRIPTION OF THE ASSIMILATION SYSTEM

The assimilation has been done using a modified version of the 3D-var assimilation system used operationally at CMC since June 1997 to produce its operational forecasts (Gauthier *et al.*, 1998b; Laroche *et al.* 1998). The model used in the assimilation cycle is a T199 spectral model comprising 29 levels in a "hybrid" vertical coordinate. This is a modified version of the operational model used at CMC until October 1998 for global analyses and forecasts (Ritchie and Beaudoin, 1994).

a. Description of the model

The characteristics of the model are as described in Ritchie and Beaudoin (1994) except that the terrain following $\sigma = p/p_s$ coordinate has been replaced by a "hybrid" vertical coordinate that matches the σ coordinate at low levels while becoming closer to a pressure coordinate in the stratosphere where it is more appropriate. This coordinate η is defined as

$$\frac{\eta - \eta_T}{1 - \eta_T} = (p - p_T)/(p_s - p_T)$$

with p_s and p_T being the pressure at the bottom and the top of the atmosphere respectively. The levels were taken to be spaced vertically as shown on Fig.1: this was a configuration that had been experimented with previously. Finally, the top of the model has been set such that $p_T = 1$ hPa with $\eta_T = 0.001$. Due to the increased wind velocities at these levels, the time steps of the semi-Lagrangian

algorithm had to be reduced to 900s (instead of the original 1800 s). When a longer timestep is used, noise was observed over mountainous areas like the Andes.

It was beyond the scope of this project to develop an assimilation cycle for all the dynamical variables of the model. Instead, for winds, temperature and specific humidity, the model was reinitialized every 24-hr using the 18 levels of the UKMO analyses located above 700 hPa (Swinbank and O'Neill, 1994). The CMC analyses were used at 1000, 925, 850 and 700 hPa for which no analysis was available from UKMO. The vertical levels are equally spaced in Δln p in the stratosphere thereby implying a uniform vertical resolution of approximately 2 km in this region. Fig.1 represents the vertical levels of both the analysis and the model. The model interpolates the analysis to its own levels before initiating its integration. It is important to stress that the UKMO analyses are only defined at low resolution. However, it is well known (Orsolini *et al.*, 1997) that advection with low resolution winds can nevertheless lead to ozone fields with filamentary structure that require a high resolution to be correctly represented. A horizontal resolution of T199 was used to resolve the fine scale structure of the ozone field.

b. Description of the variational assimilation system

A univariate analysis of ozone is performed using a 3D-var formulation of statistical interpolation. The analysis increments are produced every 6-h on the pressure levels shown in Fig.1 and interpolated to the model's grid and added to the high resolution background field. The 3D-var requires the forward model of the observation operator and its adjoint, the latter being necessary to compute the gradient of the cost function (see Gauthier *et al.*, 1998b, for more details). For the assimilation of total ozone measurements, the forecast of ozone is first interpolated vertically from the η levels to the analysis pressure levels. The forward interpolation then continues by interpolating horizontally the ozone field to the observation location and then integrating vertically this profile to obtain a model's integrated ozone at this location.

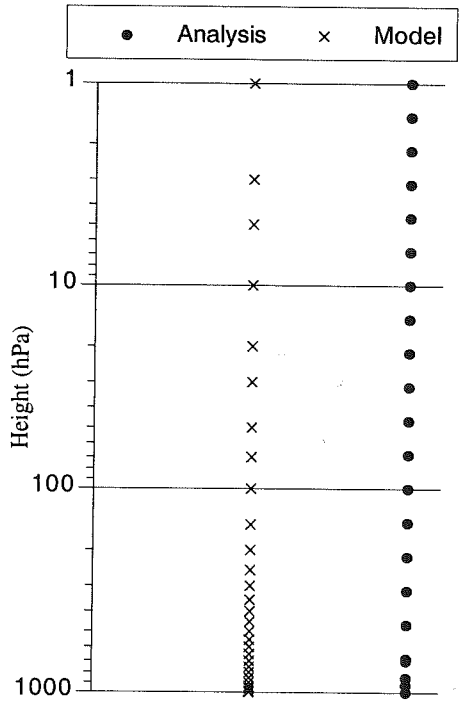


Fig.1 Vertical distribution of the analysis (•) and model (×) vertical levels.

With the top of the model being set at 1 hPa, there is a difference between this integrated ozone and the measured total ozone that includes a contribution from the ozone above 1hPa. The WOUDC data base comprises over 19,000 ozone profiles from ozonesondes for which there is also a total ozone measurement obtained from a Brewer instrument. Fig.2 shows the ratio

$$\alpha = \frac{\text{Measured total Ozone}}{\text{Integrated Ozone}}$$

as estimated from a fit to the data. The curve was extrapolated to 1 hPa and gave a value of $\alpha \cong 1.005$ which was used to convert total ozone to integrated ozone.

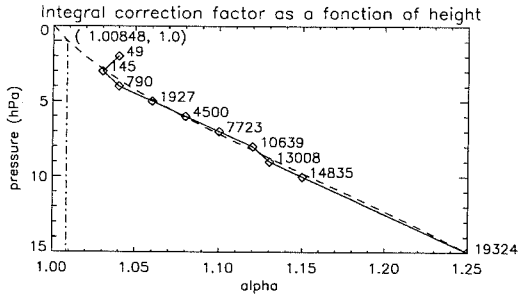


Fig. 2: Ratio between the total ozone measured by a Brewer instrument and the corresponding integrated ozone profile of ozone sondes reporting above a given level. This ratio has been computed for a total of 19,324 ozone sondes from the WOUDC. The numbers on the graph indicate how many sondes reported above this level (only 49 report above 2hPa).

c. Estimation of the background error statistics

The innovations $\boldsymbol{HX}_b - \boldsymbol{y}$ gives a measure of the misfit between the observed total ozone, $\boldsymbol{y}$, and its model equivalent, $\boldsymbol{HX}_b$. The assimilation produces a correction to the vertical profile to reduce this misfit between the forecast and the observation. The shape of the increments being determined by the

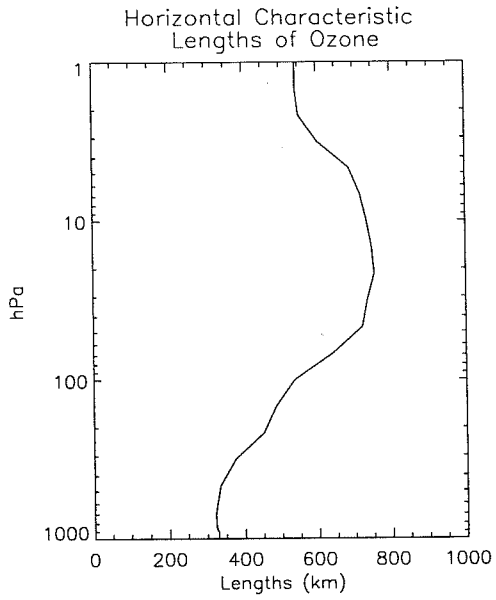


Fig.3: Estimate of the characteristic lengths of the horizontal correlations as a function of height.

background error covariances, those will control the vertical distribution of ozone and the ability of the assimilation at producing an analysis of the 3D distribution of ozone is directly related to our ability at correctly estimating the background error covariances **B**. In Riishojgaard *et al.* (1992), an ozone analysis was obtained by correcting the ozone profile according to a prescribed profile that depended on the level of the tropopause.

In a preliminary study done at AES by Génin and Mereyde (1996), the same setting as described here was used but with a lower resolution (T42) version of the model. Forecasts were made using the method of Riishojgaard to produce analyses out of which 48-h forecasts were made. Using a lagged forecast method (or NMC method) first proposed by Hoffman and Kalnay (1983) and used by Parrish and Derber (1992), Rabier *et al.* (1998), Bouttier *et al.* (1997) and Gauthier *et al.* (1998a), a set of vertical correlations were computed assuming them to be homogeneous over the whole globe. As for the horizontal characteristic length (Fig.3), they vary with height, reaching a maximum value of 755 km at the 20 hPa level. Details about this preliminary work can be found in Génin and Mereyde (1996). The resulting estimate of the standard deviation shown on Fig. 4 is confined above the tropopause, its maximum being around the level 10 hPa. The low values observed near the top level are attributed to the highest diffusion introduced in this layer.

The impact of the background-error covariances on an analysis increment is shown in Fig.5 in response to a single observation of total ozone.

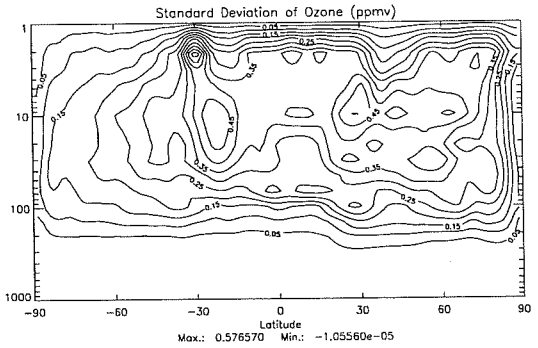


Fig.4: Estimate for the standard deviation of the zonal average of ozone forecast error (in ppmv) obtained with from a series of lagged forecasts (NMC method).

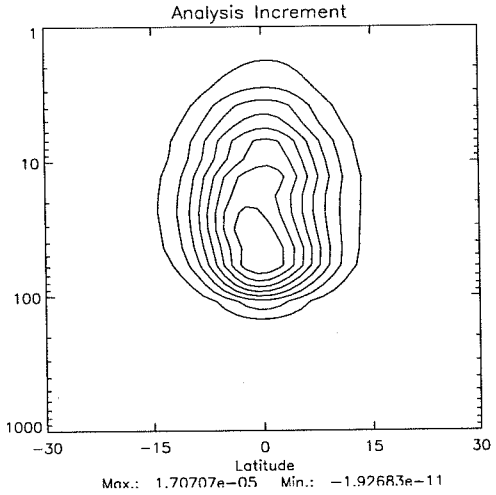


Fig.5: Analysis increment in response to a single observation of total ozone when the background-error standard deviation corresponds to that shown in Fig.4.

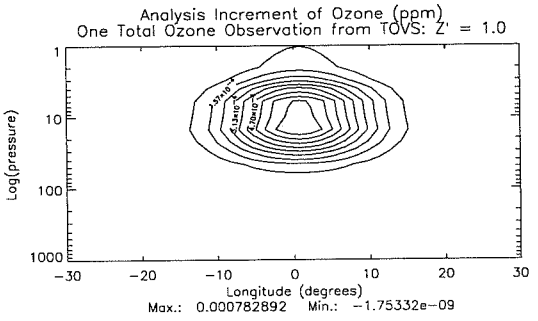


Fig.6: Same as Fig.5 but when the background-error standard deviation of Fig.corresponds is modulated according to the climatology.

There is a clear maximum located at 20 to 30 hPa. This means that any misfit between the forecast and the total ozone measurement will introduce ozone at levels that are lower than the expected 10 hPa of the climatology. To avoid this, the standard deviation were modulated according to the climatology. Fig.6

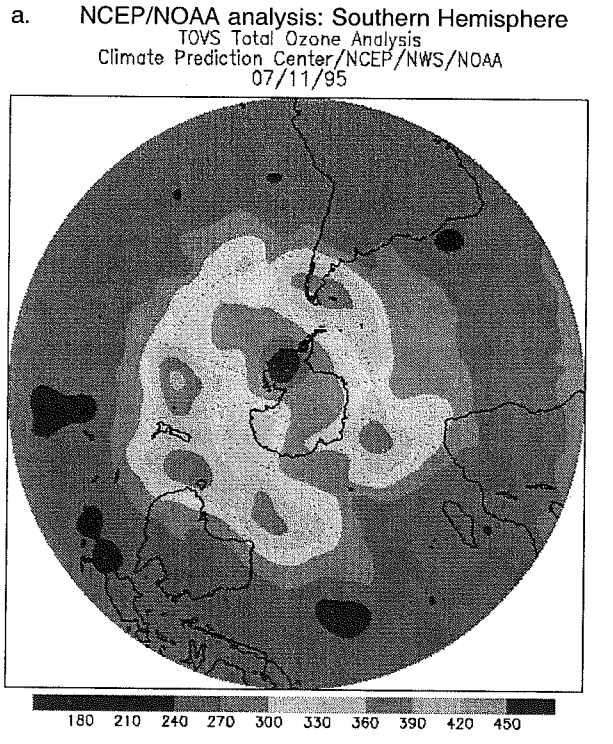


Fig.7: Comparison of total ozone analyses for July 11, 1995. Polar stereographic view of the Southern Hemisphere. Top panel is a NCEP analysis obtained from observations only, collected over a period of 24-h, bottom panel is the 3D-var analysis.

shows the response of the assimilation to a single observation of total ozone in that case. The maximum is now located around 10 hPa and the analysis would be climatologically more reasonable than the one obtained with the statistics of Fig.4.

3. TOTAL OZONE ANALYSES

Assimilation experiments using only TOVS total ozone data have been made. The results from a one-month assimilation cycle are presented for the period extending from June 21, 1995 to July 21, 1995. This period was chosen because of uninterrupted availability of TOVS data within the CMC archives. Fig. 7 compares the total ozone analyses on July 11, 1995 obtained with the 3D-var (bottom panel) with total ozone analyses from NCEP derived from data only collected over a period of 24-h (top panel). Polar stereographic views over the Antarctic are shown. The 3D-var analyses have local extrema that agree well with the data quantitatively for the amplitude and intensity. The 3D-var analyses also show a finer structure in the horizontal distribution that is considered to be more realistic of what one would expect of the distribution of total ozone. Since the data-only analysis comprises data over 24-h, this makes it difficult to resolve such fine structures.

The analysis shown in Fig.7 is representative of all total ozone analyses obtained with the 3D-var. Diagnostics based on the time evolution of these analyses indicate that they do not show any particular problem. Such is not the case with the 3D ozone distribution as will be seen in the next section.

4. VERTICAL DISTRIBUTION OF OZONE

Relying only on total ozone measurements, the vertical distribution of ozone is controlled by the background-error covariances and by the dynamics of the model. Using lagged forecasts generated by the model, the variance estimate obtained with this method does carry the signature of how the model behaves. For instance, the presence of a sponge layer near the top of the model is clearly seen in Fig.4 which shows a significant decrease in variance as one would expect in presence of dissipation. As explained earlier, this particular standard deviation field leads to an analysis increment acting as a source of ozone at a level that is too low. This is why the standard deviation was chosen to be proportional to the climatology of ozone. The proportionality constant was chosen so that the background and the observations had about an equal weight.

Fig.8 shows the zonal average of our ozone analyses averaged over the last two weeks of the assimilation to avoid the initial spin-up of the assimilation cycle. The model places the maximum of ozone around 3 hPa which is higher that that of the climatological 10 hPa observed in July. Fig.9 shows a similar zonal average obtained from a "nature" run in which the SFE model was coupled

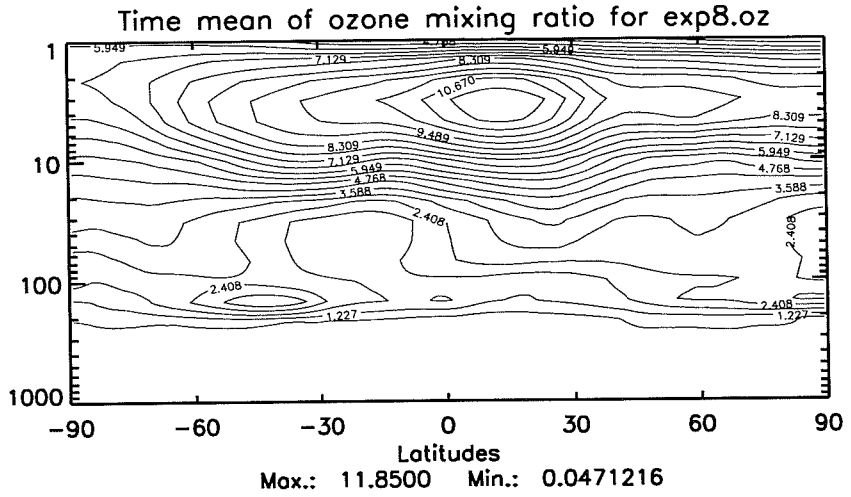


Fig.8: Zonal and time mean of the last two weeks of analyses based on the SFE model with ozone treated as a passive tracer. Only total ozone data were used.

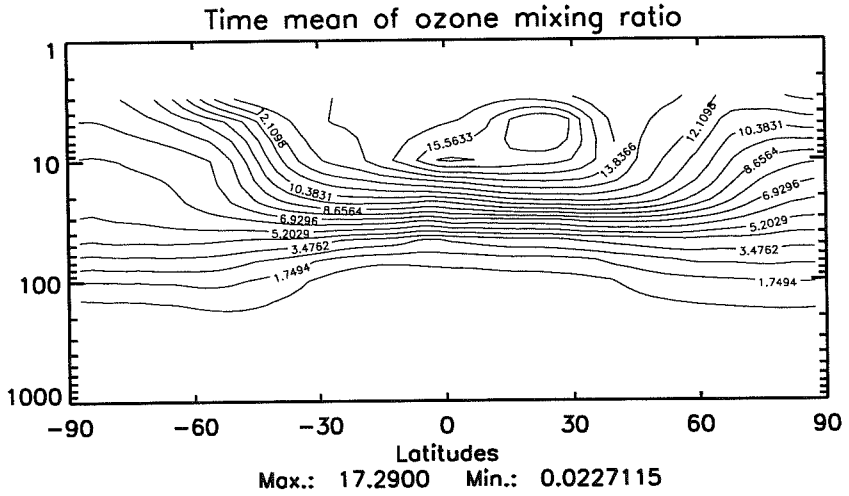


Fig.9: Zonal and time mean of the last two weeks of the nature run based on the SFE model coupled to a full chemistry.

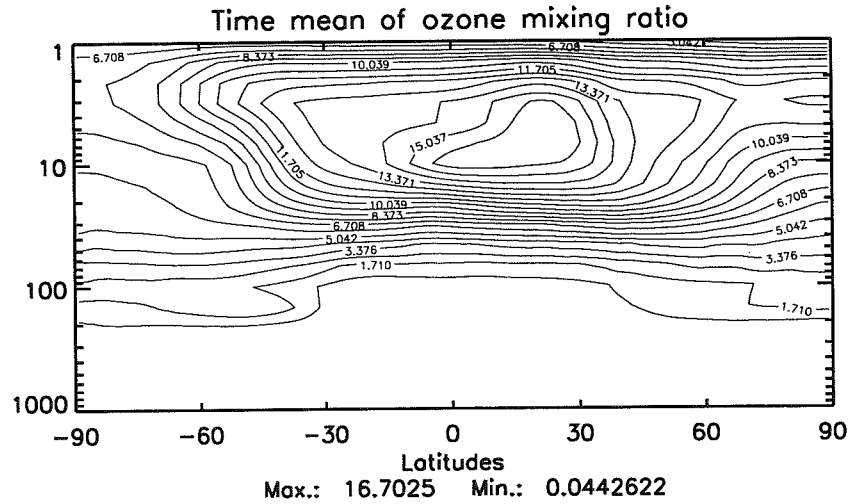


Fig.6: Zonal and time mean of the analyses of simulated ODIN data of ozone profiles. Ozone was treated as a passive tracer in the assimilating model.

to a complex chemistry (Kaminski *et al,* 1998). The presence of chemistry is beneficial to the ozone distribution, the maximum settling now around 5-6 hPa. More careful diagnostics of the results seem to indicate that there still is some problems with this integration.

As a final experiment, an OSSE was carried out and measurements of ozone profiles were generated from this nature run (Kaminski *et al.*, 1998). They were subsequently assimilated with the SFE/3D-var but without any chemistry. Fig.10 showns the zonal mean of the resulting ozone analyses and they compare quite well with that of the nature run (Fig.9). From this we can conclude that treating ozone as a passive tracer in data assimilation can lead to analyses that seem to have a good vertical distribution. It is of course important to remember that this statement depends on the nature and quality of the observed ozone profiles. Morevoer, these results fall within the "close cousins" category because the assimilating model has the same dynamics as that used to create the nature run.

5. CONCLUSION

There are many reasons why the Canadian Atmospheric Environment Service is interested in ozone. In general terms, the monitoring of ozone over the Arctic would be greatly improved if a reliable assimilation of the data could be performed. The assimilation system is also used in assisting the developement of new instruments such as ODIN and ORACLE, the latter being a laser instrument that would provide measurements of ozone profiles with a high vertical accuracy.

In this paper, a 3D-var assimilation system was used to assimilate total ozone data. The model used is a full model in which the dynamical variables are restored to archived analyses from the UKMO. The total ozone analyses agree well with those of NCEP qualitatively and quantitatively but finer scale structures can be observed in the 3D-var analyses which are more representative of the instantaneous total ozone field. The vertical distribution of ozone is however incorrect in placing the maximum too high. Introducing chemistry in the model has reduced the problem but at the present time, it is felt that the erroneous dynamics of the model is detrimental to the simulation with chemistry.

Preliminary experiments have been conducted with the global model that became operational at CMC in October 1998. This model has quite different characteristics than the spectral model used in the experiments presented here. It is in this context that conservative schemes for ozone are studied altogether with the introduction of a "true" hybrid vertical coordinate which reverts fully to pressure in the stratosphere.

6. REFERENCES

Bouttier, F., J. Derber and M. Fisher, 1997: The 1997 revision of the Jb term in 3D/4D-Var. *ECMWF Tech. Memo No. 238.*

Gauthier, P., M. Buehner and L. Fillion, 1998a: Background-error statistics modelling in a 3D variational data assimilation scheme: estimation and impact on the analyses. In the *Proceedings of the ECMWF Workshop on Diagnosis of Data Assimilation Systems,* 2-4 November 1998, Reading, U.K.

————, C. Charette, L. Fillion, P. Koclas and S. Laroche, 1998b: Implementation of a 3D variational data assimilation system at the Canadian Meteorological Centre. Part I: The global analysis. (to appear in *Atmosphere-Ocean*).

Génin, Y. and F. Mereyde, 1996: Assimilation variationnelle 3D du contenu total d'ozone. *Note de travail de l'Ecole Nationale de la Météorologie,* Météo-France, Toulouse.

Hoffman, R.N. and E. Kalnay, 1983: Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus,* **35A**, 100-118.

Kaminski, J.W., J.C. McConnell, G. Brunet and P. Gauthier, 1998: Assimilation of ODIN ozone data in the CMC weather forecast model with chemistry: a feasibility study. Proceedings of the *Workshop on Chemical Data Assimilation,* 9-10 December 1998, KNMI, de Bilt, The Netherlands.

Laroche, S., P. Gauthier, J. Saint-James and J. Morneau, 1998: Implementation of a 3D variational data assimilation system at the Canadian Meteorological Centre. Part II: The regional analysis. (submitted to *Atmosphere-Ocean*).

Levelt, P.F., M.A.F. Allaart and H.M. Kelder, 1996: On the assimilation of total ozone satellite data. *Ann. Geophysicae,* **14**, 1111-1118.

Orsolini, Y.J., G. Hansen, U.-P. Hoppe, G.L. Manney and K.H. Fricke, 1997: Dynamical modelling of wintertime lidar observations in the Arctic: ozone laminae and ozone depletion. *Quart. J.R. Meteor. Soc.,* **123**, 785-800.

Parrish, D.F. and J.C. Derber, 1992: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747-1763.

Rabier, F., A. McNally, E. Andersson, P. Courtier, P. Undén, J. Eyre, A. Hollingsworth and F. Bouttier, 1998: The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part II: structure functions. (to appear in *Quart. J.R. Met. Soc.*)

Riishojgaard, L.P., F. Lefèvre, D. Cariolle and P. Simon, 1992: A GCM simualtion of the Norhtern Hemisphere ozone field in early February 1990, using satellite total ozone for model initialisation. *Annales Geophysicae,* **10**, 54-74.

Ritchie, H.R. and C. Beaudoin, 1994: Approximations and sensitivity experiments with a baroclinic semi-Lagrangian spectral model. *Mon. Wea. Rev.*, **122**, 2391-2399.

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

Data assimilation - ozone

THE GEOS DATA ASSIMILATION SYSTEM FOR OZONE OBSERVATIONS

L. P. Riishøjgaard[1], I. Štajner[2] and G.-P. Lou[2]
Data Assimilation Office, NASA Goddard Space Flight Center
and
[1] Joint Center for Earth Systems Technology, Baltimore, MD
[2] General Sciences Corporation, Laurel, MD

ABSTRACT

A system for assimilation of observations of ozone and other minor constituents with a three-dimensional chemistry and transport model has been developed in the Data Assimilation Office at the NASA Goddard Space Flight Center. The primary scope of the system is to provide reliable background estimates of ozone profiles to users in the satellite instrument community, but the datasets generated by the system are expected to have applications in a wide range of research areas. This article provides an overview of the system design, along with sample results from the validation.

1. INTRODUCTION

In recent years the application of advanced data assimilation techniques typically developed for numerical weather prediction (NWP) to observations of ozone and other minor constituents has become a field of intense activity. These techniques are useful for generating regular gridded data from sparse, irregularly, and heterogeneous observational datasets (e.g. Levelt et al. 1996), and they may ultimately lead to a more widespread use of constituent observations directly for NWP applications (e.g. Riishøjgaard 1996). A three-dimensional data assimilation system for ozone observations oriented towards the former of these applications has been developed in the Data Assimilation Office (DAO) of NASA Goddard Space Flight Center as part of the Goddard Earth Observing System Data Assimilation System (GEOS-DAS). The gridded ozone fields produced by this system will be used as input to the retrieval algorithms of a number of satellite-borne sensors. A secondary objective is to use the system to generate long-term ozone datasets for global change studies. It is planned to modify the system for analyzing measurements of other chemical species as well.

In this article an overview of the system will be given, and its performance will be illustrated by sample results from the system validation. A more extensive description of the prototype version is provided by Riishøjgaard and Štajner (1998), while Štajner and Riishøjgaard, (1998) contains a description of further developed version, with particular emphasis on the forecast and observation error covariance models. A brief outline of future work related to the system concludes the article.

2. THE DATA ASSIMILATION SYSTEM

The ozone data assimilation system operates in a way that is similar to most meteorological systems: At the analysis time, an equation is solved for the analysis, based on a background state – generally a forecast issued from the previous analysis - and the observations available at that time, under given assumptions about the forecast and observations error characteristics. The analyzed state is propagated forward in time by the forecast model until the next analysis time, and the cycle is repeated.

In the ozone assimilation system, the analyzed as well as the prognostic variable is the ozone mixing ratio. Observations of related quantities can be assimilated through the use of observation operators. The total ozone column – which is readily observable both from space and from the ground – is an example of such a quantity.

2.1. The forecast model

The ozone mixing ratio μ is forecast using the advection equation Rood et al. (1991):

$$\frac{\partial \mu}{\partial t} + \vec{u} \cdot \nabla \mu = P - \mu L \,, \qquad (1)$$

where $\vec{u}$ is the velocity field, and the P and L terms account for production and loss of ozone by photochemical processes. The global model grid has a horizontal resolution of 2° latitude by 2.5° longitude. In the vertical there are 29 levels, the lowest 9 of which are hybrid $\sigma - p$ levels, while the remaining 20 (extending from 150 hPa to the top of the atmosphere) are pressure levels.

The transport component of eq. (1) is solved using the flux-form semi-Lagrangian scheme of Lin and

Rood (1996). The input to the transport model consists of analyzed wind and surface pressure fields from the GEOS data assimilation system (DAO 1996). These fields are available every 6 hours and are interpolated linearly in time to the model time. In the first version of the system (Riishøjgaard and Štajner 1998), the values of P and L were tabulated according to the net results from a two-dimensional photochemical model as described by Douglass et al (1996). As shown in the subsequent detailed validation (Štajner and Riishøjgaard, 1998), this lead to a substantial bias in the upper stratosphere. Pending a solution to this problem, it was decided to run the initial operational version of the system with no active chemistry, i.e with P and L set to zero.

2.2. The ozone observations

Currently two types of ozone observations are used, total ozone from the Total Ozone Mapping Spectrometer (TOMS) and vertical ozone profiles from the Solar Backscatter Ultra Violet instrument (SBUV). Each of these instruments provides regularly distributed coverage of the globe within 24 hours, both have been providing measurements since 1978, and both have been validated extensively against other ozone measurements (e.g. Fleig et al. 1988 for TOMS; McPeters et al. 1994 for SBUV).

The TOMS instrument has flown on a number of platforms since the launch of Nimbus-7 in 1978, and at present (December 1998) Earth Probe TOMS is operative. There are plans for a future instrument on the Russian Meteor-3M satellite scheduled for launch in the year 2000. The instrument operates in a cross-track nadir scanning mode, and on most platforms it provides full coverage of the sunlit portion of the globe once every 24 hours. For the Nimbus-7 orbit, the nadir pixel size is on the order of 50 by 50 km, increasing to about 50 by 190 km at either extremity of the scan. For details on the instrument and the total ozone retrieval algorithm, see McPeters et al. (1996).

According to McPeters et al. (1996), the random part of the error of the retrieved product is multiplicative and on the order of 2% of the reported value. In addition there may be an unknown time invariant bias estimated to be less than 3%, and a possible residual drift in the calibration of less than 1% per decade. Errors in the correction applied to cloudy scenes may cause observation errors to be spatially correlated. A scan angle dependent error known to exist when the aerosol loading of the stratosphere is high introduces horizontal error correlations as well. However, no quantitative results have been reported about any of these correlations.

Since the spatial resolution of the TOMS data is higher than that of the model, all observations in the level 2 data files obtained within a given model grid cell valid for a given analysis time are averaged as part of the observation preprocessing. The minimum number of measurements required within a grid cell for the observations to be used varies with latitude from 3 near the poles to 11 near the equator. This ensures that the grid cell is reasonably well observed, and it allows us to calculate the sample standard deviation of observations within the cell, which is used in the analysis algorithm as a measure of the error of representativeness (Cohn 1997). In addition to this error, a fixed relative error of 1.5% and mean zero is assumed. These two errors are assumed to be mutually as well as spatially uncorrelated.

The SBUV instrument was operating on the Nimbus-7 satellite from December 1978 through September 1990. Its successor SBUV/2 has flown on NOAA-9, NOAA-11, and NOAA-14 (currently operative) and data from these instruments have been processed and disseminated by NESDIS in near-real time. The next SBUV/2 instrument will be on NOAA-L (to be renamed NOAA-16 once operational), which is scheduled for a year 2000 launch. The ozone profile retrieval algorithm for SBUV is described by Bhartia et al. (1996). The retrieved product consists of partial ozone columns of 12 individual layers roughly concurrent with the so-called Umkehr layers (Mateer and DeLuisi 1992). At present, only the 10 upper levels are used in the assimilation, providing information from the top of the atmosphere down to 127 hPa.

The observation error covariance for SBUV is determined from the difference between a sample covariance of observation-minus-forecast increments and the modeled forecast error covariance. The covariance model allows for altitude-dependent observation error variances and for vertical error correlations, while the observation error is assumed to be horizontally uncorrelated. The observation error model is described in detail by Štajner and Riishøjgaard, (1998). Since the instrument field of view – roughly 200 by 200 km – is close to the grid size of the assimilating model, and the vertical resolution is lower than that of the model, the error of representativeness is assumed to be zero for SBUV.

2.3. The analysis

At each analysis time, the analyzed state w^a, a vector of length n, is calculated based on a prior estimate of the state given by a forecast w^f valid at that time, and a p-vector of observations, w^o, according to the statistical analysis equation:

$$w^a = w^f + K(w^o - Hw^f)\,. \qquad (2)$$

Here H is the matrix of the observation operator that maps the forecast variables onto the observation variables. Under certain assumptions about the errors of the forecast and the observations (Cohn 1997), the minimum error variance analysis is obtained by using the following $n \times p$ gain matrix K

$$K = P^f H^T (H P^f H^T + R)^{-1}\,, \qquad (3)$$

where P^f and R are the forecast and observation error covariance matrices, respectively.

The ozone analysis system uses the PSAS equations (Physical-space Statistical Analysis System; Cohn et al. 1997) to obtain a solution of eq. (2) without

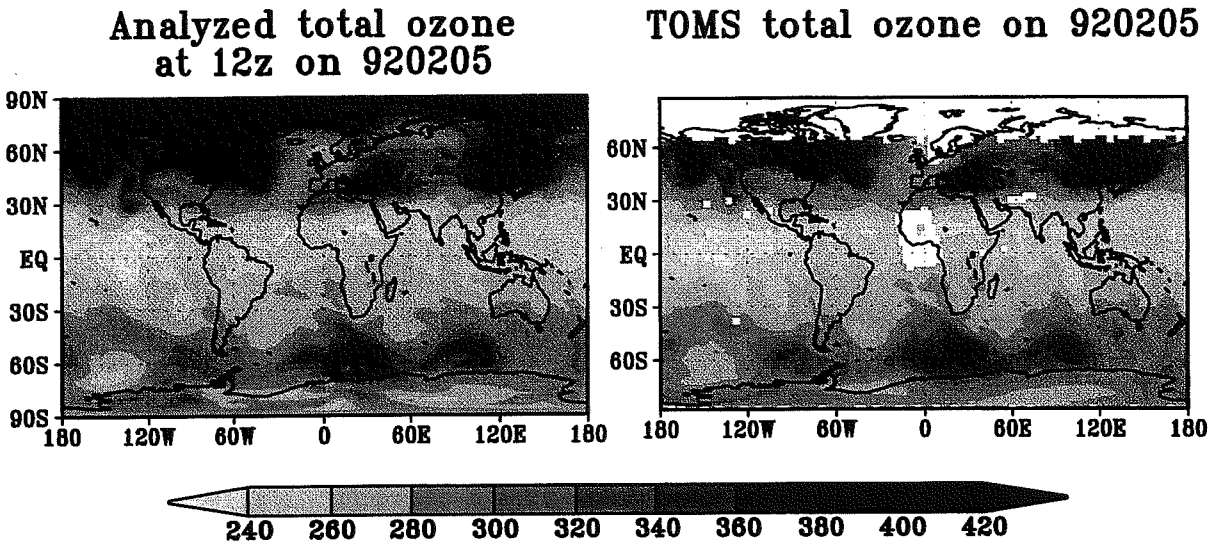


Figure 1. GEOS-DAS total ozone analysis and TOMS total ozone for Feb 5, 1992.

the costly construction of the matrix K. The PSAS algorithm consists of solving the innovation equation

$$(HP^fH^T + R)y = (w^o - Hw^f)\,, \qquad (4)$$

for the vector y, which is then mapped back onto the model state space by

$$w^a = w^f + P^fH^Ty\,. \qquad (5)$$

In contrast to the six-hourly analysis/forecast cycle typically used in operational meteorological data assimilation systems, the analysis equation of the ozone system is solved following each model time step (i.e. every 15 minutes), using observations obtained within 7.5 minutes from the analysis time. In this way, the satellite observations – that are generally obtained continuously in time – are ingested by the system closer to the actual observation time. An additional advantage is the decrease in the total computational burden achieved by analyzing a number of smaller batches of observations, rather than a single large batch all at once (Riishøjgaard and Štajner 1998).

The observation operator for TOMS is a discrete approximation to the pressure-weighted integral of the ozone mixing ratio over the depth of the model atmosphere, assuming this to be constant within a model layer. The l^{th} component, corresponding to a TOMS observation, of the p-vector Hw^f is

$$(Hw^f)_l = C\sum_{k=1}^{29} w^f_{i(l)\,j(l)\,k}\Delta p_k\,, \qquad (6)$$

where $i(l)$ and $j(l)$ are the latitude and longitude indices corresponding to the observation location, Δp_k is the thickness of the k^{th} model layer, and C is a factor for conversion from units of pressure to Dobson units.

The observation operator for SBUV consists of two parts. First a simulated forecast profile is constructed at the center of the instrument field of view by bilinear horizontal interpolation from the four surrounding model profiles. Next, this profile is integrated vertically over each of the 10 SBUV layers used in the assimilation. This part is described by eq. (6), with Δp_k now denoting the pressure increment of the intersection of the k^{th} model layer and the SBUV layer in consideration.

The forecast error variance is assumed to be proportional to the value of the mixing ratio itself, and the coefficient of proportionality has been found by experimentation. The horizontal forecast error correlation is elongated, with length scales at the equator of 770 km in the zonal direction and 385 km in the meridional direction (Štajner and Riishøjgaard, 1998). In the vertical, the forecast error correlations span at most 11 layers, and they are modeled using a compactly supported 5^{th} order piecewise rational function of the logarithm of the ambient pressure (Gaspari and Cohn 1998).

3. EXPERIMENTAL RESULTS

Results are presented for the period December 13, 1991 through February 28, 1992. During the course of the assimilation, the three-dimensional ozone fields are interpolated onto 29 standard pressure levels (850, 700, 600, 500, 400, 300, 250, 200, 170, 150, 130, 115, 100, 85, 70, 50, 40, 30, 25, 20, 15, 10, 7, 5, 3, 2, 1, 0.5, and 0.2 hPa) and written to file every 6 hours. Comparisons with independent observations are based on the nearest neighbor in time and space in this dataset.

An example of a spatial comparison with TOMS is presented in Fig. 1. The left panel shows the total ozone analysis for February 5, 1992, at 12 Z. The observed total ozone from Nimbus-7 TOMS for February 5, 1992 is shown to the right. The two fields are very similar, even though there is evidence of phase differences caused by the fact that the analysis is instantaneous, while the TOMS measurements are accumulated over 24 hours in order to get the near-global coverage shown. The assimilated total ozone fields represent a substantial improvement over both climatological ozone fields that tend to be nearly zonal, and the simulated fields of Douglass et al (1996), that have qualitatively correct non-zonal structures but strong biases. The RMS difference between the observed and assimilated total ozone fields is generally on the order of 10 DU. This value is close to 3 % of the global mean total ozone, and is largely consistent with the assumptions about the forecast and observation errors. For comparison, the RMS difference between the simulated ozone fields of Douglass et al (1996) and TOMS are on the order of 35 DU.

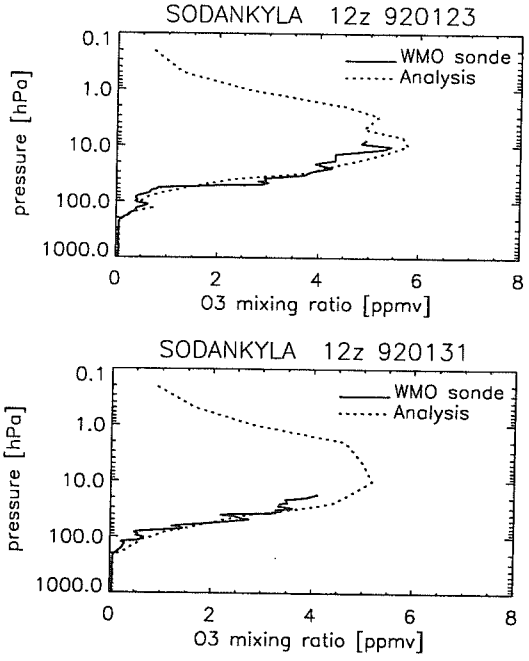


Figure 2. Sonde vs. analysis ozone mixing ratio profiles; Sodankyla (67.4° N 26.6° E)

The benefit of the data assimilation is particularly evident in the winter hemisphere, where the data void in the polar night region is filled in, and where the large excursions (up to ± 100 DU) from the zonal average seen in the TOMS total ozone are well captured.

As part of the external validation against independent data, profiles from the WMO balloon-borne ozone sondes have been compared to the corresponding profiles from the analysis and the simulation. During the main validation period, a total of about 400 sondes were launched, mostly at extratropical northern latitudes during the EASOE measurement campaign (European Arctic Stratospheric Ozone Experiment; see *Geophys. Res. Lett.*, special issue, vol. 21, no. 13, 1994). Even this unusually intense measuring activity resulted in an average of only five sondes per day. In spite of the sparse coverage provided, the sonde data are crucial for the validation of the assimilated profiles in the lower stratosphere, since they provide the sole source of high resolution information there.

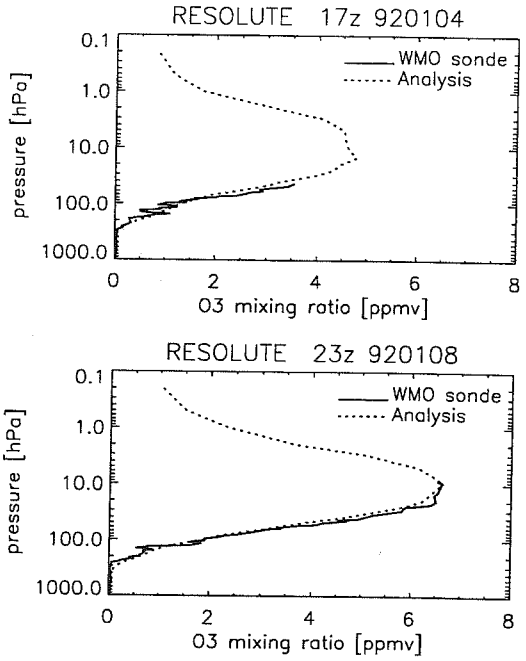


Figure 3. Sonde vs. analysis ozone mixing ratio profiles; Resolute (74.7° N 95.0° W)

The profiles shown in Figs 2 and 3 were all obtained in the polar night region (between 5 and 16 degrees north of the northernmost TOMS and SBUV/2 observations available to the system at their respective times), and they therefore also serve to illustrate the performance of the assimilation system in unobserved regions. Overall, the agreement between the sondes and the analysis profiles is excellent throughout the lower stratosphere. In all profiles, the tropopause height and the slope of the profile in the lower stratosphere is captured by the assimilation, and in one of Sodankyla profiles (Fig. 2, top panel), the analysis captures one of the major laminar features that are often seen in winter and spring extratropical profiles. Laminae of the vertical extent shown here are generally captured by the assimilation.

In Fig. 4 the mean of all profiles for Februay 15, 1992, obtained with the HALOE (Halogen Occultation Experiment; Russell *et al*, 1993) instrument on NASA's Upper Atmosphere Research Satellite (UARS) are shown together with the mean of the corresponding analysis profiles. The HALOE instrument has a vertical resolution of about 2.5 km in the stratosphere above 100 hPa, albeit with a relatively sparse horizontal coverage of 28-30 observations per day and a highly non-uniform sampling pattern, with the observations tending to concentrate in two narrow latitude bands most of the time. Nonetheless these profiles support the notion that the analyzed profiles

agree well with independent observations in the lower stratosphere.

In the upper stratosphere, the disagreement is somewhat larger. The analyzed profiles tend to peak at a higher altitude than the HALOE profiles, and the peak is more pronounced in the analysis. With the "no chemistry" forecast model, the shape of the ozone profile near the peak is mostly determined by the SBUV/2 observations, so the disagreement between the HALOE and the analysis profiles near 10 hPa essentially reflects a discrepancy between the two instruments. HALOE and SBUV are based on different measurement principles, have different viewing geometries, and their respective platforms are in different orbits. A direct comparison between the two instruments is often problematic due to the need for transforming one observation type into something that can be meaningfully compared with the other, and due to the lack of temporally and spatially collocated observations (McPeters et al. 1994). The potential for a data assimilation system defining the common ground in cross-validation of different measurements of the same geophysical variable is clearly illustrated in Fig. 4.

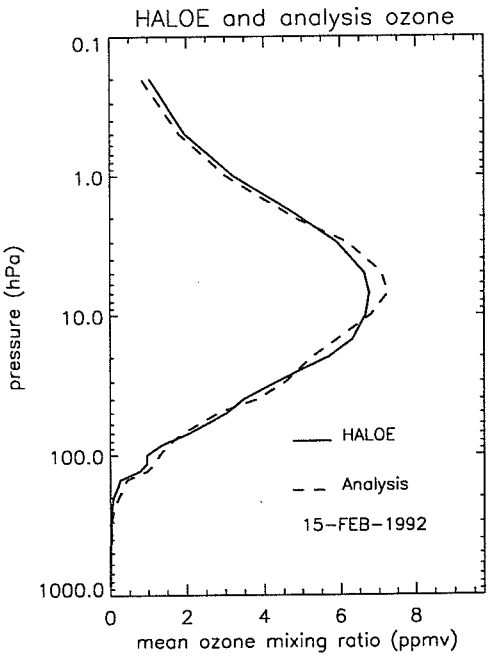


Figure 4. Mean HALOE vs. analysis profiles

4. SUMMARY AND OUTLOOK

The ozone data assimilation system developed at NASA Goddard Space Flight Center has been presented. The system generates total ozone and ozone profiles in near-real time for satellite remote sensing applications.

Examples of both internal and external validation have been shown. The data assimilation produces high quality total ozone fields, with the model forecasts generally close to the TOMS observations. The assimilated profiles compare well with balloon profiles and with HALOE profiles in the lower stratosphere. The peak values are higher than those observed by HALOE, and the peak is generally found at a higher altitude.

Future development will take place in a number of areas:

Based on requests from user groups in the satellite remote sensing community,the horizontal resolution of the system will be increased to 1° latitude by 1° longitude in the near future.

An important issue is the presence of biases in the forecast model and between different instruments. The present version of the system is based on the assumption of forecast and observation errors having zero mean. Biases – here taken as differences with non-zero mean – between comparable ozone observations of different types have been reported in numerous studies, and one therefore cannot *a priori* assume a given observation type to be unbiased with respect to the true atmospheric state. Errors with non-zero mean can be treated either by correcting the datasets for known biases prior to the analyses, or by redesigning the analysis to include non-zero mean errors (Dee and da Silva, 1998).

The forecast error covariance model in the present system is crude. The forecast error correlations are static, and although the forecast error variances are state-dependent, the variance model does not account for error growth during the range of the forecast, nor does it account for the reduction in error achieved through the analysis. Recently proposed strategies for addressing both these problems will be tested in the system (Riishøjgaard, 1998; Riishøjgaard, 1998b).

Current plans for analyzing other constituents include CH_4 and CO. Future work is likely also to include N_2O, SO_2, and aerosols.

ACKNOWLEDGMENTS

The authors wish to thank their colleagues at the Goddard Space Flight Center, in particular Ricky Rood and Steve Cohn, for their support and encouragement during the development of the new system. Anne Douglass and Eric Nielsen provided the code for the chemistry and transport model, Steve Cohn, Richard Ménard and Greg Gaspari offered useful suggestions on various aspects of the system, and Dick Dee helped with tuning the parameters of the forecast and observation error covariance models. Financial support for this work was obtained through the Cooperative Agreement (NCC 592) between University Maryland Baltimore County and NASA Goddard Space Flight Center, and NASA grants no. 621-82-31-20 and 622-59-07-21.

REFERENCES

Bhartia, P. K., McPeters, R. D., C. L. Mateer, L. E. Flynn, and C. Wellemeyer, Algorithm for

the estimation of vertical ozone profiles from the backscattered ultraviolet technique, *J. Geophys. Res.*, **101**, 18793 (1996).

Cohn, S. E., An introduction to estimation theory, *J. Meteorol. Soc. Jap.*, **1B**, 257 (1997).

Cohn, S. E., A. da Silva, J. Guo, M. Sienkiewicz, and D. Lamich, Assessing the Effects of Data Selection with the DAO Physical-space Statistical Analysis System, *DAO Office Note 97-08.* Data Assimilation Office, NASA Goddard Space Flight Center (1997).

DAO, *Algorithm Theoretical Baseline Document Version 1.01.* Data Assimilation Office, NASA Goddard Space Flight Center (1996).

Dee, D. P., and A. M. da Silva, Data assimilation in the presence of forecast bias, *Q. J. R. Meteorol. Soc.*, **124**, 269 (1998).

Daley, R., *Atmospheric data analysis*, Cambridge University Press, Cambridge, NY (1991).

Douglass, A. R., C. J. Weaver, R. B. Rood, and L. Coy, A three-dimensional simulation of the ozone annual cycle using winds from a data assimilation system, *J. Geoph. Res.*, **101**, 1463 (1996).

Fleig, A. J., D. S. Silberstein, C. G. Wellemeyer, R. P. Cebula, and P. K. Bhartia, An assessment of the long-term drift in TOMS total ozone data, based on comparison with the Dobson network, *Geophys. Res. Lett.*, **15**, 1133 (1988).

Gaspari, G. and S. E. Cohn, Construction of correlation models in two and three dimensions, *Q. J. Meteorol. Soc., in press.*

Levelt, P. F., M. A. F. Allaart, and H. M. Kelder, On the assimilation of total-ozone satellite data, *Ann. Geophysicae.*, **14**, 1111 (1996).

Lin, S. J., and R. B. Rood, Multidimensional flux-form semi-Lagrangian transport schemes, *Mon. Wea. Rev.*, **124**, 2046 (1996).

Mateer, C. L., and J. J. DeLuisi, A new Umkehr inversion algorithm, *J. Atmos. Terr. Phys.*, **54**, 537 (1992).

McPeters, R. D., T. Miles, L. E. Flynn, C. G. Wellemeyer, and J. M. Zawodny, Comparison of SBUV and SAGE II ozone profiles: Implications for ozone trends, *J. Geophys. Res.*, **99**, 20513 (1994).

McPeters, R. D, A. J. Krueger, P. K. Bhartia, J. R. Herman, B. M. Schlesinger, *et al.*, Nimbus-7 Total Ozone Mapping Spectrometer (TOMS) data products user's guide, *NASA Reference Publication 1384*, National Aeronautics and Space Administration, Washington DC (1996).

Riishøjgaard, L. P., On four-dimensional variational assimilation of ozone data in weather prediction models, *Q. J. R. Meteorol. Soc.*, **122**, 1545 (1996).

Riishøjgaard, L. P., A direct way of specifying flow-dependent background error correlations for meteorological analysis systems *Tellus*, **50A**, 42 (1998).

Riishøjgaard, L. P., A method for estimating the error variance of a physical space analysis system, *manuscript submitted to Q. J. Roy. Meteorol. Soc.* (1998b).

Riishøjgaard, L. P., and I. Štajner, The GEOS ozone data assimilation system, *manuscript submitted to Adv. Space Res.* (1998).

Rood, R. B., A. R. Douglass, J. A. Kaye, M. A. Geller, C. Yuechen, *et al.*, Three-dimensional simulations of wintertime ozone variability in the lower stratosphere, *J. Geophys. Res.*, **96**, 5055 (1991)

Russell, J. M. III, L. L. Gordley, J. H. Park, S. R. Drayson, D. H. Hesketh, *et al.*, The Halogen Occultation Experiment, *J. Geophys. Res.*, **98**, 10777, (1993).

Štajner, I., and L. P. Riishøjgaard, The GEOS ozone data assimilation system: Specification of error statistics, *manuscript in preparation* (1998).

SOME RESULTS FROM THE STATISTICAL INTERPOLATION OF SATELLITE OZONE DATA

S. Grainger

Meteorology CRC, Monash University, Wellington Road, Clayton, VIC 3168, Australia
Phone: +61-3-9905-9680, Fax: +61-3-9905-9689
Email: grainger@vortex.shm.monash.edu.au

ABSTRACT

Statistical Interpolation has been used to obtain global analyses of the ozone distribution. This has been done as a 2D analysis - using TOVS total ozone, and as a 3D analysis - using all available satellite ozone mixing ratio data. First-guess ozone fields are obtained by the implicit (2D) or explicit (3D) advection of the analysed ozone field.

Results are presented from both analyses for September-November 1994. They show good agreement between the analyses and correlative measurements. Issues that require further investigation are indicated at the end of the paper.

1. INTRODUCTION

Ozone is one of the most important atmospheric trace gases. It is a powerful absorber of UV radiation and an important greenhouse gas. Changes to the ozone distribution therefore have the potential to change global climate. However, the lack of a historical dataset inhibits an accurate estimate of ozone radiative forcing (IPCC, 1994), for example.

The purpose of this paper is to present results from global analyses of ozone produced from satellite data. Statistical Interpolation (SI) is used to combine total ozone and ozone mixing ratio data with a first-guess field to create a global ozone distribution every 6 hours. The SI algorithm and the methods used to obtain the first-guess fields are described in the next section. Section 3 presents results from the 2D analysis of TOVS total ozone. Results from the operational implementation of the total ozone forecast are described in section 4. Section 5 presents results from the 3D ozone mixing ratio analysis. Finally, section 6 examines some of the issues that need to be investigated in order to improve the ozone analysis.

2. METHODS

Statistical Interpolation (SI) attempts to fit a set of observations to a first-guess field such that the expected analysis error is minimised. Observations are converted to increments normalised by the first-guess error. The increment at each grid point is found by inverting a matrix of error correlations between the observations and the grid points. The algorithm used is described in Lorenc (1981).

The analysis is performed every 6 hours on a global 2.5 degree horizontal grid. The analysis can be run either in 2D mode to obtain the total ozone distribution, or in 3D mode to simultaneously analyse the ozone mixing ratio and total ozone distributions. In the 3D analysis there are 19 pressure levels between 1000 and 0.3hPa. This includes 8 levels in the lower stratosphere between 100 and 10hPa. Details of the implementation of the SI algorithm are given in Grainger (1998).

An important part of the ozone analysis scheme is the ability to generate good quality first-guess fields. This is required so that data sparse regions, with little or no satellite coverage, contain recent information about the ozone distribution. Since ozone is relatively long-lived in the lower stratosphere, first-guess fields can be found by considering it as a passive tracer.

For the 2D analysis, total ozone is implicitly advected using a PV equivalent latitude transformation scheme, similar to that used by Lary et al. (1995). Total ozone is first converted to mean ozone mixing ratio above 450K. This is converted to a function of equivalent latitude using the analysed 450K PV. Analysed PV 6 hours later is then used to convert the effective mixing ratio back to total ozone.

Advection of the ozone mixing ratio distribution is done explicitly using Reverse Domain Filling (RDF), first described by Sutton et al. (1994). The ozone mixing ratio at the new grid point is found by obtaining the location of the grid point at the previous analysis time. This is done on 15 isentropic levels between 340 and 750K, located approximately between the tropopause and 10hPa. Persistence (the previous analysis) is used below 340K and above 750K. Persistence is also used in the tropics, since analysed wind fields are of poor quality. The first-guess total ozone field is found by integrating the first-guess ozone mixing ratio field.

3. 2D ANALYSIS

The 2D analysis has been run using TOVS total ozone for the period from September to November 1994. The TOVS data was obtained over the Global Telecommunication System (GTS) from NOAA NESDIS, and is derived using the algorithm described

by Neuendorffer (1996). The PV coordinate transformation method was run using analysed fields from the Australian Bureau of Meteorology's Global Assimilation and Prediction (GASP) model. Figure 1 shows the daily mean difference of the TOVS analysis with respect to data from the Meteor-3 TOMS instrument averaged over 15 September to 24 November 1994.

Comparisons were only made for days when TOMS was present. Much of the difference at high latitudes in the Northern Hemisphere can be explained by Meteor-3 TOMS observing low total ozone at high solar zenith angles. Over much of the globe, TOMS minus TOVS differences are less than ±2.5%. Areas of larger differences highlight some of the drawbacks of the TOVS data. For example, the areas of significant differences, here defined as the difference exceeding zero by greater than two standard deviations, in the tropics is a result of the TOVS algorithm retrieving ozone amounts that are too low as a result of cirrus contamination. The large areas of significant differences over deserts, particularly the Sahara occurs because the TOVS algorithm also cannot account for the high surface albedo in the 9.7μm ozone channel. Modifying the background radiance used by the algorithm can reduce these effects (not shown) but does not eliminate them entirely. The results here reflect those shown by Neuendorffer (1996).

The TOVS analysis also sees too much ozone over high altitude regions, such as the Tibetan Plateau, the Andes and Antarctica. It is possible that this is because the retrieval algorithm does not explicitly account for altitude. A related explanation is that TOVS cannot retrieve brightness temperatures very well over ice, resulting in poor quality or missing total ozone over high-altitude areas. Differences over the Antarctic are complicated by the fact that much of the information there is from the first-guess field only, and it is believed that this introduces additional errors in the analysis.

4. TOTAL OZONE FORECASTS

As noted in section 2, the PV coordinate transformation scheme can be used to generate total ozone forecasts. Operationally, this has been implemented at the Australian Bureau of Meteorology as part of the UV forecast scheme (Lemus-Deschamps et al., 1997). 36 hour forecasts of total ozone are generated using forecast PV fields from GASP, which are then used to obtain the global UV distribution for the following day.

In this section we look at the effectiveness of the PV coordinate transformation scheme in forecasting total ozone over 36 hours. The variance explained is found by calculating the reduction in standard deviation when the forecast is compared to the analysis relative to the persistence-analysis comparison. This is calculated as follows

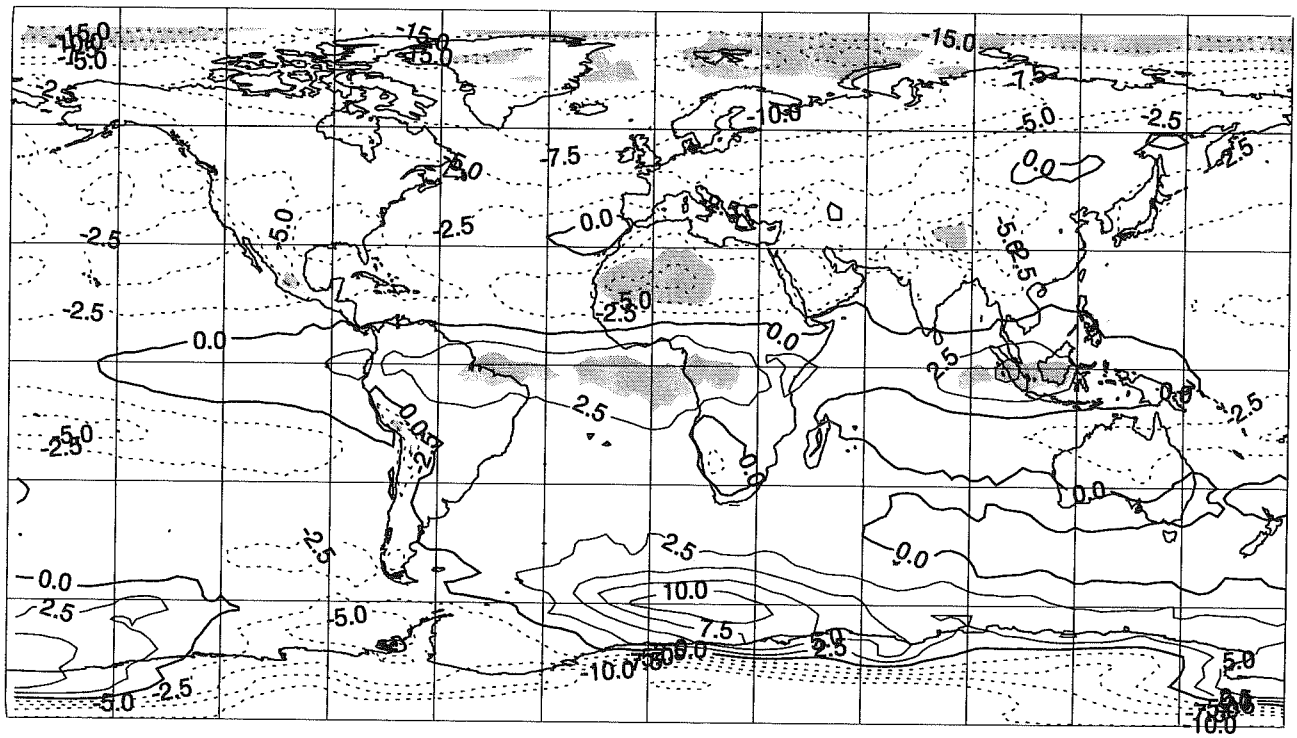


Figure 1. Daily mean TOMS minus TOVS analysis difference averaged over 15 September to 24 November 1994. The difference is shown as a percentage of TOMS. Shading indicates areas where the difference exceeds zero by greater than two standard deviations.

$$\text{Var. Ex.} = \left(1 - \frac{\sigma_f^2}{\sigma_p^2}\right) \times 100 \qquad (1)$$

where σ_f is the forecast minus analysis standard deviation and σ_p is the persistence minus analysis standard deviation for 36 hours. The zonal mean variance explained for daily 23Z analyses averaged over 2 September to 21 December 1997 is plotted as a solid line in figure 2. The dotted line shows the zonal mean σ_p averaged over the same period. The ozone forecast scheme can explain up to 55% of the variance over 36 hours at Southern Hemisphere high latitudes, and around 40% of the variance at mid-latitudes in the Northern Hemisphere.

In the tropics the scheme does less well, explaining less than 20% of the variance. This is partly because PV is nearly zonal in the tropics. However, the variability of tropical ozone is low, with σ_p around 10DU, so it is difficult to improve on persistence in any case. That the forecast scheme captures the dynamic variability of total ozone is also shown in figure 2 by comparison with the persistence standard deviation. The largest reduction in forecast error occurs where the day-to-day variability is greatest.

5. 3D OZONE ANALYSIS

3D ozone analyses have been obtained for September and October 1994. Ozone mixing ratio data was used from all available satellite instruments, namely MLS, HALOE, SBUV/2, SAGE II and POAM II. In addition, total ozone from the TOVS and SBUV/2 instruments was used, although this has little impact on the ozone mixing ratio analysis. The RDF runs to obtain the first-guess fields were made using GASP analysed winds smoothed to T21 truncation.

In order to validate the analysis, comparisons were made with ozonesonde data primarily obtained from the World Ozone and Ultraviolet Data Center (WOUDC). Since ozonesonde launches are relatively infrequent, generally only once per week, sonde profiles for the two month period have been grouped into geographic regions. Mean analysis minus sonde differences for 4 regions are shown in figure 3. In general, differences in the lower stratosphere are less than ±10%, and often less than ±5%. For comparison, instrument error in the lower stratosphere is typically 5-10%. Above 10hPa, the analysis increases relative to the sonde data because of the decreased pump efficiency of the sonde, resulting in the sonde observing less ozone. Differences are much larger below 100hPa, although the ozone mixing ratio is very low in the troposphere. The large differences are because, in general, satellite ozone profiles are not retrieved below about 100hPa.

Comparisons with ozonesondes over the Antarctic (not shown) are less good. Analysed ozone is up to 1.5ppmv too low above 20hPa at the South Pole, and around 1ppmv too high above 15hPa at Syowa and

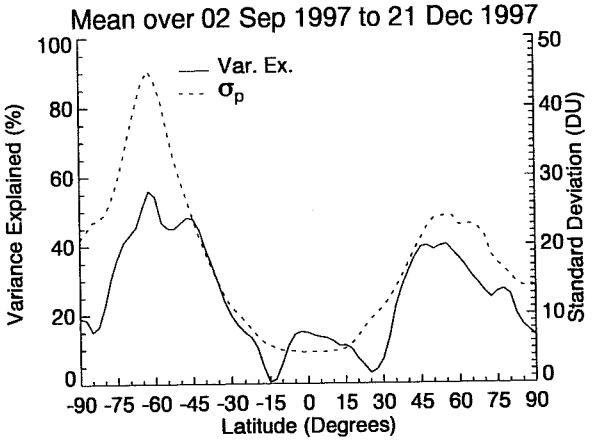


Figure 2. Zonal mean of the variance explained (solid line) and σ_p, (dotted line) for 36 hours to 23Z averaged over 2 September to 21 December 1997.

Marambio on the Antarctic coast. Possible reasons for this are investigated in the next section.

6. FURTHER ISSUES

Figure 4 shows the analysed ozone mixing ratio at 40hPa for 23Z 24 September 1994. There are two features that highlight flaws in the ozone analysis.

The first is seen in the 2.5ppmv contour in the tropics. This shows a physically unrealistic wave-14 structure. In fact, a polar-orbiting satellite has 14 orbits per day, indicating that this structure is due to satellite tracks. One of the assumptions of SI is that all observations are unbiased. Thus the fact that MLS is too low at 46hPa, whereas SBUV/2 is probably too high in the lower stratosphere, is not accounted for by the ozone analysis scheme.

A second unrealistic feature is the negative ozone mixing ratio, shown by the shaded regions in figure 4, within the Antarctic ozone hole. It is possible that this caused by horizontal extrapolation, but the addition of POAM II data did not remove this problem. It is more likely to be caused by the vertical error correlation function. This was derived from tropospheric radiosonde data. It appears to be too sharply peaked in the lower stratosphere, thus decoupling the analysis levels.

The first-guess error field also needs careful consideration. Previous experiments have shown that if it is too high, it causes problems in the analysis. Manual construction of the first-guess error field may also lead to misrepresentation of the tropical vertical gradient.

Finally, there is the problem of the asymptote at zero mixing ratio. This means that the SI assumption of normally distributed errors breaks down in extreme conditions, making physically unrealistic values possible. The analysing the log of mixing ratio could be investigated as a way around this problem.

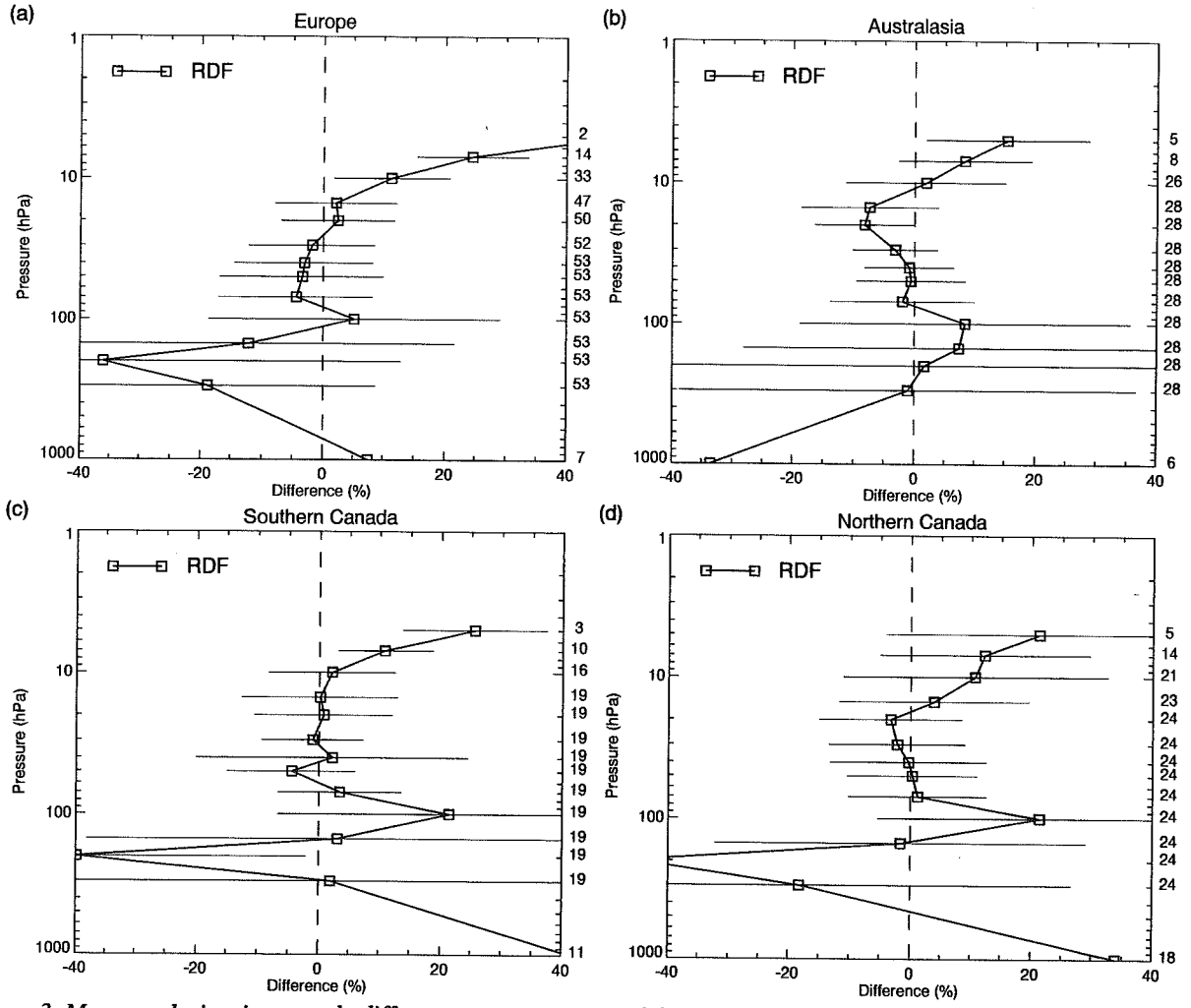


Figure 3. Mean analysis minus sonde difference as percentage of the sonde for 4 regions. Horizontal lines indicate the 1σ standard deviation. The number of comparisons at each level is given at the right of the figure.

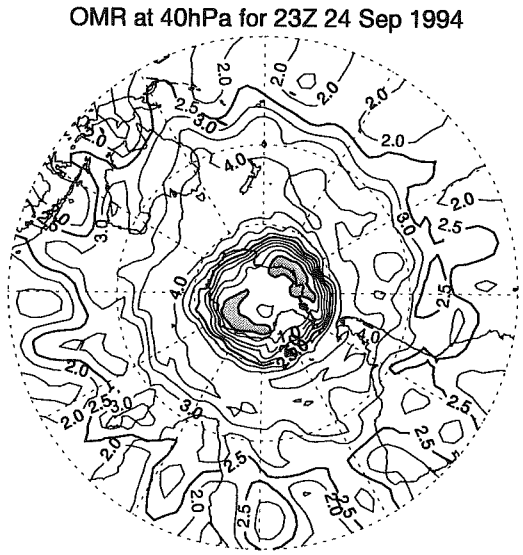


Figure 4. Analysed ozone mixing ratio at 40hPa for 23Z 24 September 1994. Shading indicates regions of negative ozone mixing ratio.

ACKNOWLEDGMENTS

I am particularly grateful for the efforts of the teams around the world responsible for the running the ozonesonde program. Attendance of the SODA Chemical Data Assimilation Workshop was partially funded by KNMI, The Netherlands.

REFERENCES

Grainger, S. 1998, PhD thesis, Monash University, Australia, 182pp.

IPCC, 1994, Climate Change 1994, CUP, 339pp.

Lary D.J. et. al 1995, Q.J.R. Met. Soc, 121, 187-210.

Lemus-Deschamps et al. 1997, WCRP, WMO/TD-No.814, 549-552.

Lorenc, A.C. 1981, Mon. Wea. Rev., 109, 701-721

Neuendorffer, A.C. 1996, J.G.R., 101, 18807-18828.

Sutton R.T. et al. 1994, J. Atmos. Sci.., 51, 2995-3005.

CHEMICAL DATA ASSIMILATION USING THE UKMO UNIFIED MODEL

P. M. Connew[1]

[1]UK Meteorological Office, Middle Atmosphere Group, Bracknell, Berkshire, RG12 2SZ, UK
telephone: +44 (0) 1344 854052 fax: +44 (0) 1344 854046 email: pmconnew@meto.gov.uk

ABSTRACT

Observations of ozone profiles from the Upper Atmosphere Research Satellite Microwave Limb Sounder (UARS MLS) instrument have been assimilated in the United Kingdom Meteorological Office Unified Model (UKMO UM). The stratosphere-troposphere data assimilation system uses a 42 vertical level configuration of the UM. Initial ozone fields are produced by rescaling 2 dimensional chemical model results according to the local potential vorticity. The operational Analysis Correction (AC) assimilation scheme is used to adjust the model fields towards the observation values using small increments on each model timestep. The observation processing includes a component to assess the quality of each observation based on the observation error and the accuracy of the background (or 'initial') fields. In the assimilation, data are weighted according to these errors.

1. STRATOSPHERIC ASSIMILATION SYSTEM

The stratospheric assimilation system in use at the UKMO was set up to provide independent daily analyses, 1.5 days behind real time, to the UARS research community. It is now part of the UKMO operational numerical weather prediction suite producing analyses nearer real time. It uses the UM (Cullen, 1993) and the AC scheme (Lorenc et al., 1991). The stratospheric configuration of the UM comprises of 42 vertical levels with a horizontal resolution of 2.5° latitude by 3.75° longitude. The model domain includes the entire troposphere and stratosphere with the top level situated at 0.28hPa. Some parameterizations in the stratospheric configuration differ from those in the main operational forecast model at the UKMO, including amendment to the longwave radiation scheme and changes to represent effects of Doppler line broadening. It was also necessary to replace the gravity wave drag with Rayleigh Friction at higher altitudes (Swinbank & O'Neill, 1994).

The tracer advection scheme used is positive definite (Cullen & Barnes, 1997). The scheme is mass conserving, preserving the total amount of tracer and the concentration is bounded by initial values. There is no chemistry included in the UM-assimilation system used in these runs.

In addition to the routine production of assimilated fields, the assimilation system has been adapted to re-run past cases with temperature and wind data from the UARS satellite. These include a 3-month assimilation with temperatures from the Improved Stratospheric And Mesospheric Sounder (ISAMS), and periods where stratospheric winds were available (at higher than usual temporal resolution) from the High Resolution Doppler Imager (HRDI) (Connew et al., 1998). The stratospheric assimilation system has also been extended to include observations of atmospheric constituents. Assimilations have now begun with ozone, undertaken as part of the project 'Studies of Ozone Distributions based on Assimilated satellite measurements' (SODA).

2. OZONE ASSIMILATION

MLS views the atmosphere in a direction perpendicular to the orbit path. Latitudinal coverage ranges between 80° in one hemisphere and 34° in the other, also UARS yaws at intervals of around 36 days, switching which hemisphere has the greater coverage. The SODA assimilations use ozone measurements from the 205 GHz channel. The 'level 3AT' data used consist of profiles of ozone observations, measured approximately once per minute along the track. The profile comprises of observations on alternate standard UARS pressure levels from 100hPa to 0.1hPa. Observation errors are estimated as part of the MLS retrieval, and these estimated errors are used to help quality control the data and to calculate the weight given to each observation in the subsequent assimilation.

UARS observations were not available in near enough real time to include in operational assimilation, so a past case was chosen to coincide with ozone initial data availability. Changes were made to the assimilation code to treat ozone profiles in a similar way to radiosonde and satellite temperature profiles, as layer means. An observation profile is taken and it is assumed that the quantity varies linearly with log pressure between observation levels for vertical interpolation, and an average assimilation increment is calculated for each model layer.

3. OBSERVATION PROCESSING

The stratospheric assimilation suite uses NESDIS temperature retrievals from NOAA polar orbiters (with 120km horizontal resolution), radiosonde temperature, winds and humidity, aircraft reports of temperature and winds, satellite observed cloud-track winds and surface

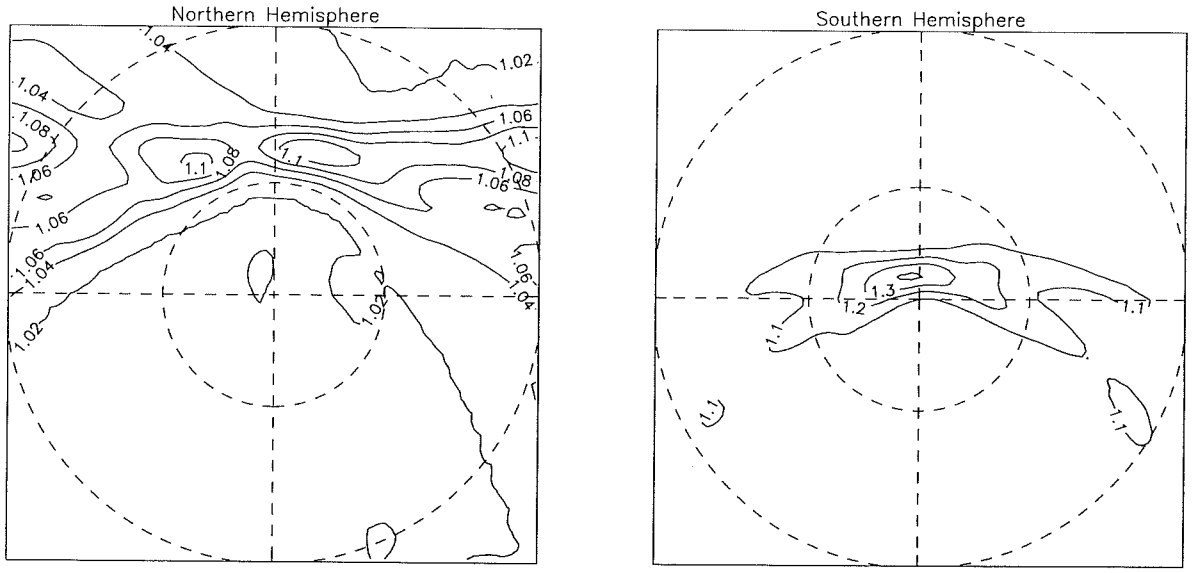


Figure 1. Analysis fields after 6 hours of assimilating MLS ozone fields at 46hPa valid at 18Z 15th March 1995 (arbitrary units).

pressure observations. observations are obtained from the UARS Central Data Handling Facility. Once retrieved and reformatted, all observations are then evaluated in a quality control routine.

A short term forecast is run from initial data fields produced by mapping 2D fields from a chemical model onto the current synoptic situation, using smoothed PV from the operational stratospheric analyses (e.g. Lary et al., 1995). These forecasts provide 'background' data used to compare with observations as part of the quality evaluation, where the value of the observation minus the background value at the observation location, is obtained.

The background error calculation uses the model forecast ozone value plus the local spatial and temporal gradient to produce errors on a 5° by 5° grid. The observation processing uses the given observation error and calculated background errors in a Bayesian statistical quality control method, based on the operational scheme at the UKMO (Lorenc & Hammon, 1988), to calculate a final probability of gross error for each observation. This assumes an a priori initial probability of gross error for each observation type. The quality control algorithms also check for, and discard, observations with negative concentrations.

4. MLS ASSIMILATION RESULTS

Preliminary test assimilations, run over a period of 6 hours, used an initial field constant on pressure levels. This run showed clearly that the observations were affecting the analysis at the insertion point and their influence was spreading between successive observations and tracks (figure 1). Other parameters, such as the a priori initial probability of gross error, were adjusted to give a realistic rejection rate.

The assimilation was then run for 12 hours with the initial fields generated by the PV initialisation technique. The observations are seen to have values similar to the background fields in the lower to mid stratosphere. Figure 2 shows the T+0 initial fields and T+6, T+12 assimilated fields at 46hPa, where the influence of the observations can be seen predominantly at low latitudes. There is a known inadequacy in the MLS observations used, at the lower end of the retrieved profile (Froidevaux et al., 1996). As larger observation errors imply less weight given to those observations in the assimilation, so observations may be expected to have less impact lower down. Figure 3 shows fields valid for the same times at 10hPa, and here the advection of low ozone near the pole is apparent in the time series. Locations of individual observations cannot be identified in the analyses at this level. At 1hPa (figure 4) the impact of the observations is clearly seen in the assimilated fields with consistently lower values, again, most evident at low latitudes. Orbital tracks of the spacecraft can be easily identified when the T+6 fields are compared with the first panel of figure 1, valid for the same time and hemisphere.

Further runs, varying the horizontal correlation scale of the ozone in the assimilation from its initial value of 600km, attempted to improve the observation influence

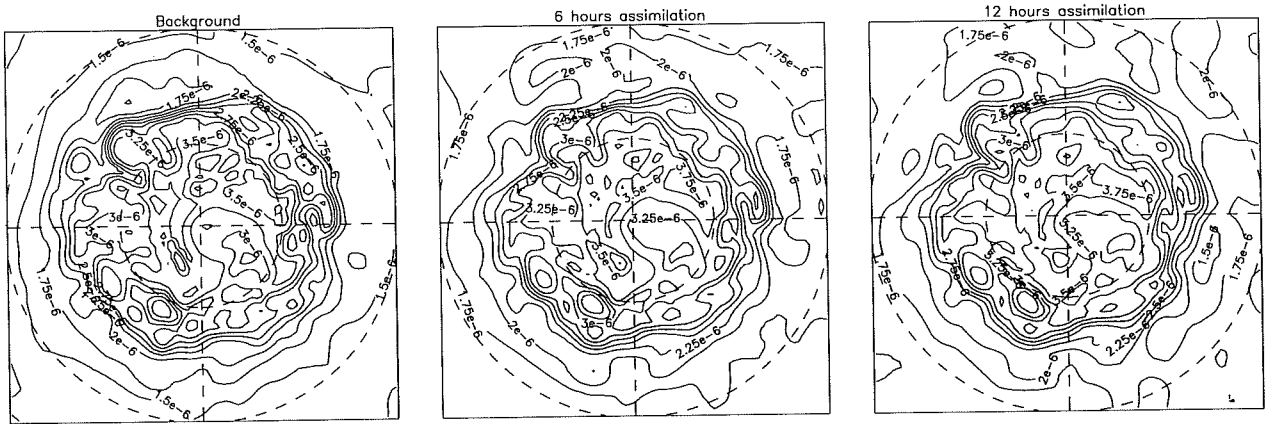


Figure 2. Northern hemisphere ozone concentration (vmr) after T+0, T+6, and T+12 hours of assimilation, at 46hPa.

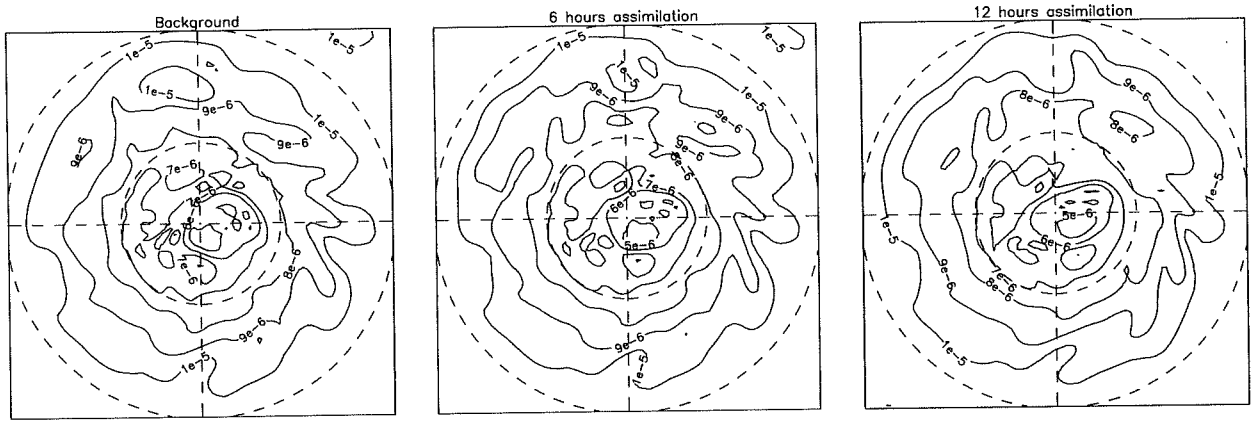


Figure 3. Northern hemisphere ozone concentration (vmr) after T+0, T+6 and T+12 hours of assimilation, at 10hPa.

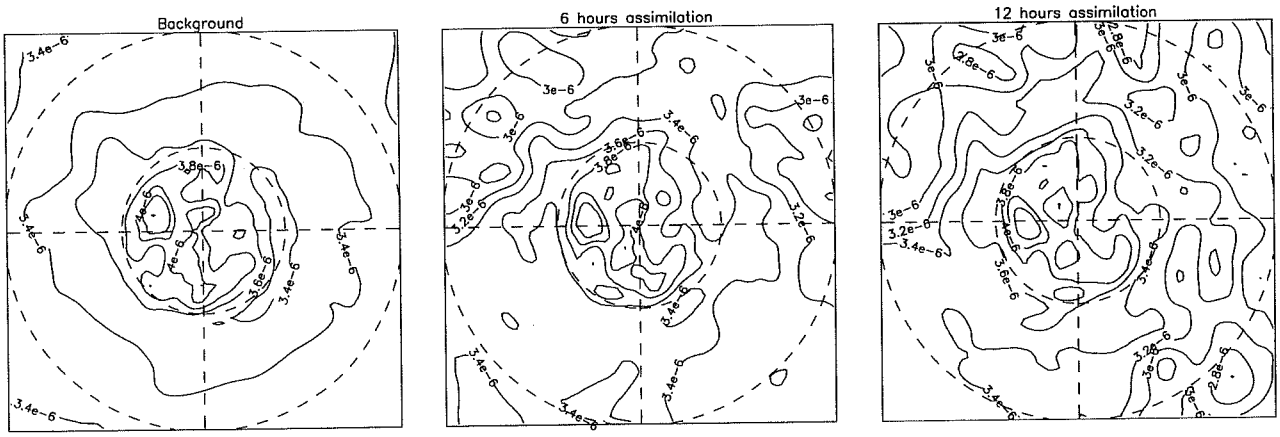


Figure 4. Northern hemisphere ozone concentration (vmr) after T+0, T+6 and T+12 hours of assimilation, at 1hPa.

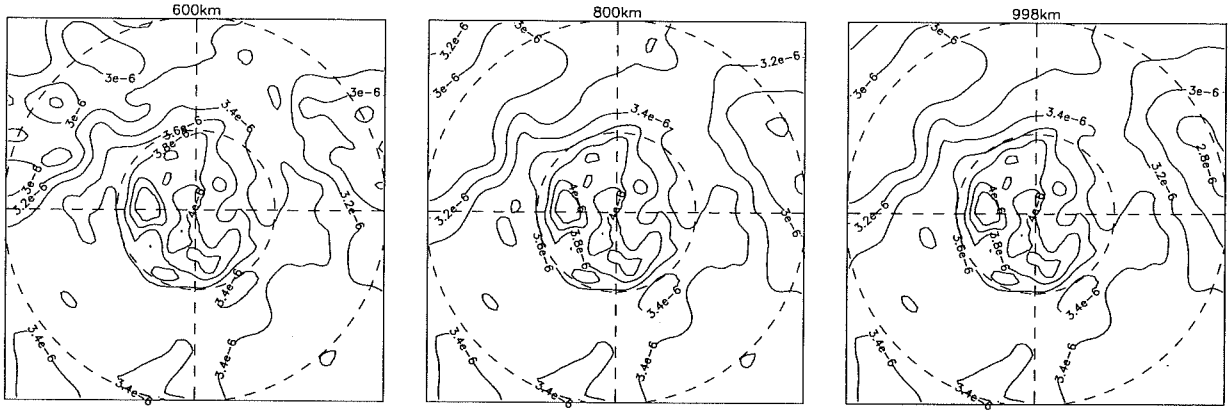


Figure 5. 1hPa Northern Hemisphere ozone concentration (vmr) showing increased spread of observation influence after 6 hours assimilation as horizontal correlation scale is increased. 1st panel 600km, 2nd panel 800km, 3rd panel 1000km.

in data sparse areas. An increase in the horizontal spread of the observations in the analysed field was seen (figure 5) as the horizontal correlation scale increased.

CONCLUSION

MLS observations were successfully assimilated into the UKMO UM for a 12 hour period. These short term assimilation runs showed the analyses were being influenced by the observations. This impact was more apparent at 1hPa, where individual observation locations could be identified, than at lower levels. An improved horizontal spread of observations was achieved by increasing the horizontal correlation scale in the assimilation.

The tracer fields appeared to be advected satisfactorily over the short time periods assimilated. A longer run is necessary to assess whether the tracer advection scheme is too diffusive, although this may not be an issue when enough observations are being assimilated. Introduction of simple chemistry would possibly be beneficial, especially at higher levels.

In collaboration with the Centre for Global Atmospheric Modelling (CGAM) at the University of Reading it is intended to build on the UKMO UARS assimilations to develop data assimilation techniques for both UARS data, and future research satellite data such as Envisat. A new observation processing scheme currently limits the periods of past cases that can be run, but has the potential to become a powerful standalone component for a portable UM-assimilation system.

An upgrade to the stratospheric model is planned as initial studies have shown a benefit in increasing the number of vertical levels in the stratosphere and raising the top of the model. Model physics is soon to be updated to use new schemes recently developed by the UKMO Hadley Centre. 3D variational data assimilation is to be implemented operationally late 1999, and a new dynamics formulation is currently being developed.

REFERENCES

Connew, P., Swinbank, R., Ortland, D., 'Assimilation of Directly Measured Stratospheric Winds into the UK Meteorological Office Unified Model' (submitted to Monthly Weather Review, corrections pending. 1998)

Cullen, M. J. P., 'The Unified Forecast/Climate Model', Meteorological Magazine, 122, 1449, 81-94, 1993

Cullen, M. J. P., Barnes, R. T. H., 'Positive-Definite Advection scheme' Unified Model Documentation paper 11, (Unpublished, copy available from the head of Numerical Weather Prediction division, UKMO), 1997

Froidevaux, L., et al. 'Validation of UARS Microwave Limb Sounder ozone measurements' J. Geophys. Res., 101, 10017-10060, 1996

Lary, D. J., Chipperfield, M. P., Pyle, J. A., Norton, W. A., Riishøjgaard, L. P., 'Three-dimensional tracer initialization and general diagnostics using equivalent PV latitude-potential-temperature coordinates', Quart. J. Roy. Meteor. Soc., 121, 187-210, 1995

Lorenc, A. C., Hammon, O., 'Objective quality control using Bayesian methods', Quart. J. Roy. Meteor. Soc., 114, 515-543, 1988

Lorenc, A. C., Bell, R. S., Macpherson, B., 'The meteorological office analysis correction data assimilation scheme', Quart. J. Roy. Met. Soc., 117, 59-89, 1991

Swinbank, R. S., O'Neill, A., 'A Stratosphere-Troposphere Data Assimilation System', Monthly Weather Review, 122, 686-702, 1994

ASSIMILATION OF TOVS AND GOME TOTAL OZONE WITH THE TM3 MODEL

Henk Eskes and Ad Jeuken

Royal Netherlands Meteorological Institute, Postbus 201, 3730 AE De Bilt, The Netherlands

ABSTRACT

In this paper we describe a data assimilation scheme for the analysis of total ozone satellite observations (GOME, TOVS) based on the global three-dimensional tracer-transport model TM3. TM3 is driven by ECMWF meteorological fields. A parametrized chemistry is added to the model in order to simulate realistic ozone profiles. The assimilation scheme is a simplified Kalman filter approach with a fixed correlation function and a time and space dependent forecast error which is advected by the model and which is recomputed every time observations are analysed. For the vertical covariance three different approaches are tested, but the results are shown to be relatively insensitive to the choice made. Despite the limited resolution in the stratosphere the model is able to capture the dynamical features in the lower stratosphere and the profiles compare favourably with sondes. The model error is surprisingly small.

1. INTRODUCTION

Assimilation of satellite measurements of trace gases into atmospheric models is receiving increased attention in the recent literature. Data assimilation fills the gaps in the data and provides value-added satellite products, i.e. global synoptic maps. The analysis of observations of one or several chemical species with advanced assimilation techniques may provide important information about the consistency and completeness of the chemical modelling (this proceedings). Ozone is also an interesting compound for the numerical weather prediction community. The inclusion of O_3 in the model improves the description of the atmospheric radiation, and may well have a positive effect on the lower stratospheric winds (Rishøjgaard, 1996).

The bulk of the retrieved satellite ozone measurements consists of total column data. In this contribution we describe a data assimilation tool for the analysis of total ozone observations only, based on an off-line three-dimensional tracer transport model. Some preliminary results will be shown for the TIROS Operational Vertical Sounder (TOVS) instruments and the Global Ozone Monitoring Experiment (GOME) spectrometer (Burrows et al., 1993) on board ERS-2. A more detailed account of the TOVS assimilation can be found in a recent paper Jeuken et al., 1998 which will appear in *Journal of Geophysical Research.*

Figure 1 shows the GOME assimilated field on March 9, 1997. Although GOME covers the globe in about three days, the assimilation scheme collects and transports the ozone observations, resulting in synoptic global maps of total ozone. The figure shows a large streamer event, with equatorial air mixed into the NH vortex, leading to low O_3 values above Scotland and Scandinavia. These ozone features compare well with details of the corresponding ADEOS-TOMS total ozone maps for the same day.

2. THE TM3 MODEL

The tracer transport model (TM3) has been adapted from the global Tracer transport Model TM2 (Heimann, 1995), and calculates the horizontal and vertical transport of tracer masses. It is driven by the meteorological fields (wind, surface pressure, temperature and humidity) from the European Centre for Medium-Range Weather Forecasts (ECMWF) model. These fields are updated every 6 hours. In this study the TM3 model has 19 vertical sigma-pressure layers between the surface and 10 hPa and a horizontal resolution of $3.75° \times 5°$. In the near future the horizontal resolution will be improved to 2.5°. Apart from three-dimensional advection the model accounts for subgrid-scale convection and diffusion. The advection of tracers in the model is calculated with the slopes scheme (Russell and Lerner, 1981). In this linear scheme extra information about the spatial distribution of the tracer is stored by means of three directional derivatives that are recomputed every time step. With this, the scheme is able to preserve small-scale features of the order of one grid cell.

2.1. Chemistry parametrization

The production and loss of ozone in the stratosphere is described by the chemistry parametrization developed by Cariolle and Déqué, 1986. The mixing ratio of ozone (χ) is calculated by

$$\frac{d\chi}{dt} = \langle S \rangle + \langle \frac{\partial S}{\partial \chi} \rangle (\chi - \langle \chi \rangle)$$

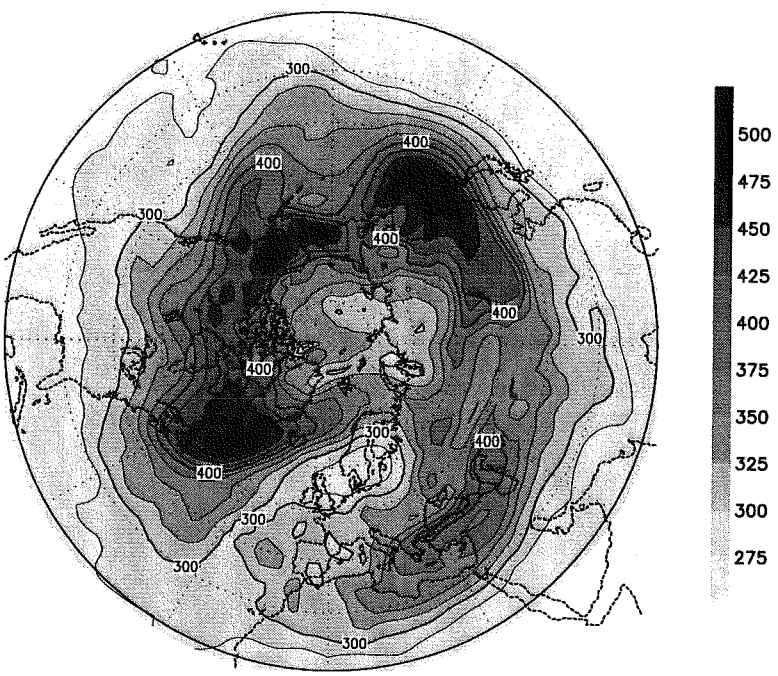


Figure 1. Assimilated northern hemisphere ozone field, using GOME total ozone observations. March 9, 1997. Scale in Dobson Units

$$+\langle\frac{\partial S}{\partial T}\rangle(T-\langle T\rangle)+\langle\frac{\partial S}{\partial \Phi}\rangle(\Phi-\langle\Phi\rangle) \quad (1)$$

The right-hand side consists of a source term, a driving force towards an equilibrium ozone mixing ratio $\langle\chi\rangle$, a response to a change in temperature T and a radiation term depending on the amount of ozone $\langle\Phi\rangle$ above the current position. The relaxation time (related to $\langle\frac{\partial S}{\partial \chi}\rangle$) depends strongly on height, being of the order of months in the lower stratosphere, and shorter than the model time step for the highest Cariolle layers.

In our implementation with the TM3 model we found that the coefficients of Cariolle give rise to a model bias. Therefore we decided to replace the ozone equilibrium values $\langle\chi\rangle$ by an ozone climatolology (Fortuin and Kelder, 1998). As discussed by Wauben et al., 1998 the tropospheric production and loss nearly balance, and the budget is described largely by the influx from the stratosphere and deposition at the surface. Therefore we only apply a dry surface deposition and neglect tropospheric chemistry.

2.2. Vertical resolution

The ECMWF meteorological fields provide a state-of-the-art description of the meteorological fields. However the present operational model is not developed to describe the stratosphere. The vertical resolution in the higher atmosphere is coarse with only 3 layers between 50 and 10 hPa. This is a point of concern since the top layers contain a major part of the atmospheric ozone. In Fig. 2 an example of a model ozone distribution (19 ECMWF layers) is compared with an ozone sounding at De Bilt. Note that there are a reasonable number of layers to describe the lower stratosphere. Because the chemistry is strongly height dependent, we choose to perform the stratospheric chemistry step on the layers as defined by Cariolle. Before the chemistry step the ozone in the top model layers is distributed over the Cariolle levels according to the ozone distribution $\langle\chi\rangle$, next a chemistry step is applied, and finally the ozone is collected in the top model levels. This procedure gives an ozone distribution as shown on the left side of the picture.

3. DATA ASSIMILATION APPROACH

Apart from analysing the ozone field, an estimate of the space and time dependence of the forecast error is made. Since a data assimilation scheme is very efficient in removing biases between the model forecast and the observations that are analysed, this 'error' should be interpreted as a skill of the model to predict the next GOME retrieved ozone observation. Biases that may exist between the GOME retrieved product and retrieved total ozone from other instruments, such as for instance TOMS or TOVS, will be present in the assimilated result as well.

The basic Kalman filter equations serve as the starting point of our approach. The first two equation describe the application of the model M, and the analysis of the ozone field by the incorporation of the

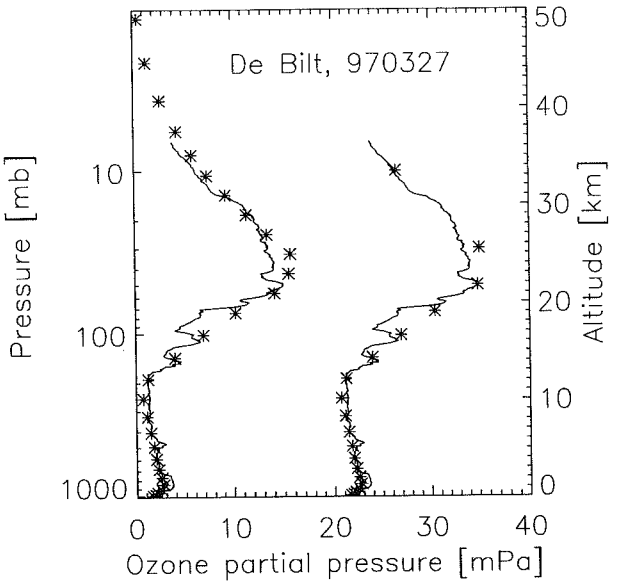


Figure 2. Ozone on the 19 model layers (right) and for an extended set of layers in the stratosphere using the Cariolle vertical levels (left). For comparison ozone sonde data from De Bilt is shown.

observations $\vec{y}^{obs}$ in the model forecast $\vec{x}^f$

$$\vec{x}_{t+1} = \mathbf{M}\vec{x}_t \tag{2}$$

$$\vec{x}_t^a = \vec{x}_t^f + \mathbf{B}\mathbf{H}^T \left(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{O}\right)^{-1} \left(\vec{y}^{obs} - \mathbf{H}\vec{x}_t^f\right) \tag{3}$$

The second equation is the statistical interpolation formula (Daley, 1991). **H** calculates (interpolates) a prediction of the observation from the model ozone values. **O** is the observation error covariance.

The third Kalman equation describes the propagation of the forecast error covariance **B**. (**L** is the linearized model, a matrix. Q is the model error)

$$\mathbf{B}_{t+1} = \mathbf{L}\,\mathbf{B}_t\,\mathbf{L}^T + \mathbf{Q} \tag{4}$$

The diagonal elements of this equation are modelled by ($\sigma_i = \sqrt{B_{ii}}$)

$$\vec{\sigma}^{t+1} = \mathbf{M}\,\vec{\sigma}^t + \text{model error term} \tag{5}$$

Thus the standard deviation is transported in the same way as the ozone field.

The equation that describes the decrease of the forecast covariance due to the new observations

$$\mathbf{B}^a = \mathbf{B}^f - \mathbf{B}^f\,\mathbf{H}^T\,[\mathbf{O} + \mathbf{H}\,\mathbf{B}^f\,\mathbf{H}^T]^{-1}\,\mathbf{H}\,\mathbf{B}^f \tag{6}$$

is solved for the diagonal elements only.

The off-diagonal elements of the covariance are parametrized by means of a correlation function that depends only on the distance between the two positions. The forecast covariance is modelled as

$$B_{im;jn} = \sigma_i\sigma_j\rho(|\vec{r}_i - \vec{r}_j|)\beta_{mn}^{(i)(j)} \tag{7}$$

Note that the standard deviations σ_i here are two-dimensional total-column quantities, expressed in Dobson units, and i is a horizontal lat-lon coordinate. The normalised factor $\beta_{mn}^{(i)(j)}$ describes the vertical covariances between layers m and n, and depends (weakly) on the horizontal coordinate as well.

GOME is flying on board of the ERS-2 polar-orbiting satellite. Since GOME measures the reflected solar radiation in the visible and UV, it measures only about half of the time. To profit optimally from Eq. 3, all observations of one GOME track are analysed together at the appropriate model time step (see figure 3)

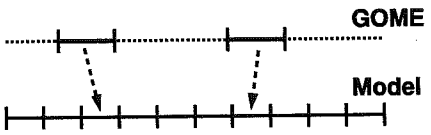


Figure 3. The observations of one GOME track (about 40-50 minutes) are analysed together in one time step.

4. THE VERTICAL PROBLEM

The three-dimensional TM3 model field is updated with two-dimensional measurements only. The theory on data assimilation (see above) shows that the analysis increment is distributed over the vertical with a weight proportional to the vertical covariances (factor β_{mn}). These vertical factors, however, can not be determined from the ozone columns, and additional information is needed. In order to test the sensitivity of the assimilation approach to the choice of the vertical variances, we introduce three different forms. 1) The variance β_{mm} is proportional to the actual ozone mass in layer m. 2) The vertical variance is proportional to the time variation of the model ozone field, a function of the latitude. 3) We adopt an NMC approach (a comparison of two forecasts for the same time, but started from two analyses with a 24 hour difference) to estimate the vertical covariance. These three choices are shown in Fig. 4. Since most of the variation occurs in the lower stratosphere, methods 2 and 3 have there maximum at lower altitudes than method 1.

5. RESULTS

The critical parameter in data assimilation is the ratio between the forecast and observation error. One approach to estimate this ration is to vary the forecast error such that the forecast skill is optimised. This is shown in Fig. 5 in the case of TOVS total ozone. As can be seen the optimum value is around 0.55. This suggest that the model forecast is almost twice as accurate as the observations the forecast is based on! In the case of GOME this ratio at the minimum is roughly 1.5, showing that GOME has less noise and is probably more accurate.

In figure 6 we show the TOVS forecast minus observation departure statistics as a function of the horizon-

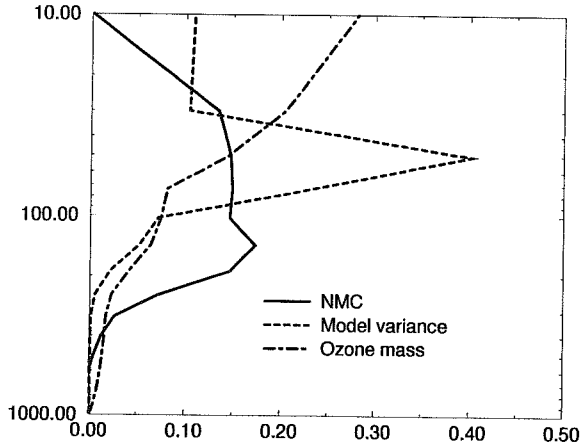


Figure 4. Different choices for the vertical distribution β_{mn}, based on the ozone amount, the variance of ozone, and an NMC estimate of covariances.

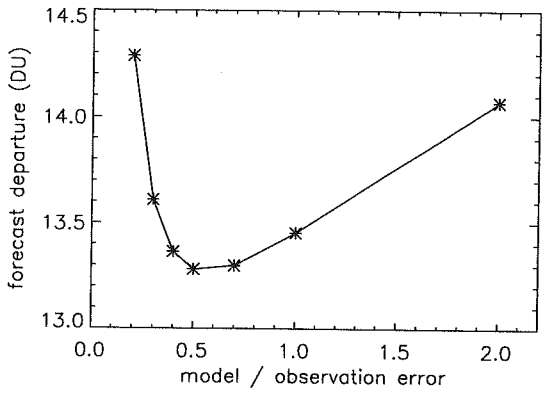


Figure 5. The forecast skill as a function of the ratio between the model and observation error, in the case of TOVS data. Period: April 1992.

tal distance between the two observations (stars). In the assimilation the data of two satellites, NOAA 11 and 12, is used. This top curve is the sum of the forecast and the observation covariance. When the ratio of Fig. 5 and the NMC method are combined, we can make a rough estimate of the forecast part (filled diamonds). The remainder may then be attributed to correlations between the TOVS observations. This shows a rapid decay, with a long-range correlation tail (distances > 500 km). The tail is changing when only NOAA 11 or NOAA 12 data is used, supporting this interpretation. These regional-scale correlations may well reflect the sensitivity of the retrieval on sand (desert), orography (mountains), clouds, and snow or ice (polar regions).

Figure 5 suggests that the model has a good forecast skill. This can be shown more directly by plotting the departures as a function of the forecast time. This is illustrated in Figure 7 for GOME data, showing in-

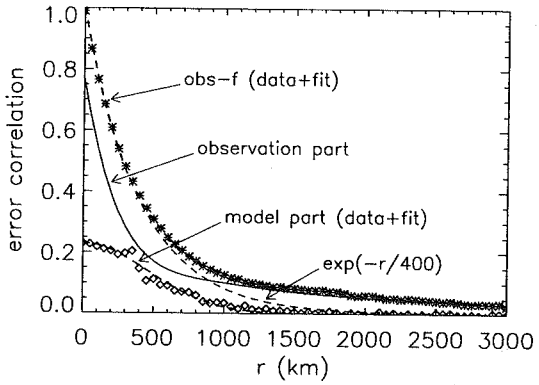


Figure 6. The correlation functions for the assimilation of TOVS observations, April 1992. The top curve (stars) is the forecast minus observation statistics. The lower curve is an estimate of the forecast error covariance, and the remainder is attributed to the TOVS observations (solid line).

deed a surprisingly small deterioration of the forecast skill versus time. Although in this particular latitude band for this particular period the bias is almost zero, in general part of the increase of the departure between the model and observations is due to the bias. This bias is latitude dependent and may be (partly) related to imperfections in the parametrized chemistry scheme. A typical departure increase (model error + bias) is 0.6 DU per day. An important reason for this good result is the realistic meteorological analysis of the ECMWF model that drives the ozone transport model. Increasing the model resolution may affect the model error.

In the last figure we show a validation with an ozone sonde from De Bilt. Three model runs, with the three vertical weight functions discussed in Fig. 4, are compared. As can be seen, despite the large differences between the vertical variance functions, the assimilation results are very similar. One reason for this is the small increments made in the analysis step. Secondly the model tends to an "equilibrium" profile: high in the stratosphere the profile is determined by the chemistry, and in the troposphere by the deposition and vertical mixing. In the lower stratosphere the (chaotic) mixing dictated by the wind field and the patterns of low and high pressure force the ozone towards a distribution which partly erase the information in the starting ozone field.

6. CONCLUSIONS

In conclusion, we have presented details of the assimilation of total ozone satellite observations with the TM3 model. An approach is implemented to estimate the time and space dependence of the forecast error. The forecast error of TM3 was shown to be small for total ozone. This is evidenced by a small optimal ratio between the model and observation error, and by a slow deterioration of the model forecast. This may come as a surprise, given the course resolution of

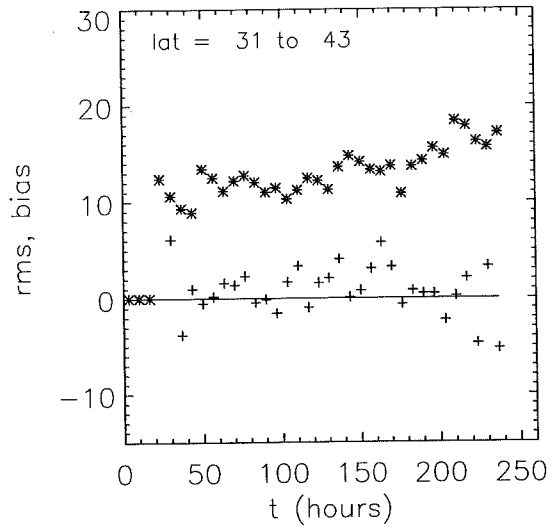


*Figure 7. The forecast - GOME observation mismatch (bias, +, and error, *) as a function of the time since the last analysed measurement. Period: 10 days. Data collected for latitudes between 31 and 43 degrees north. The average slope of the curve provides an estimate of the model error.*

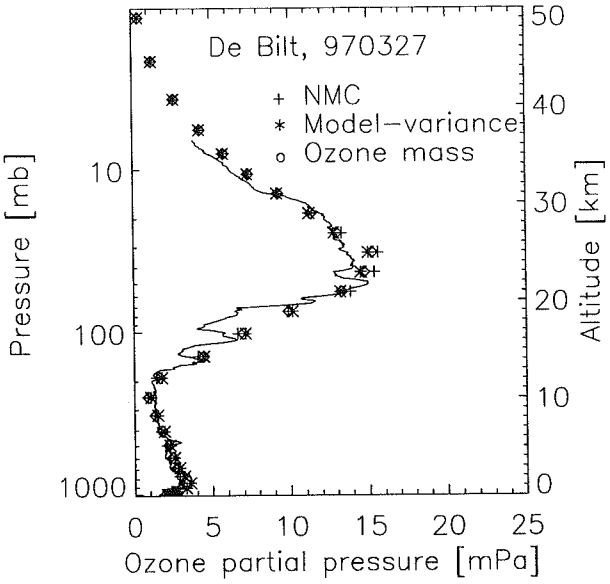


Figure 8. Comparison of the assimilation of GOME data with an ozone sonde at De Bilt. The different curves correspond to the different vertical weight functions β.

the current operational ECMWF model in the stratosphere. The comparison with sondes show that the vertical ozone profiles are realistic. In general, ozone transport models driven by realistic meteorological fields are well capable of describing the main spatial and temporal variation of the ozone layer occurring in the lower stratosphere.

REFERENCES

Burrows, J., *et al.*, Global Ozone Monitoring Experiment, Interim science report, Tech. Rep. SP-1151, ESA, ESTEC, Publication Divisions, Noordwijk, The Netherlands, 1993.

Cariolle, D. and M. Déqué, Southern Hemisphere medium-scale waves and total ozone disturbances in a spectral general circulation model, *J. Geophys. Res., 91,* 10825, 1986

Daley, R., *Atmospheric Data Analysis,* 457 pp., Cambridge Univ. Press, New York, 1991.

Eskes H.J., A. Piters, P. Levelt, and M. Allaart, Variational assimilation of ozone total column satellite data in a 2D lat-lon tracer-transport model, preprint, 1998.

Fortuin, J.P.F. and H.M. Kelder, An ozone climatology based on ozonesonde and satellite measurements, accepted for publication *J. Geophys. Res.,* 1998.

Heimann, M., The global atmospheric tracer model TM2, *Technical report no. 10,* Deutsches Klimarechenzentrum (DKRZ), Hamburg, Germany, 1995.

Jeuken, A. B. M., H. J. Eskes, P. F. J. van Velthoven, H. M. Kelder, and E. V. Hólm, Assimilation of total ozone satellite measurements in a three-dimensional tracer transport model 1998, accepted for publication in *J. Geophys. Res.*

Lorenc, A. C., Optimal nonlinear objective analysis, *Q. J. R. Meteorol. Soc., 114,* 205, 1988.

Neuendorfer, A. C., Ozone monitoring with TIROS-N operational vertical sounders, 1996, *J. Geophys. Res. 101,* 18807

Parrish, D.F. and J.C. Derber, The National Meteorological Center's spectral statistical-interpolation analysis system, *Mon. Wea. Rev., 120,* 1747, 1992.

Riishøjgaard, L.P., On four-dimensional variational assimilation of ozone data in weather-prediction models, *Q. J. R. Meteorol. Soc., 122,* 1545, 1996.

Russell, G.L. and J.A. Lerner, A finite difference scheme for the tracer transport equation, *J. Appl. Meteorol., 20,* 1483, 1981.

Wauben, W.M.F., J.P.F. Fortuin, P.F.J. van Velthoven and H.M. Kelder, Comparison of modelled ozone distributions with sonde and satellite observations, *J. Geophys. Res., 103,* 3511, 1998.

AN O.S.S.E. USING TOVS TOTAL OZONE DATA

A. Peuch, J.-N. Thépaut, J. Pailleux

Centre National de Recherches Météorologiques, Météo-France,
42 avenue G. Coriolis, 31057 Toulouse Cedex, France

Aline.Peuch@cnrm.meteo.fr
Jean-Noel.Thepaut@meteo.fr
Jean.Pailleux@meteo.fr

ABSTRACT

An Observing System Simulation Experiment (OSSE) has been performed to investigate the impact of total ozone data upon analysis (especially of winds) in Numerical Weather Prediction models (NWP). The assimilation system used is the four-dimensional variational (4D-VAR) algorithm available in the NWP IFS/ARPEGE. Ozone columns have been generated by the model to be representative of TOVS data coverage, and assimilated over a twelve-hour period, with different observational errors. Results show that very accurate measurements of ozone columns provide information on thermodynamical fields in areas with strong horizontal gradients of total ozone, and that the 4D-VAR system is able to derive well-adapted increments which improve the initial state. However, some negative effects have been observed, between 150 and 250 hPa, for the zonal component. The crude initialisation of the tridimensional ozone field is suspected to be responsible for this degradation of the analysis. A more realistic initial ozone field has been built to perform new assimilation experiments at a higher resolution.

Key words: four-dimensional variational assimilation; ozone columns.

1. INTRODUCTION

NWP models forecast the evolution of the atmospheric state from an initial point called analysis. The analysis is performed using different informations on the atmosphere, such as model outputs and observations. Observations coverage is not homogeneous over the world. In particular, few wind observations are available in the upper troposphere and lower stratosphere. Any indirect information upon wind in these areas is potentially interesting to be assimilated in NWP models.

As an important greenhouse gas, atmospheric ozone is attentively observed, in particular through remote-sensing. The variability of ozone is dominated by transport processes in the upper troposphere and lower stratosphere. In these areas, its lifetime is relatively long and ozone can be considered as a good tracer of the flow. Under these conditions, one can expect to be able to derive some information on the winds from successive observations of ozone structures (Riishojgaard 1996). This idea can be extended to observations of total ozone columns, since ozone quantities in the upper troposphere and lower stratosphere contribute significantly to total ozone values (Vaughan & Price 1991).

The interest for ozone observations is increased by the development of 4D-VAR assimilation systems. They are particularly well-adapted to the assimilation of satellite data which are asynoptic. Moreover they allow to take into account various data, and not necessarily measurements corresponding directly to prognostic variables of model.

2. THE ASSIMILATION SYSTEM

The 4D-VAR algorithm used has been developed jointly by ECMWF and Météo-France for IFS/ARPEGE (Rabier et al. 1996). This system is already operational at ECMWF and tested in research mode at Météo-France with an operational goal towards 1999-2000.

In the 4D-VAR technique, a cost function J is defined, which measures the distance between the model trajectory and different informations about the atmospheric state (direct observations and a previous forecast called background). During the assimilation process, J has to be minimised trough an iterative algorithm. J is a function of a vector x which is formed with some model variables at the initial time of the assimilation. The part of J relative to the distance between the model trajectory and the assimilated observations is computed according to Equation 1.

$$J_o(x) = (H(x) - y^o)^t R^{-1} (H(x) - y^o) \qquad (1)$$

with x the control variable of the minimisation, y^o the assimilated observations, and H an observation

operator which permits to compute the model equivalents to the observations. R is the matrix of the covariances of the observational errors.

In our experiments, the control variable x consists of wind, temperature and surface pressure. It means that only the initial fields of these parameters are likely to be modified to allow the subsequent forecast to fit the assimilated observations better and reduce the cost function. In particular, ozone is not in the control variable of the 4D-VAR minimisation. The initial ozone field is implicitely considered as perfect, since it cannot be modified by the minimisation.

When ozone column observations are assimilated, the model equivalents to the observations are computed using H formed of the forecast of the ozone mixing ratios at the observation points and of the sum of the ozone mixing ratios over the model layers. The only parameter of the control variable to which ozone is directly linked, is the wind. See Equation 2. Ozone is a prognostic variable of the model. It is considered as a passive tracer, simply advected by winds.

$$\frac{\partial R_{O_3}}{\partial t} + V.\nabla R_{O_3} = 0 \qquad (2)$$

with R_{O_3} the mixing ratio of ozone and V the wind vector.

3. METHODOLOGY : AN OSSE

3.1. Set-up of the experiments

Assimilation experiments are made in the context of an Observing System Simulation Experiment (OSSE). Assimilated observations are not real, but generated by the model. Preliminary experiments are made at low resolution : T21/L27. The model top is at 10hPa. The date chosen for the study is February 1997, 19th ; it corresponds to a well-documented case of the FASTEX experiment. An initial tridimensional ozone field has been crudely built : an observed total ozone field from TOVS has been split up along the vertical, using a latitudinal climatology of vertical profiles. As a consequence, the vertical profiles for a given latitude are all homothetic.

3.2. Simulation and assimilation of observations

Total ozone observations have been simulated to be representative of TOVS data coverage, using a model forecast from the operational analysis on 19/02/1997, at 0h UTC. This analysis is considered as the "truth", the exact atmospheric state at 0h UTC. Note that it had been performed with the operational assimilation system used at that time at Météo-France, based upon an optimal interpolation algorithm. The values of the simulated observations are taken equal to the exact values of the columns predicted by the model, at times and locations of real TOVS observations. No noise is added : observations are thus supposed to be perfect.

Two assimilation experiments of simulated ozone observations have been performed. In the first one, the standard deviations of the observational error are taken equal to 10% of the column values (OZ10 experiment) ; in the second one, to 1% (OZ01 experiment). In such a way, we can assess the impact of total ozone data for a wide range of measurement accuracy. Assimilation experiments start from the operational background on 19/02/1997, at 0h UTC, which is a six-hour forecast. The assimilation window is twelve-hour long.

3.3. A reference experiment

In order to be able to assess the amount of information provided by simulated ozone columns, we have performed a reference experiment. Data which were assimilated operationally on 19/02/1997 have been simulated. They have also been assimilated over twelve hours : this is the OPER experiment.

Figure 1 schematizes how observations have been simulated and assimilated in the three experiments OZ10, OZ01 and OPER.

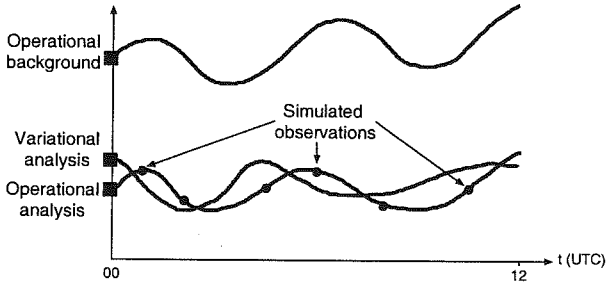


Figure 1. Simulation and assimilation of observations.

4. RESULTS

4.1. Study of wind increments

Given the set-up of the experiments, operational increments (difference between operational analysis and background) are the ideal increments : they correspond to the exact correction to add to the background to obtain the "truth". As a consequence, the impact of assimilated observations in OZ10, OZ01 and OPER can be evaluated by comparing increments generated in each of these three experiments to the ideal increments. Figures 2, 3, 4 and 5 show ideal, OPER, OZ10 and OZ01 wind increments at 200 hPa, on 19/02/97 at 0h UTC.

Ideal and OPER wind increments present a very good agreement, showing that the 4D-VAR system works correctly. Increments generated by OZ10 and OZ01 are smaller, which simply illustrates that ozone observations cannot bring as much information as all the observations assimilated in NWP. They are also smaller in OZ10 than in OZ01, according to the confidence given to observations by the assimilation system. Some reasonable agreement can be locally observed between ideal increments and increments gen-

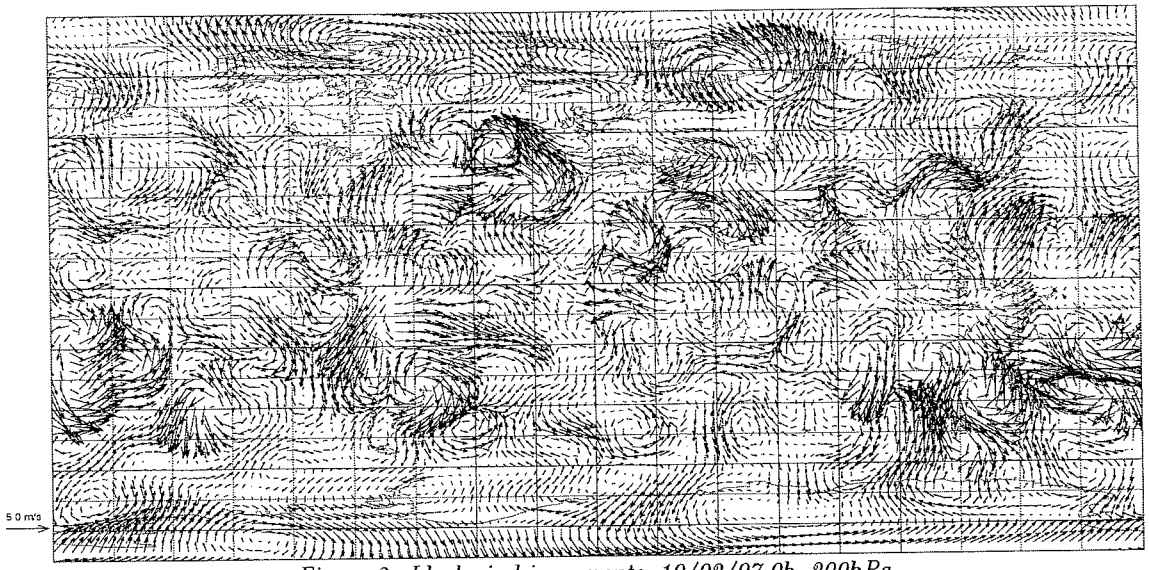


Figure 2. Ideal wind increments, 19/02/97 0h, 200hPa.

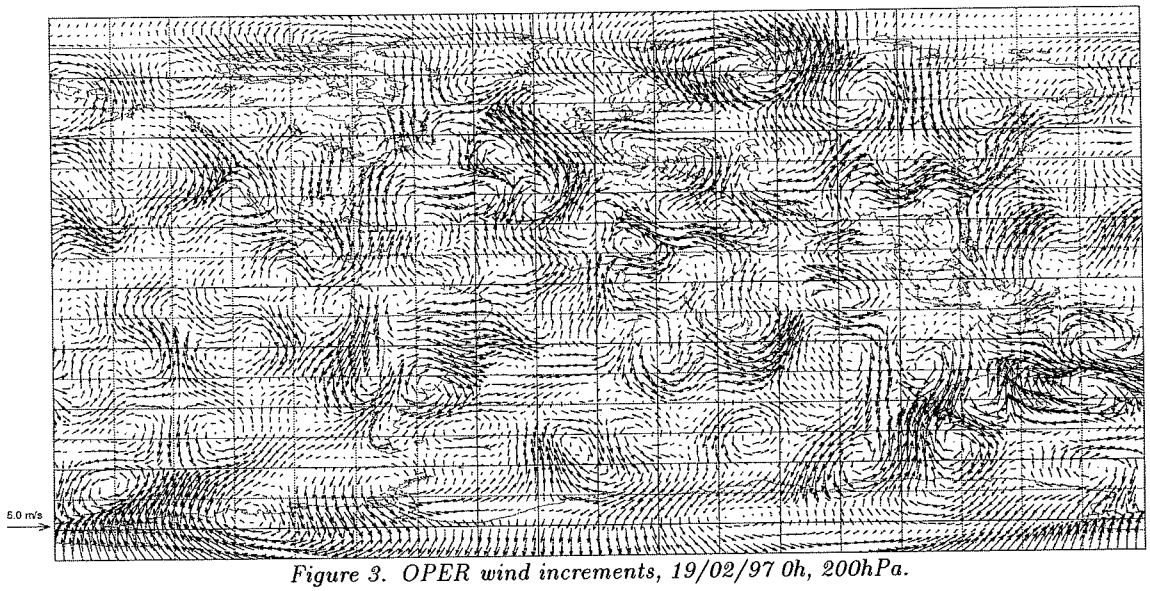


Figure 3. OPER wind increments, 19/02/97 0h, 200hPa.

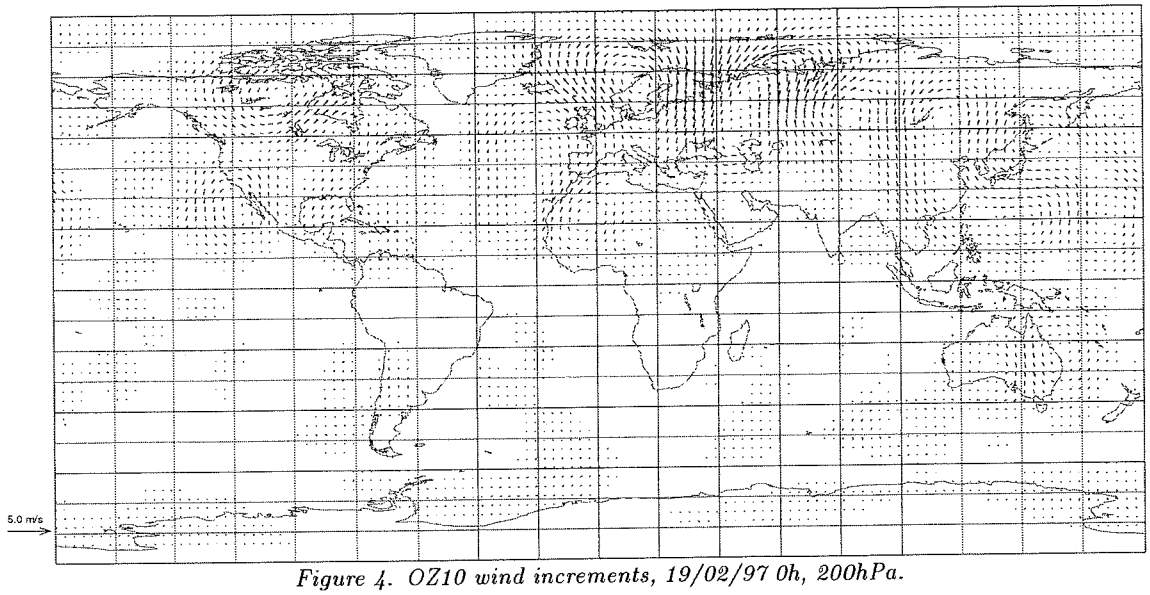


Figure 4. OZ10 wind increments, 19/02/97 0h, 200hPa.

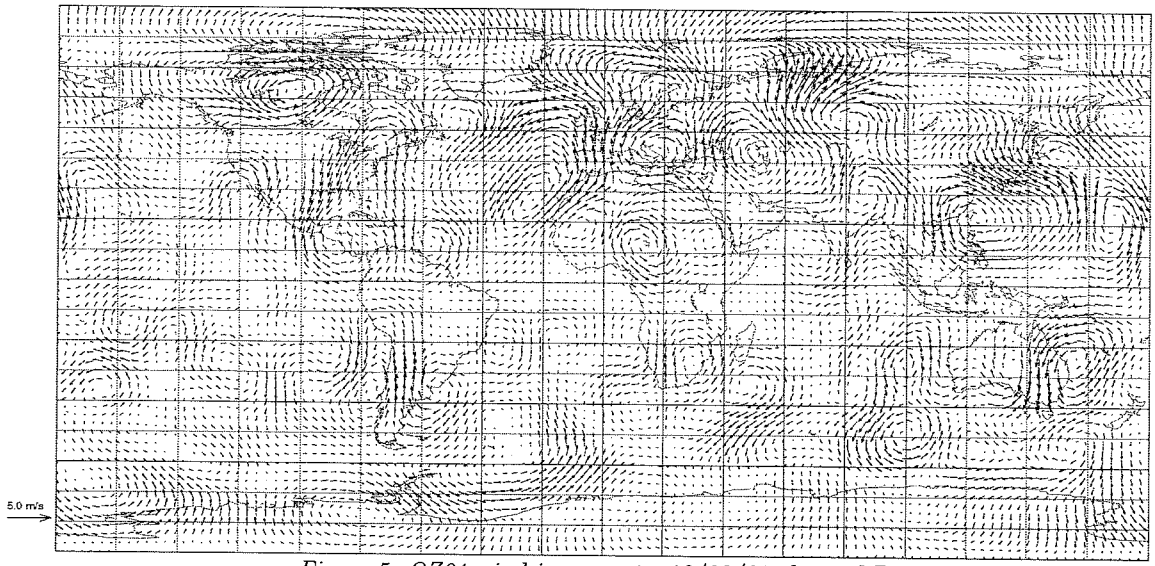


Figure 5. OZ01 wind increments, 19/02/97 0h, 200hPa.

erated by OZ01, especially over the northern hemisphere where the meteorological activity is important at the date studied (e.g. the cyclonic circulation over the north Atlantic ocean). They correspond to areas where total ozone horizontal gradients are strong.

Correlation coefficients between wind increments have been computed. In particular, the correlation between ideal and OZ01 increments computed over the northern hemisphere, is similar to the correlation between ideal and OPER increments computed over some Pacific Ocean areas, where remote-sensing data only are available.

4.2. Fit of wind fields to observations

An other way to assess the impact of ozone columns upon wind analysis, is to investigate how wind fields fit wind observations, before and after the minimisation. Results of the comparison of wind background and analysis to wind observations from radiosoundings are presented for OZ01, in Figure 6. Results are given over the northern hemisphere only ; results on the tropics and the southern hemisphere are rather inconclusive because of the scarcity of radiosoundings observations. We can observe a significant improvement of the definition of the initial state for the meridional component over an important part of the vertical, and for the zonal component above 50 hPa. However, there is also a degradation of the initial state for the wind zonal component between 150 and 250 hPa. We suspect the crude initialisation of the 3D ozone field to be responsible for that, because it weakens the variability of the ozone field, and of the horizontal ozone gradients, in the zonal direction. Concerning OZ10, results are similar, although less significant.

4.3. Impact upon the temperature fields

Temperature is included in the control variable. As a consequence, the assimilation of total ozone observations is likely to generate temperature increments. Although there is not a direct link between ozone and temperature, temperature increments can be expected to be induced by observed wind increments, through the geostrophic relation.

Figure 7 gives the fit of temperature fields to temperature observations from radiosoundings, over the northern hemisphere, at 0 and 12h UTC. A slight positive impact can be noticed upon temperature analyses in OZ01, above 50 hPa. No impact has been detected in OZ10.

5. CONCLUSIONS AND PROSPECTS

Assimilation experiments made in the context of OSSE have shown that ozone columns with TOVS data coverage can provide information on flow in areas with strong horizontal gradients of ozone. A significant positive impact has been observed on the wind meridional component, and also on the zonal component over a more limited part of the vertical. In fact, some negative effects on the zonal component have been noticed between 150 and 250hPa. The crude initialisation of the tridimensional ozone field is suspected to be responsible for that. One can note also a slight positive impact upon temperature analyses above 50hPa, when total ozone observations assimilated are very accurate.

To study if the ozone initialisation is at the origin of the negative effects of the assimilation of total ozone contents on the wind zonal component, OZ01 has been tested with a better initial tridimensional ozone field. Resolution is increased, both on the horizontal (T63, instead of T21) and on the vertical (31 levels, instead of 27). The model top is now at 5hPa. The initialisation uses outputs from REPROBUS, the tridimensional chemical transport model used at Météo-France(Lefèvre et al. 1994). Two initial ozone conditions for 19/02/1997 are built : one is simply the interpolation to our resolution of REPROBUS ozone field on 19/02/1997 ; the other one is a two-

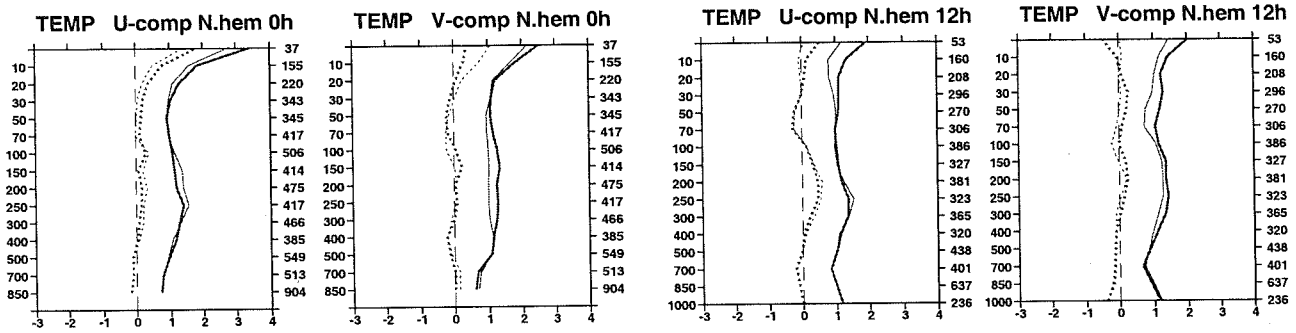


Figure 6. Fit of wind background (bold) and analysis (thin) to observations from radiosoundings, over the northern hemisphere, for the zonal and meridional components, at 0 and 12h UTC : bias (dotted) and standard deviations (solid) of the difference between fields and observations are presented. The number of observations treated at each different level are given at the right of the diagrams. Unit : m/s.

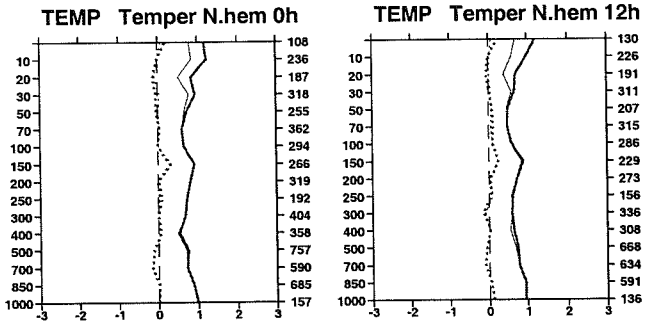


Figure 7. Fit of temperature background and analysis to temperature observations from radiosoundings, over the northern hemisphere, at 0 and 12h UTC. Legend : same as Figure 6. Unit : K.

day forecast from the combination of ARPEGE analysis and interpolated ozone field from REPROBUS on 17/02/1997. The results will be presented in a near future.

The next step will be to adapt developments made by Elias Holm in IFS at ECMWF to the ARPEGE code. It consists principally in putting ozone in the control variable of the minimisation, which implies an additional term due to ozone in the part of the cost function relative to the distance between the model trajectory and the background. Then the assimilation of real total ozone observations and its impact on the model predictions will be investigated.

ACKNOWLEDGMENTS

This work has benefited from several contributors. In particular, we would like to thank Fernand Karcher and Vincent-Henri Peuch for providing information on ozone and for the REPROBUS outputs.

REFERENCES

Lefèvre, F., Brasseur, G.P., Folkins, I., Smith, A.K., Simon, P. 1994, Chemistry of the 1991-1992 stratospheric winter : three-dimensional model simulations, J. Geophys. Res., 99, D4, 8183-8195

Rabier, F., Thépaut, J.-N., Courtier, P. 1996, Extended assimilation and forecast experiments with a four-dimensional variational assimilation system, Q. J. R. Meteorol. Soc., 124, 1861-1887

Riishojgaard, L.P. 1996, On four-dimensional variational assimilation of ozone data in weather prediction models, Q. J. R. Meteorol. Soc., 122, 1545-1571

Vaughan, G., Price, J.D. 1991, On the relation between total ozone and meteorology, Q. J. R. Meteorol. Soc., 117, 1281-1298

MULTIVARIATE OZONE ASSIMILATION IN FOUR-DIMENSIONAL DATA ASSIMILATION

E. V. Hólm, A. Untch, A. Simmons, R. Saunders, F. Bouttier, E. Andersson
ECMWF, Shinfield Park, Reading RG1 9AX, UK

ABSTRACT

Ozone has been included in the ECMWF 4D-VAR data assimilation. The system includes a multivariate coupling between ozone and vorticity background errors. The ozone observations used are retrieved TOVS products. Comparison of preliminary results with TOMS total ozone data show that the ozone chemistry parameterization used for the ozone forecasts are essential for the assimilation of TOVS ozone. However, the parameterization also causes a negative bias in the tropics and positive bias over the poles which the TOVS data can not completely remove. A modified TOVS product, a layer ozone between 31hPa and 600hPa can be obtained from the original total ozone TOVS. Using this layer ozone product gives a closer fit with TOMS. Further improvements of the ozone chemistry parameterization and more ozone data with some vertical resolution are needed before the feedback of ozone on dynamics can be evaluated.

1. INTRODUCTION

At ECMWF ozone has been included in the four-dimensional data assimilation (4D-VAR) system as a part of the stratospheric extension of the centres weather forecast model (Untch et al. 1998). Apart from being able to provide three-dimensional ozone forecasts and ozone fields consistent with the atmospheric dynamics, the inclusion of interactive ozone is expected to have an impact on atmospheric dynamics. First, the use of ozone measurements will affect the model wind fields through 4D-VAR data assimilation. Second, the use of model ozone in the calculation of temperature, humidity, and ozone profiles from TOVS radiances will improve the accuracy of the data from this source. Third, the use of model ozone in the radiation parameterization of the model will directly affect the model temperature.

Prerequisites for investigating the interaction of ozone with dynamics are the use of high quality ozone observations and the modelling of ozone with an accuracy comparable with that of the dynamical model fields. For modelling ozone a good transport module and a reasonably realistic parameterization of sources and sinks arising from atmospheric chemistry and surface deposition are needed.

In this paper we will look at the first ECMWF 4D-VAR version which also includes multivariate ozone assimilation. This version uses TOVS retrieved total ozone products as the only ozone observations, and has a stratospheric ozone chemistry parameterization from Cariolle and Déqué, 1986. The present system is still not complete enough (in particular with respect to observations) to investigate the influence of ozone on dynamics. However, we are able to test the behaviour of the ozone field, and identify some problems which need to be solved before progressing further.

We will start with a short theoretical discussion on the interaction of ozone and dynamics in 4D-VAR, followed by preliminary tests with and without ozone chemistry parameterization, as well as testing two alternative retrieved TOVS ozone products.

2. COUPLING OF OZONE AND DYNAMICS

As an illustration of how ozone (or any other trace gas) and dynamics are coupled in 4D-VAR, we consider the transport of a passive tracer φ by a time-dependent windfield u in a one-dimensional periodic domain of length L. Denoting the model variables as $\vec{q} = (u, \varphi)$, the periodic boundary conditions are $\vec{q}(0,t) = \vec{q}(L,t)$. Then the model with initial conditions becomes

$$\frac{\partial \vec{q}}{\partial t} + u\frac{\partial \vec{q}}{\partial x} = \vec{0} \tag{1}$$

$$\vec{q}(x,0) = \vec{q}^{\,0}(x,0) + \vec{q}^{\,\prime}(x,0) \tag{2}$$

Here $\vec{q}^{\,0}(x,0)$ is the first guess initial condition, which in data assimilation terminology is called the background field and is provided from a model forecast from an earlier timepoint. The goal of 4D-VAR is to determine the correction $\vec{q}^{\,\prime}(x,0)$ to the initial condition in an optimal way using available observations of the atmosphere in the timewindow $0 \leq t \leq T$. These observations we collect in a data vector $\vec{d}$,

$$\vec{d} = \vec{H}(\vec{q}) + \vec{d}^{\,\prime} \tag{3}$$

where observation d_m is associated with an observation operator H_m and an observation error d'_m. What determines the relative fit between background fields and observations are the background error covariance matrix $\mathbf{B}$ and the observation error covariance matrix $\mathbf{O}$. We only need to consider two direct observations,

d_u and d_φ respectively, for illustrating the coupling of tracers and dynamics. Since the measurements are direct, the observation operators are just identity. The observations are independent, with error standard deviations σ_u and σ_φ respectively.

It can be shown that with the above assumptions the optimal solution $\hat{q}$ should fulfill the following set of Euler-Lagrange equations (see Bennet, 1992),

$$\vec{q}^*(x,T) = \vec{0} \tag{4}$$

$$-\frac{\partial u^*}{\partial t} - \frac{\partial}{\partial x}(\hat{u}u^*) = -\vec{q}^*\frac{\partial \hat{q}}{\partial x} + \delta_u \frac{d_u - \hat{u}}{\sigma_u^2} \tag{5}$$

$$-\frac{\partial \varphi^*}{\partial t} - \frac{\partial}{\partial x}(\hat{u}\varphi^*) = \delta_\varphi \frac{d_\varphi - \hat{\varphi}}{\sigma_\varphi^2} \tag{6}$$

$$\hat{q}(x,0) = \vec{q}^0(x,0) + \mathbf{B} \circ \vec{q}^* \tag{7}$$

$$\frac{\partial \hat{q}}{\partial t} + \hat{u}\frac{\partial \hat{q}}{\partial x} = \vec{0} \tag{8}$$

where the background error calculation is

$$\mathbf{B} \circ \vec{q}^* = \int_0^L \mathbf{B}(x,x')\vec{q}^*(x',0)dx' \tag{9}$$

Here $\vec{q}^*$ are the adjoint model variables, and the boundary conditions are $\hat{q}(0,t) = \hat{q}(L,t)$ and $\vec{q}^*(0,t) = \vec{q}^*(L,t)$. Furthermore δ_u and δ_φ are Dirac delta functions at the time and place of the observations.

The background error covariance matrix $\mathbf{B}$ has four elements. The diagonal element b_{uu} describes the univariate spatial covariance of u', and $b_{\varphi\varphi}$ the covariance of φ'. The off-diagonal elements $b_{u\varphi}$ and $b_{\varphi u}$ describe the multivariate covariances of u' and φ'. If the background errors u' and φ' were uncorrelated, then there would only remain the diagonal elements of $\mathbf{B}$ which would make the background error calculations simpler. Derber and Bouttier have shown (see Derber and Bouttier, 1998) that this simplification of the background error calculation can be achieved even if u' and φ' are correlated. Given the background error u', separate the background error φ' into two parts. The first (balanced) part is proportional to u', $\varphi'' = \alpha u'$, and the second (unbalanced) part ϕ' is uncorrelated with u'. The background error covariance matrix now only considers the covariances of u' and ϕ' and is diagonal as wanted. Now the original background error calculation

$$\begin{pmatrix} u' \\ \varphi' \end{pmatrix} = \begin{pmatrix} b_{uu} & b_{u\varphi} \\ b_{\varphi u} & b_{\varphi\varphi} \end{pmatrix} \circ \begin{pmatrix} u^* \\ \varphi^* \end{pmatrix} \tag{10}$$

is transformed to the simpler form

$$\begin{pmatrix} u' \\ \varphi' \end{pmatrix} = \begin{pmatrix} 0 \\ \alpha u' \end{pmatrix} + \begin{pmatrix} b_{uu} \circ u^* \\ b_{\phi\phi} \circ \phi^* \end{pmatrix} \tag{11}$$

We will now examine how the tracer and dynamics are coupled by the data assimilation by considering first only a wind observation, then only a tracer observation, and finally wind and tracer observations together.

2.1. Wind observation only

In this case we see that the adjoint tracer φ^* starts from zero at $t = T$ and remains identically zero during the integration back to $t = 0$. This is because the only source of φ^* is observations involving φ, and there are no such observations. The resulting solution from 4D-VAR should now fulfill the following simpler set of Euler-Lagrange equations (using $\varphi^* \equiv 0$)

$$u^*(x,T) = 0 \tag{12}$$

$$-\frac{\partial u^*}{\partial t} - \frac{\partial}{\partial x}(\hat{u}u^*) = -u^*\frac{\partial \hat{u}}{\partial x} + \delta_u \frac{d_u - \hat{u}}{\sigma_u^2} \tag{13}$$

$$\hat{q}(x,0) = \vec{q}^0(x,0) + \begin{pmatrix} b_{uu} \circ u^* \\ \alpha u' \end{pmatrix} \tag{14}$$

$$\frac{\partial \hat{q}}{\partial t} + \hat{u}\frac{\partial \hat{q}}{\partial x} = \vec{0} \tag{15}$$

Here we see that a wind observation changes the tracer through the background error coupling. More generally in a weather forecast model, all observations can affect a passive tracer through the background error coupling.

2.2. Tracer observation only

The simplified Euler-Lagrange equations for this case are

$$\vec{q}^*(x,T) = \vec{0} \tag{16}$$

$$-\frac{\partial u^*}{\partial t} - \frac{\partial}{\partial x}(\hat{u}u^*) = -u^*\frac{\partial \hat{u}}{\partial x} - \varphi^*\frac{\partial \hat{\varphi}}{\partial x} \tag{17}$$

$$-\frac{\partial \varphi^*}{\partial t} - \frac{\partial}{\partial x}(\hat{u}\varphi^*) = \delta_\varphi \frac{d_\varphi - \hat{\varphi}}{\sigma_\varphi^2} \tag{18}$$

$$\hat{q}(x,0) = \vec{q}^0(x,0) + \begin{pmatrix} b_{uu} \circ u^* \\ \alpha u' + b_{\phi\phi} \circ \phi^* \end{pmatrix} \tag{19}$$

$$\frac{\partial \hat{q}}{\partial t} + \hat{u}\frac{\partial \hat{q}}{\partial x} = \vec{0} \tag{20}$$

Both adjoint variables start from zero at $t = T$, but as soon as the tracer observation appears during the integration back to $t = 0$ then φ^* becomes nonzero. The adjoint wind u^* then also becomes nonzero through the source term $-\varphi^*\frac{\partial \hat{\varphi}}{\partial x}$. There is no feedback from u^* to φ^* during the adjoint integration, but at $t = 0$ the background error calculation couples the windfield changes back to the tracer as for the wind observation only case.

2.3. Wind and tracer observations

We are now back to the full Euler-Lagrange equations for the system. Collecting together the above results we see that the wind affects the tracer during the forward integration and background error calculation, whereas the tracer affects the wind during the adjoint calculation.

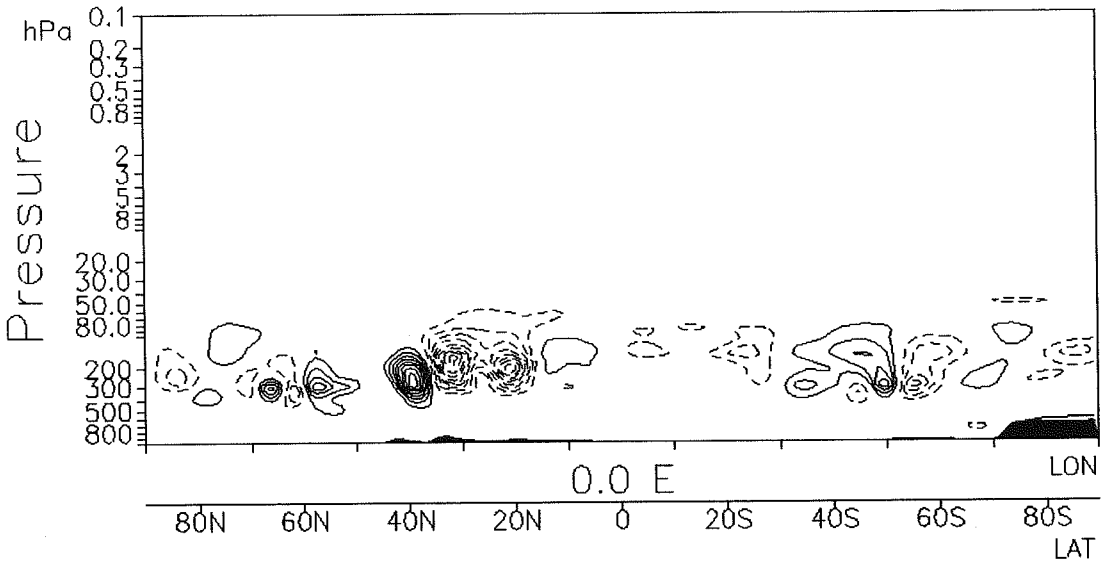


Figure 1. Ozone icrements, (analysis - first guess)/climatology, from multivariate background error coupling with vorticity. The cross section is along 0E on 2/12 1997, 18h (after seven days of data assimilation). The contour interval is 2%, with negative contours dashed (no zero).

2.4. Extension: Ozone in radiation

So far we have considered a general tracer. In the case of ozone, there is a direct feedback on the dynamics through the absorption of radiation in the stratosphere. This means that ozone can interact with the dynamics during a forward model integration by using the model ozone in the radiative transfer parameterization. Returning to the present simplified model, let us represent this by a source term $\kappa\varphi$ on the right hand side (rhs) of the wind equation, leading to an additional source term κu^* on the rhs of the adjoint ozone equation. Now any change of the adjoint wind will affect the ozone as well during the adjoint integration. Thus even in the case of a wind observation only, the adjoint ozone is nonzero. The inclusion of model ozone in the radiation parameterization will thus make the ozone and dynamics interactive during both the forward and adjoint integrations, while there is still only a one-way interaction from dynamics to ozone in the background error calculation.

3. DATA ASSIMILATION INCLUDING OZONE

After the introductory notes on the coupling between tracers and dynamics in 4D-VAR, we now consider the actual data assimilation in the IFS forecast model including a prognostic equation for ozone mass mixing ratio φ

$$\frac{d\varphi}{dt} = R_\varphi \tag{21}$$

where R_φ is a parameterization of sources and sinks of ozone from Cariolle and Déqué, 1986. All experiments have been performed with a 43 level model version with the top model level at 0.05hPa. The background and observation errors used for ozone will not be discussed here, since they are still under development and not essential for the conclusions.

We will present the results of four preliminary data assimilation experiments which were performed to test the implementation of ozone in the ECMWF data assimilation system:

1. No ozone observations, $R_\varphi = 0$.
2. TOVS total ozone observations, $R_\varphi = 0$.
3. TOVS total ozone observations, $R_\varphi \neq 0$.
4. TOVS layer ozone observations, $R_\varphi \neq 0$.

All experiments include a background error coupling between ozone and vorticity similar to the formulation discussed above, and the use of observations is normal except for ozone. All experiments are 4D-VAR, except for experiment 1 which is 3D-VAR, since with no ozone observations the interaction of dynamics and ozone only takes place through the background error coupling which is the same for 3D-VAR and 4D-VAR.

The difference between experiments 3 and 4 is the formulation of the TOVS ozone observation operator. From NOAA/NESDIS we obtain a retrieved product called total ozone. From the paper of Neuendorffer, 1996, we learn that this product consists of a retrieved layer ozone between 31hPa and 600hPa with a constant climatology added on top. A layer ozone product $d_{\varphi l}$ can be obtained from the total ozone product $d_{\varphi t}$ through the simple formula below (this is more accurate than Eq. 16 of his paper, according to a personal communication from Neuendorffer)

$$d_{\varphi l} = d_{\varphi t}(0.85 + 0.05\gamma) - 71.4\gamma - 18 \tag{22}$$

$$\gamma = 0.9 + 1.1\cos(latitude) \tag{23}$$

The associated observation operators are (apart from a horizontal interpolation to the observation loca-

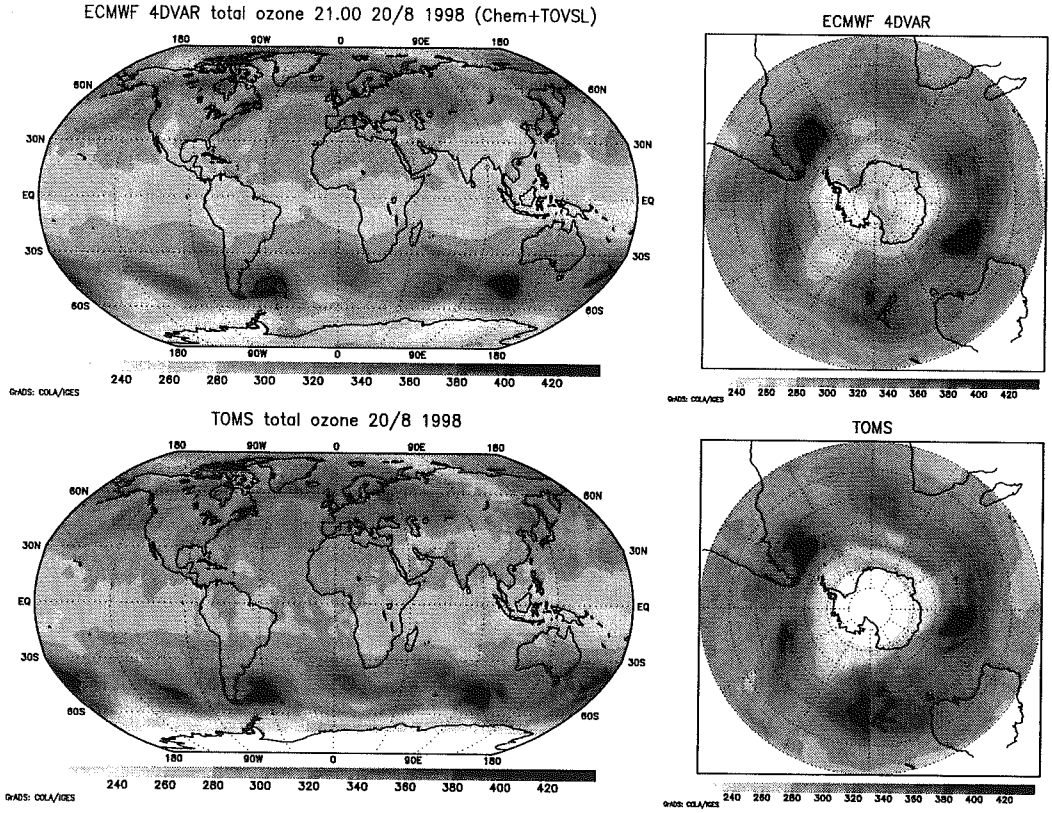


Figure 2. Total ozone after 24 hours of data assimilation compared with TOMS measurements (Dobson units). TOVS layer ozone observations are used and the ozone chemistry parameterization is turned on.

tion)

$$H_{\varphi t} = -\int_{p_s}^{0} \frac{\varphi(p)}{g} dp \qquad (24)$$

$$H_{\varphi l} = -\int_{\min(600hPa, p_s)}^{31hPa} \frac{\varphi(p)}{g} dp \qquad (25)$$

where p_s is the surface pressure and g is the gravitational acceleration.

The discrepancy in the total ozone product can explain a lot of the differences between TOMS and TOVS retrieved total ozone. When the southern hemispheric polar jet is active above 30hPa for example, the TOVS total ozone product will completely miss the associated synoptic ozone variations which can be clearly seen in TOMS. This is to be expected, since ozone above 30hPa is represented by a constant in the TOVS product.

3.1. The background error coupling

Experiment 1 was performed to see how the background error coupling from vorticity to ozone works. From the background error statistics of the model we expect the largest relative impact of vorticity on ozone in the lower stratosphere and upper troposphere between 150hPa and 300hPa, where the vorticity error variance explains ca. 30% of the ozone error variance. Furthermore there should be no accumulation of ozone due to the background error coupling over time.

Figure 1 shows a cross section of the ozone analysis increments normalized by climatology from Fortuin and Kelder, 1998. The background error coupling has largest relative effect above the tropopause as it should. The peak values are 20% of climatology, with typical values in the lower stratosphere of 3-5% of climatology. When we add up all the ozone increments from the seven days of the experiment (28 6-hour assimilation cycles) the result is unbiased and mostly below 5%, with peak value 30%, of climatology. Dividing these numbers by the 28 assimilation cycles we see that on average there is negligible accumulation of ozone due to the background error coupling.

3.2. The effect of ozone observations

In this section we will look at the effect of the TOVS observations on the ozone field in data assimilation. We will compare the effect of the total ozone and layer ozone observations and also look at how important the ozone chemistry parameterization R_φ is.

Figure 2 shows total ozone after 24 hours (four 6-hour cycles) of data assimilation using TOVS layer ozone observations and including the chemistry parameterization (experiment 4). The starting field was zonally symmetric climatological ozone from Fortuin and Kelder, and we see a 6-hour forecast from the fourth 4D-VAR output. The agreement with TOMS is good, although there are some differences associated with the southern polar jet. These differences come partly from the fact that the TOMS figure is a collection of measurements over 24 hours.

Figure 3 shows the bias and standard deviation between TOMS and the different experiment configurations 2 to 4 as a function of latitude. We see that

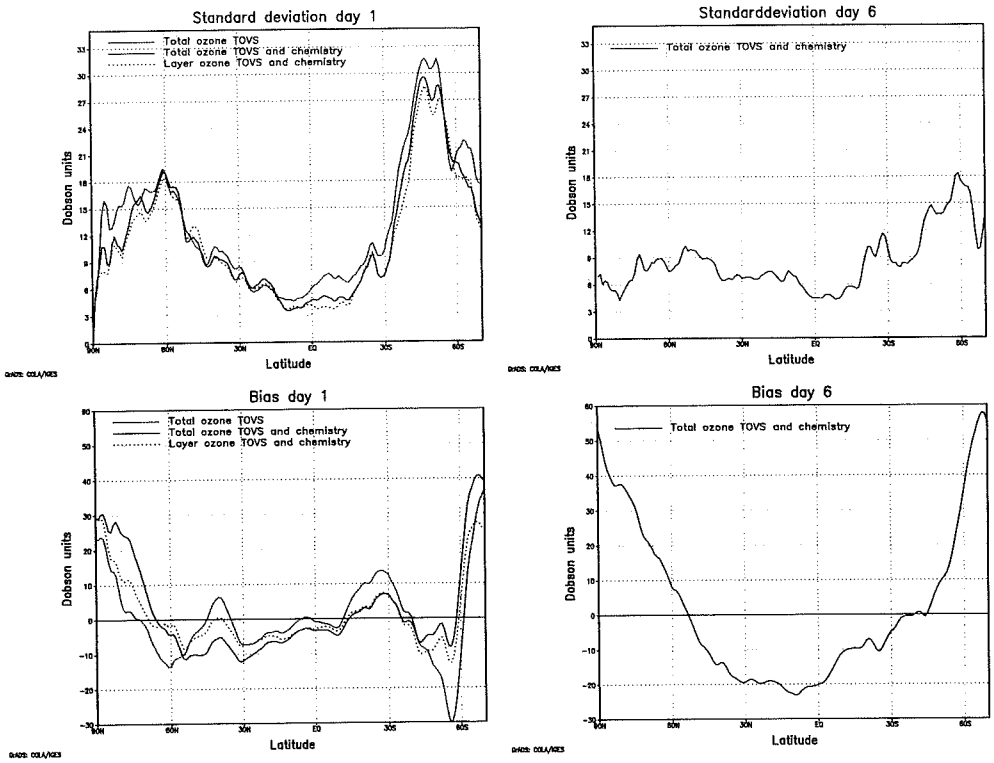


Figure 3. Standard deviation and bias of the ozone assimilation with respect to TOMS after one and six days of assimilation (20/8 and 25/8 1998).

the exclusion of the parameterization and the use of the total ozone TOVS both increase the error of the forecast.

We continued experiment 3 for six days. As found in earlier experiments with a pure ozone forecast, ozone decreases in the tropics and increases at the poles. We had hoped to get away from this bias by including total ozone observations in the data assimilation system, but this has evidently not worked. A possible reason for this is the construction of the ozone chemistry parameterization. In Figs. 4– 5 we see zonally averaged ozone assimilation increments with and without chemistry, experiments 2 and 3. There is a competition between data assimilation and ozone chemistry parameterization, with similar tendencies of opposite sign. This is evident in the much reduced amplitude of the analysis increments when chemistry is included. The parameterization is by construction drawing ozone towards a climatological equilibrium state, whereas the observations by nature draw ozone to the actual state of the day (see Jeuken et al., 1998, for possible modifications of the parameterization for getting at least the total ozone right).

4. CONCLUSIONS

From the preliminary experiments performed we find that a stratospheric ozone chemistry parameterization is essential for the assimilation of ozone data with low vertical resolution. We see however that the parameterization as used presently in the ECMWF model would benefit from improvements in accuracy. In particular, it is a disadvantage to formulate the parameterization as a tendency towards an equilibrium state which may be neither representative for the ECMWF model nor the actual situation of the day. With height-resolving ozone observations, the importance of good chemistry may diminish with respect to an accurate ozone prediction. If the goal is to obtain dynamical feedback from ozone, then it is essential to include good enough chemistry parameterization. Otherwise the resulting dynamical increments will reflect ozone modelling errors instead of transport modelling errors. This is not a problem in the lower stratosphere, where the chemistry is very slow, but this will reduce the usefulness of ozone observations in the middle and upper stratosphere.

We only used TOVS retrieved ozone observations here. We saw that the original total ozone product includes a constant for the upper stratosphere and above. TOVS is an accurate lower stratosphere/upper troposphere ozone measurement, and by using a layer ozone derived by subtracting the artificial constant, we obtain a better observation. We plan to test the use of ozone in a radiative transfer model for assimilating TOVS radiances directly. The ozone radiances as well as SBUV/2 profiles and GOME total ozone and profiles will be the next step in improving the multivariate 4D-VAR ozone assimilation at ECMWF.

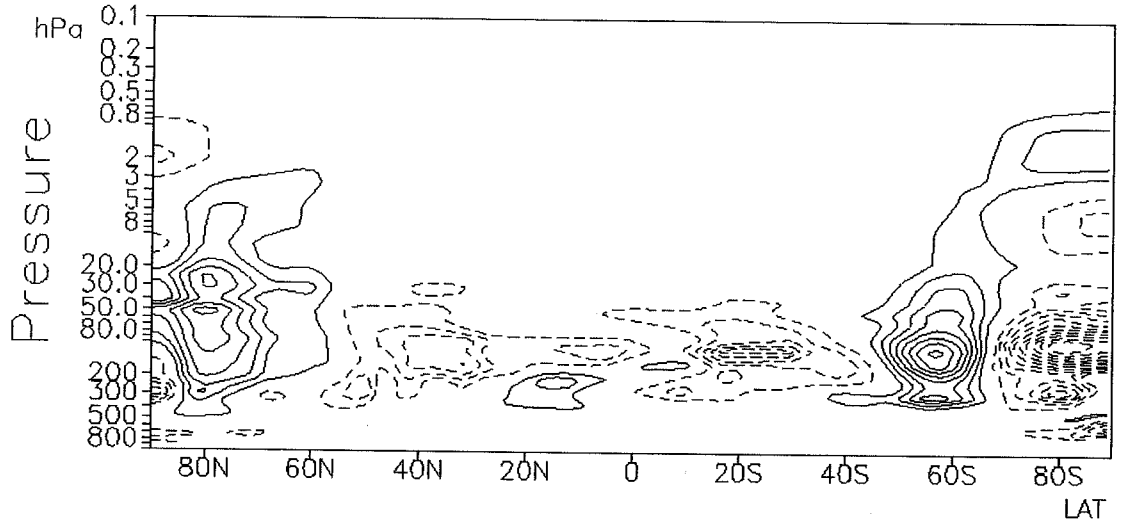


Figure 4. Ozone increments, (analysis - first guess)/climatology, using retrieved TOVS total ozone without chemistry. The zonal average is for 20/8 1998, 15h (after one day of data assimilation). The contour interval is 1%, with negative contours dashed (no zero).

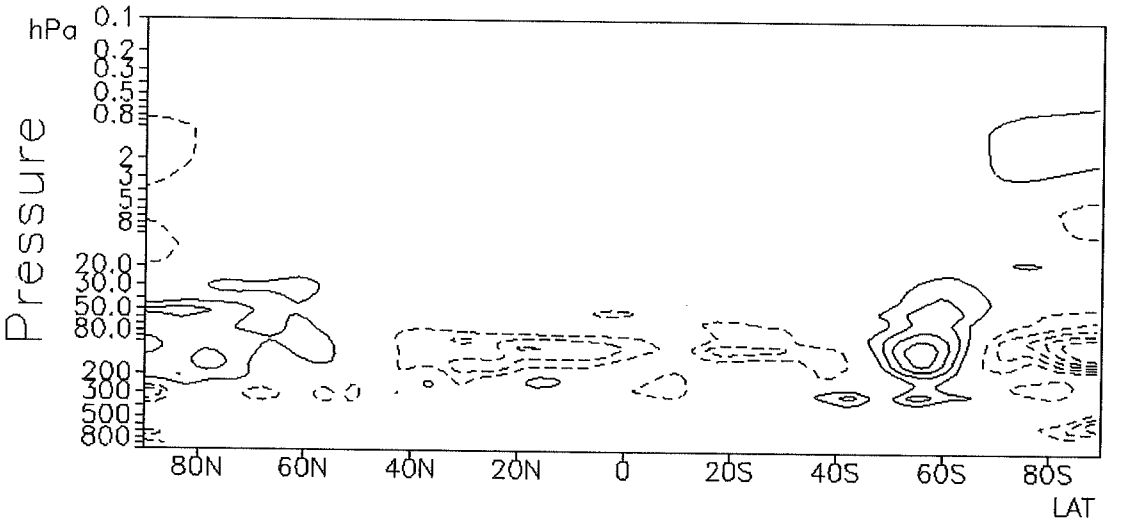


Figure 5. Ozone increments using retrieved TOVS total ozone and ozone chemistry parameterization (otherwise as Fig. 4).

REFERENCES

A. F. Bennet. *Inverse methods in physical oceanography.* Cambridge University Press, 346 pp., 1992.

D. Cariolle, M. Déqué. Southern hemisphere medium-scale waves and total ozone disturbances in a spectral general circulation model. *J. Geophys. Res.*, 91:10825–10846, 1986.

J. Derber, F. Bouttier. A reformulation of the background error covariance in the ECMWF global data assimilation system. *Accepted for publication in Tellus.*, 1998.

A. B. M. Jeuken, H. J. Eskes, E. V. Hólm, P. J. F. van Velthoven, H. M. Kelder. Assimilation of total ozone satellite measurements in a three-dimensional tracer transport model. *Accepted for publication in J. Geophys. Res.*, 1998.

J. P. F. Fortuin, H. M. Kelder. An ozone climatology based on ozonesonde and satellite measurements. *Accepted for publication in J. Geophys. Res.*, 1998.

A. C. Neuendorffer. Ozone monitoring with TIROS-N operational vertical sounders. *J. Geophys. Res.*, 101:18807–18828, 1996.

A. Untch, A. Simmons, M. Hortal, C. Jacob, and colleagues. Increased stratospheric resolution in the ECMWF forecasting system. *In this proceedings*, 1998.

DATA ASSIMILATION AT THE CENTRE FOR GLOBAL ATMOSPHERIC MODELLING: PROGRESS AND PLANS

Alan O'Neill and William Lahoz

NERC Centre for Global Atmospheric Modelling,
Department of Meteorology, University of Reading, RG6 6BB, UK

1. BACKGROUND

As part of the international effort to exploit data from NASA's Upper Atmosphere Research Satellite (UARS), the UK Met Office (UKMO) adapted its operational data assimilation system to include the stratosphere as well as the troposphere. Data from nadir-sounding operational satellites, as well as ground-based instruments, were assimilated to provide global fields of meteorological variables from the ground up to 0.3 hPa. These fields were provided to the UARS Science Team as "correlative data" to facilitate the interpretation of UARS data (Swinbank and O'Neill 1994). An archive has been established of these fields from October 1991 (shortly after the launch of UARS) to the present day. UARS data were not included in this assimilation, however. In parallel with this work, some experiments were run to assimilate UARS temperature, ozone and wind data. These assimilations dealt with data only for short periods, and there remains considerable scope to exploit UARS data in a data assimilation system.

The Centre for Global Atmospheric Modelling at the University of Reading has undertaken a project to assimilate UARS temperature and ozone data by adapting the troposphere-stratosphere data assimilation sytem used to provide correlative data for UARS. These data will be assimilated together with all available operational data, including data from nadir-sounding operational satellites. As well as providing valuable data sets for climate research, experience gained from combining nadir- and limb-sounding instruments will be valuable for the next stage of the project, which will be to assimilate temperature and trace chemical data from instruments on board the Envisat, scheduled for launch in May 2000.

2. PROJECT OUTLINE

The data assimilation system will be, in the first instance, the so-called Analysis Correction (AC) scheme developed by Lorenc et al. (1991). The AC scheme is, in effect, a nudging scheme in that the state-vector of the forecast model is adjusted iteratively down the gradient of a cost function, but there is only one iteration step per time step of the model. The AC scheme gives full weight to an observation at the time it was made, which is an advantage for treating asynoptic satellite data.

Work has begun to assimilate temperature measurements from the Microwave Limb Sounder (MLS) on UARS. The advantage of this instrument is that it provided data for several years, from the time of launch until mid 1998. Because of recent changes in operational procedures at UKMO, it may not prove possible to assimilate the complete data record from MLS. We shall therefore begin by assimilating data for the year 1997, focusing on periods of meteorological interest, such as stratospheric sudden warmings in the northern hemisphere and the break up of the polar vortex in the stratosphere of the southern hemisphere. At first, the background error covariances will be adopted that were used to produce the correlative data for UARS. An early goal will be to optimise these error covariances. The assimilation will also use information on instrument errors provided by the MLS team. In the light of the assimilation experiments, this information will be improved.

Work to assimilate ozone data from MLS is at a preliminary stage as of January 1999. Ozone data have been extracted from the British Atmospheric Data Centre, and periods of interest have been identified during 1997. Research is underway, in collaboration with UKMO and other parties, on how best to represent the background errors to assimilate ozone data.

UKMO is developing a 3dvar scheme to replace the current AC scheme in its operational data assimilation system (Andrew Lorenc, personal communication). After 3dvar becomes operational, it is our intention to repeat our UARS assimilation experiments with this improved data assimilation system.

3. PROJECT PLANS

Beyond the assimilation of UARS data, our goal is to have an effective data assimilation system in place to exploit data from instruments on Envisat. The instruments of prime interest for us are:

MIPAS (Michelson Interferometer for Passive Atmospheric Sounding), a limb sounder which measures

temperature, ozone and other chemical species with a vertical resolution of about 3 km;

GOMOS (Global Ozone Monitoring by Occultation of Stars), which measures temperature, ozone and other chemical species with a vertical resolution of about 1.7 km; and

SCIAMACHY (Scanning Imaging Absorption Spectrometer for Atmospheric Cartography), a limb- and nadir-sounder which measures vertical profiles and total column amounts of temperature, ozone and other chemical species.

The European Union has funded a concerted action, DARE (Data Assimilation in Readiness for Envisat), to co-ordinate collaboration between European groups to build a capability to assimilate Envisat data. Further and other parties, to assimilate Envisat data has been accepted by the European Space Agency. CGAM data assimilation plans involve close collaboration with its DARE partners.

REFERENCES

Lorenc, A. C., R. S. Bell, and B. MacPherson, 1991: The meteorological office analysis correction data assimilation scheme. *Q. J. R. Meteorol. Soc., 117,* 59-89.

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

Chemical data assimilation

ATMOSPHERIC CHEMICAL DATA ASSIMILATION

D.J. Lary[1,2], S. Hall[2], M. Fisher[3]

[1]Department of Geophysics & Planetary Sciences, Tel Aviv University, 69978 Tel Aviv, Israel
[2]Department of Chemistry, Lensfield Road, Cambridge, CB2 1EW, England

ABSTRACT

Since the 1960s, one of the most important sources of data on our atmosphere has been from satellites. The ability of satellites to observe the Earth ranges from nearly complete global coverage every day, to continuous observations of a particular part of the Earth. Satellites offer the only practical method of obtaining valuable data over much of the world, especially the oceans and remote land areas. To date, the analysis of chemical trace species has received little attention in comparison with the analysis of meteorological variables. Current methods of chemical analysis used for satellite data tend to treat species independently, ignoring the complex and very important balances that exist between species. Moreover, the large diurnal variations in the concentrations of many species are either accounted for in very simple ways, or avoided by analysing concentrations at fixed local time. Thus a great deal of valuable information contained in the shape of diurnal cycles and the partitioning of chemical species is completely wasted. This report outlines the methods we use, and the progress we have made in chemical data assimilation of a large number of atmospheric species relevant to ozone depletion.

Key words: *Chemical data assimilation, 4D-Var, ozone depletion, atmospheric chemistry.*

1. Introduction

The global distribution of ozone and other atmospheric constituents has changed considerably over the last decade, with record low global ozone levels in the last few years and downward trends over much of the globe (WMO 1995). The analysis of the changes in the ozone fields, and of atmospheric chemistry in general, is severely hampered by a lack of consistent data sets and especially from the lack of insight in to the role of chemistry and dynamics on the behaviour of ozone. The data assimilation package that we have developed, and is described here should be able to dramatically improve this state of affairs.

Since the 1960s, one of the most important sources of data on our atmosphere has been from satellites. The ability of satellites to observe the Earth ranges from nearly complete global coverage every day, to continuous observations of a particular part of the Earth. Satellites offer the only practical method of obtaining valuable data over much of the world, especially the oceans and remote land areas. The Earth Observing System (EOS), which will be launched in the late 1990s, will provide the most complete set of Earth observations ever taken.

The intelligent use of this data on a wide variety of chemical constituents, measurements that have cost many millions of dollars/pounds to make, is a non-trivial task as the observations are not co-located in time or space. Satellites make measurements of atmospheric constituents by a range of methods, and at a range of times and locations. The measurements are not made on a regular spatial grid or at the same times of day. Since the analysis of satellite measurement is so complex, the measurements have not been used to their full potential.

To date, the analysis of chemical trace species has received little attention in comparison with the analysis of meteorological variables. Current methods of chemical analysis used for satellite data tend to treat species independently, ignoring the complex and very important balances that exist between species. Moreover, the large diurnal variations in the concentrations of many species are either accounted for in very simple ways, or avoided by analysing concentrations at fixed local time. Thus a great deal of valuable information contained in the shape of diurnal cycles is completely wasted.

The technique of data assimilation can significantly improve the situation. Data assimilation is most highly developed in meteorology, particularly in weather forecasting. Improvements in assimilation techniques have already been instrumental in considerably improving the accuracy and range of weather forecasts, and the potential for assimilation techniques to advance our general understanding of the Earth system has already been clearly demonstrated.

The data assimilation technique is so effective since it seeks to produce an analysis which fits a set of observations taken over a time window, subject to the strong constraint that the evolution of the analysed quantities is governed by a deterministic model describing the given observations. By imposing the equations of the model as strong constraints, the

analysis problem is reduced to that of determining initial values for the model such that the subsequent evolution minimises a measure of the fit to the observations. The analysis method is therefore able to elegantly exploit information, which is not available to conventional analysis techniques, such as the shape of the diurnal cycles of atmospheric constituents. Thus intelligently using our knowledge of the processes involved has allowed us to extract much more information from the observations.

As a result, asynoptic satellite observations made at whatever location or time within a time window can be used to produce a set of self-consistent synoptic analyses of the observed species. In addition, synoptic analyses can be inferred for species included within the model even though not actually observed. In other words, an intelligent use of our knowledge of the processes involved extends the information contained within the observations to infer observations not made. This can itself be used as a stringent test of the technique.

For the first time in *Fisher & Lary* [1995] we have applied the technique of data assimilation to atmospheric chemistry. This use of the data assimilation technique for atmospheric chemistry produced a synoptic analysis of chemical species from asynoptic satellite data. In addition, the method allowed many useful insights to be gained that can not be obtained by any of the other techniques currently available. We have now considerably extended the work of *Fisher & Lary* [1995] to have a more detailed chemical scheme.

2. Objectives

The specific objective of this research has been to derive the maximum value from remotely sensed atmospheric data gathered by international agencies in understanding the effects of man-made pollutants on our atmosphere, such as the chemical mechanisms responsible for ozone depletion. This is the first time that a self-consistent chemical analysis of many atmospheric observations has been produced using data-assimilation will have been used to examine in unprecedented detail the chemical mechanisms involved in such important environmental issues as ozone depletion.

We now have a good basis for the tools that we need to make a substantial contribution to the analysis of EOS, ERS, and Envisat data, as well as to the analysis of existing data from UARS.

2.1. Analysis Method

4D-Var expresses the analysis problem as the constrained minimisation of a cost functional, $\mathcal{J}$, defined as

$$\mathcal{J} = \frac{1}{2}(\mathbf{x}_b - \mathbf{x}_0)^T \mathbf{B}^{-1}(\mathbf{x}_b - \mathbf{x}_0) + \frac{1}{2}\sum_{n=0}^{N}(\mathbf{y}_n - \mathbf{s}_n)^T \mathbf{R}_n^{-1}(\mathbf{y}_n - \mathbf{s}_n) \quad (1)$$

Here, $\mathbf{x}_0$ is the vector of initial parcel concentrations, $\mathbf{x}_b$ is an independent estimate of the initial parcel concentrations and $\mathbf{B}$ is the covariance matrix of expected errors in $\mathbf{x}_b$. The expression $(\mathbf{x}_b - \mathbf{x}_0)^T \mathbf{B}^{-1}(\mathbf{x}_b - \mathbf{x}_0)$ is generally called the "background term" of the cost functional and $\mathbf{x}_b$ is called the background.

The vector $\mathbf{y}_n$ in equation (1) consists of all observations which are considered valid at timestep n. $\mathbf{s}_n$ is a vector of "model equivalents" of the observations. That is, each element of $\mathbf{s}_n$ is an estimate of the corresponding element of $\mathbf{y}_n$, based, in our application of the method, on the parcel concentrations at timestep n. In the analyses presented in this paper, $\mathbf{s}_n$ is a linear function of parcel concentrations, $\mathbf{s}_n = \mathbf{H}_n \mathbf{x}_n$, where $\mathbf{x}_n$ is the vector of concentrations of all species for all parcels at timestep n, and $\mathbf{H}_n$ is an "observation operator". We leave further discussion of the calculation of $\mathbf{s}_n$ until later sections.

The matrix $\mathbf{R}_n$ is the covariance matrix for the random errors in $(\mathbf{y}_n - \mathbf{s}_n)$ which would be expected given a perfect analysis. That is, $\mathbf{R}_n$ accounts for the random errors in the observations and the "representativeness errors" (Lorenc, 1986) introduced in simulating the observations.

The strong constraints of the model equations are incorporated into the analysis by regarding $\mathcal{J}$ as a function of the initial concentrations only — i.e. as a function of $\mathbf{x}_0$. Concentrations at subsequent times are determined by integrating the model equations forward in time. This achieves two major simplifications. First, it replaces a constrained minimisation problem with an unconstrained problem. (Numerical algorithms for unconstrained minimisation are considerably more efficient and less prone to problems of ill-conditioning than are algorithms for constrained minimisation.) Second, the number of independent variables is reduced by a factor of $N + 1$.

The analysis scheme uses a descent algorithm to produce a convergent sequence of estimates of the vector $\mathbf{x}_0$ which minimises the cost functional. The algorithm requires the calculation of the gradient of the cost functional with respect to $\mathbf{x}_0$. This is evaluated by integrating the adjoint of the tangent linear equations for the model. These equations may be derived in a number of ways. Talagrand & Courtier (1987) derived the equations using the theory of adjoint operators. A derivation in terms of Lagrange multipliers is also possible (Daley, 1991).

The algorithm used to minimise $\mathcal{J}$ is as follows:

1. Start with an initial guess for $\mathbf{x}_0$.
2. Integrate the photochemical model to give $\mathbf{x}_n$ for $n = 1...N$.
3. Evaluate $\mathcal{J}$. If the value of $\mathcal{J}$ is small enough then STOP.
4. Iterate the ATL equations to calculate $\nabla_{\mathbf{x}_0}\mathcal{J}$.
5. Use a descent algorithm to find a better guess at $\mathbf{x}_0$ (i.e. a guess for which $\mathcal{J}$ is reduced).
6. GOTO 2

3. Chemical Model

The model used in this the model by Lary called AUTOCHEM (Lary *et al.* (1994)). The model is explicit and uses the adaptive-timestep, error monitoring, Bulirsch-Stoer (1980) time integration scheme designed by Press et al. (1992) for stiff systems of equations. The integration scheme is as accurate as the often used Gear (1971) package, but faster. Photolysis rates are calculated using full spherical geometry and multiple scattering as described by Lary & Pyle (1991a,b) after Meier et al. (1982) and Anderson (1983). For this study, photolysis rates were updated every fifteen minutes.

An extremely useful feature of the model is the existence of a code generation program which automates the process of writing numerical chemical models. Given a set of reaction databases for bimolecular, trimolecular, photolysis and heterogeneous reactions the program automatically writes the `fortran` code to calculate the time derivatives and the jacobian matrix required by the numerical integration program. This enables new reaction schemes, containing different chemical species, rates or both, to be implemented without the need for manual coding. The user specifies which reactants and products should be included. Reactions which are just upper limit estimates, with unknown products of the required species, or reactions which are endothermic by more than a given amount can be automatically excluded if required.

The analysis method presented in this paper integrates the adjoint of the tangent linear equations for the model in addition to integrating the model. The number of code changes required to implement new reactions in the analysis scheme is about double the number required to implement the reactions in the model alone. By modifying the code generation program to automatically write code for the adjoint tangent linear model, we have extended to the analysis scheme the ability to implement new reactions without manual coding. We intend to exploit this ability in future studies to determine the effect on the accuracy of the analysis of including or excluding various species and reactions.

4. Atmospheric Database

The ever-increasing quantity and availability of observational data allows us to develop more accurate and more detailed atmospheric models than ever before. Unfortunately there is a drawback to having this amount of readily accessible information, namely the time that must necessarily be invested in sorting and extracting the required data. Given that different data sets are invariably presented in different file formats this process is usually performed from scratch each time, especially when data from more than one set is required.

A full description of the database and its powerful capabilities are beyond the scope of this report. However, the basic points are:

1. Common file format

⋆	Day	⋆	Month
	Year		Day of year
	Time (GMT)	⋆	Longitude
⋆	Latitude		Temperature
	Pressure	⋆	Geopotential height
⋆	Potential temperature		Potential vorticity
⋆	Equivalent latitude		Equivalent length
	Data		Error

Table 1. Chemical database records for each observation. Those quantities marked with a star are used as co-ordinates and may be used to specify limits in the search process.

2. Flexible database import & export

3. Rapid searching of large datasets

As a first step in automating this process we first convert each new data set into a common file format. This allows our searching procedures to be re-used each time.

The file format we use is a column-based plain text file. We use one file for each day of observations from each instrument. The data usually consist of the date, time and location of the measurement along with the measurement itself and an estimate of the error on that value. In addition to these we add several other quantities describing the atmosphere at that point and time. The full list is given in Table 1. Those quantities marked with a star are used as co-ordinates and may be used to specify limits in the search process.

New data set is added to the database using a form. The location of the files is set along with the instrument and quantity being measured. The database then scans each file to determine the co-ordinate limits and creates an entry in the index table. Finally the entry in the contents table is created.

Data is extracted from the database using a the multi-tabbed form. The first tab allows the user to select which specie or other quantity is required. Any one or more observing instruments measuring that quantity may then be selected. Finally any set of the available data sets may be selected. Once these have been chosen the user can then select the limits in any set of co-ordinates which are used to filter the data. The data may be filtered according to physical location or its location in equivalent latitude - theta space. In addition, a range of dates may be specified and data may be dropped if its error is greater than a user-specified level. The third tab allows the user to refine the list of fields output buy the database.

The data may be binned in up to three co-ordinates which the user can choose, with the widths of each bin specified separately. This allows a high number of data points to be brought down to a manageable level. There is also an option to drop extreme data values.

The final form allows the output options to be selected. The data may be output in a format suitable to be used with our models or for display with the IDL plotting package. Data is automatically grouped

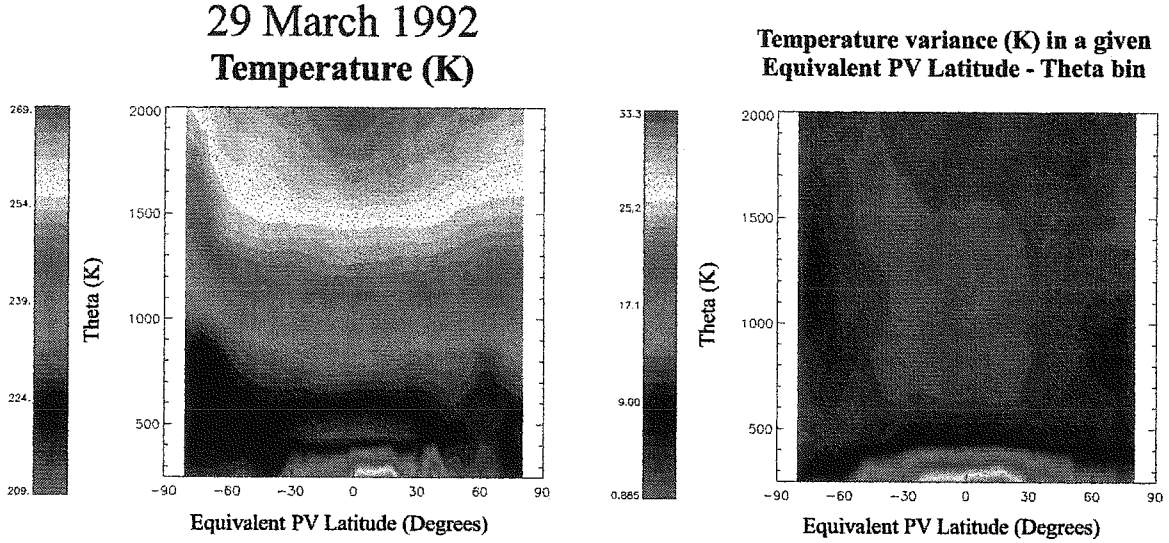


Figure 1. The equivalent PV latitude - theta vertical temperature cross-section from the UKMO analysis for 29 March 1992.

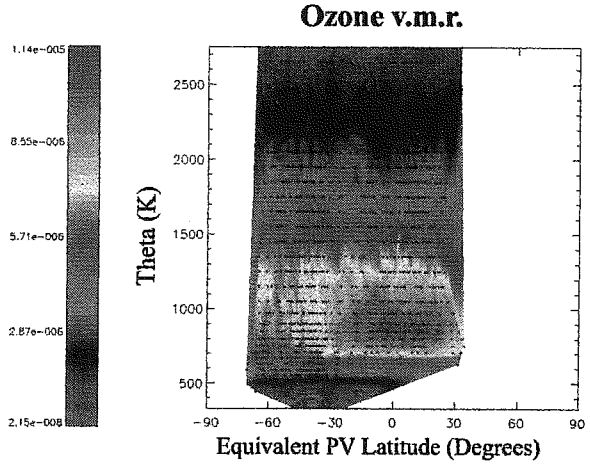


Figure 2. All the O_3 observations made during the ATLAS-1 mission by ATMOS plotted as a function of equivalent PV latitude and theta.

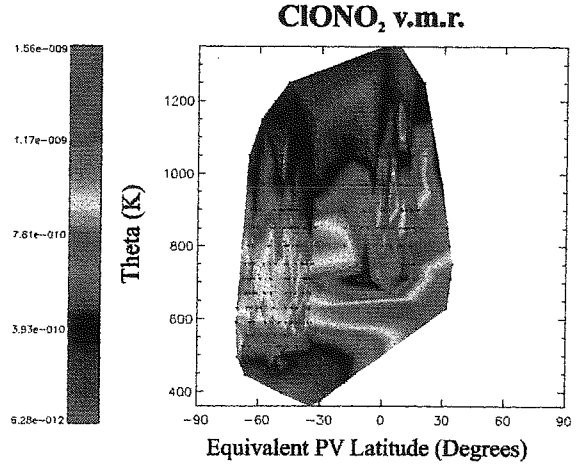


Figure 3. All the $ClONO_2$ observations made during the ATLAS-1 mission by ATMOS plotted as a function of equivalent PV latitude and theta.

according to the time each observation was made. The grouping interval can be specified here. There is also an option to ignore the dates of the observations and to treat them all as one day.

The versatility of this database allows much to be added with a minimum of effort.

5. Visualisation

Given the amount of data we have available to put into out models and the complexity of the chemical schemes we are able to run, the amount of data that is produced means that some method of rapidly visualising the results is required. To this end we have developed a package called PlotAssim using RSI's IDL software.

The primary mode of operation is to read in the output files generated by our models. Using a file selection window the path and file extension can be set, either directly or by clicking on browse and using a directory selection dialog box. The full list of files is given on the left, the list selected on the right. The files may be selected in any order.

The package has two main plotting modes, X-Y line plots and X-Y-Z surface plots. In each case a separate set of data may be loaded and overlayed. This allows model output to be compared with observations extracted from the database or with another model.

6. Catalytic cycles

Within the PlotAssim package there is a facility to calculate the properties of chemical catalytic cycles. The importance of atmospheric catalytic cycles was first recognised by *Bates and Nicolet* [1950]. Since then, it has become well established that the concentration of stratospheric ozone is controlled by the

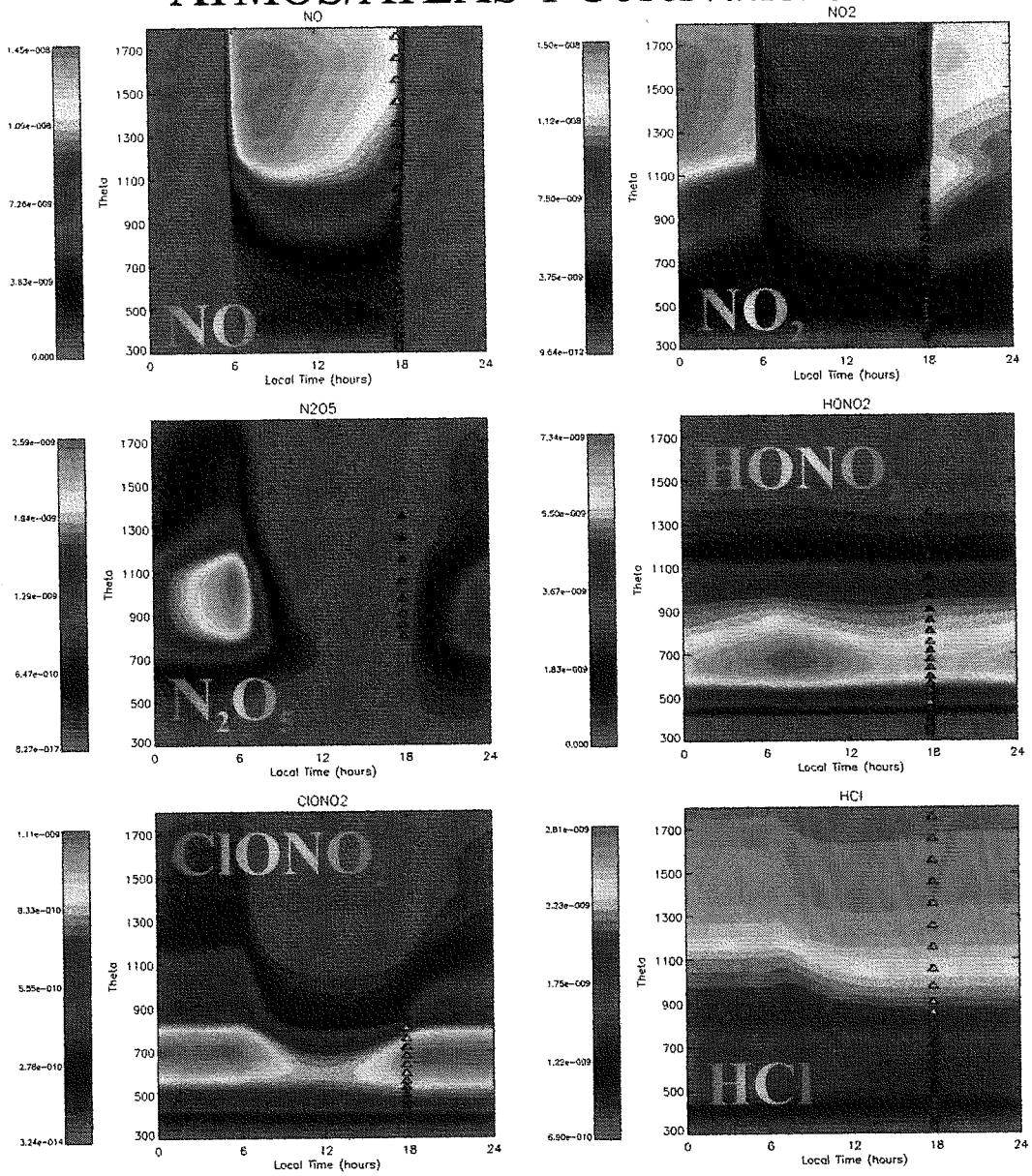


Figure 4. 4D-Var chemical analyses for 29 March 1992 of ATLAS-1/ATMOS data in an equivalent latitude band centered on 40°S. The vertical axis in each case is the potential temperature, θ, in K, the horizontal axis is the local solar time in hours. Overlaid on the analysis are the observations made by ATMOS. As the ATMOS instrument uses solar occultation the observations are only available at sunrise or sunset.

balance between its production, and its destruction, and that the destruction of ozone is mainly due to catalytic cycles involving nitrogen, hydrogen, chlorine, and bromine species.

The effectiveness of catalytic cycles in destroying ozone is controlled by two factors, the chain length of the catalytic cycles and the abundance of the radical which is the chain center. The chain length is the number of times the catalytic cyclic is executed before the reactive radical involved, the chain center, is destroyed. It is valuable to systematically consider the effectiveness of the ozone destruction cycles in the atmosphere.

Once a set of cycles has been entered several quantities are calculated for each cycle. Amongst these are the chain length and chain effectiveness for each cycle, the rate of propagation of each branch as a fraction of the total, the rate of termination of the cycle and the rate limiting step of each branch.

Some of this data is then summarised in a cycle plot. The two nodes involved in the cycle (e.g. Cl & ClO)

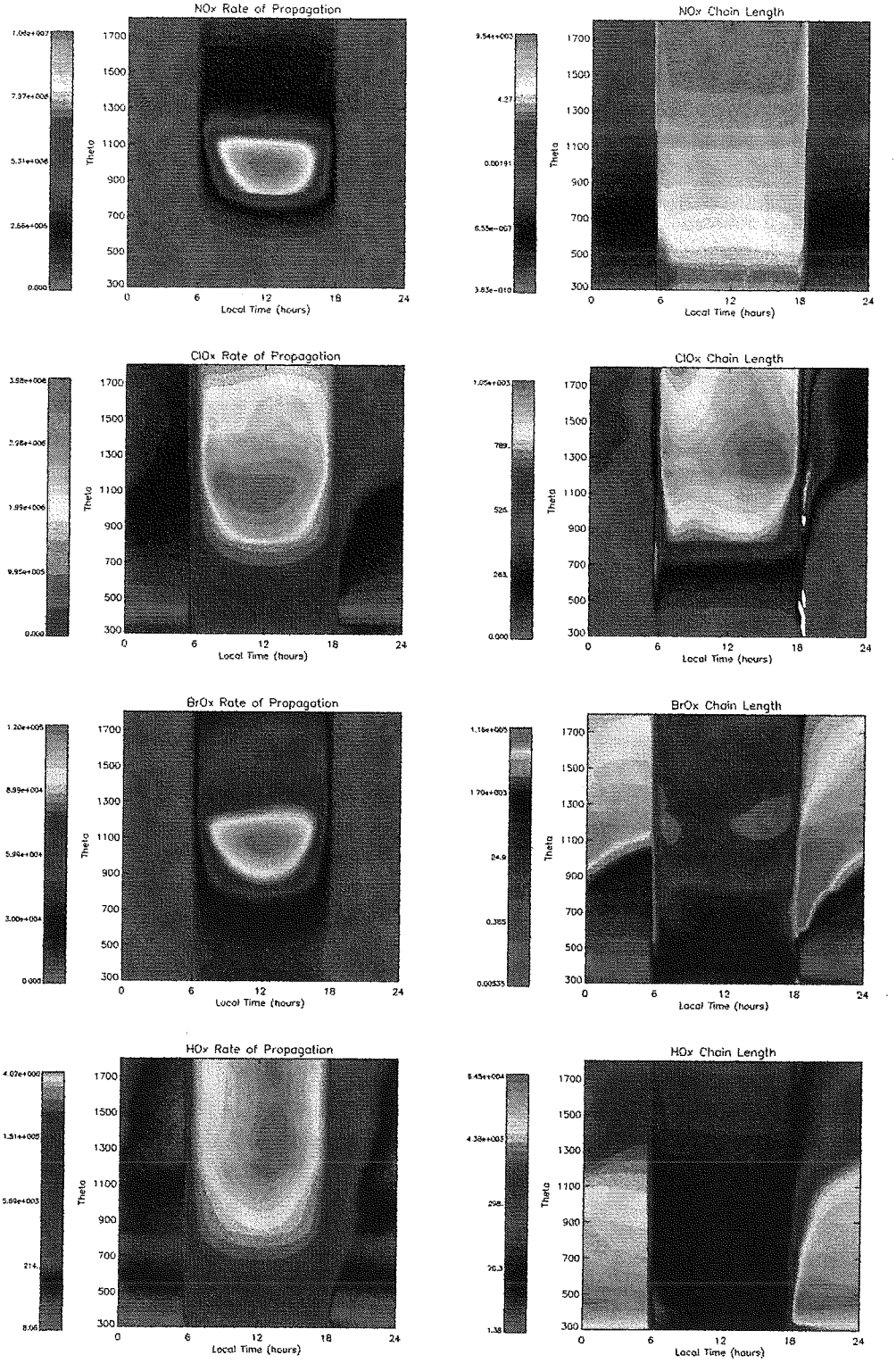


Figure 5. On the left hand side is shown the rate of propagation for 29 March 1992 of the NO_x, ClO_x, BrO_x, and HO_x cycles based on the 4D-Var chemical analysis of ATLAS-1/ATMOS data in an equivalent latitude band centered on 40°S. On the right hand side is shown the corresponding chain length (how many times the catalytic cycle is executed before the radical is removed) for the NO_x, ClO_x, BrO_x, and HO_x cycles.

are shown in the large circles in the centre of the diagrams. Each branch or group is shown as a box linking the nodes. Those going from left to right are shown at the top, those going in the opposite direction are shown at the bottom. The branch contributing most to the rate in each direction is plotted closest to the nodes. The fraction each branch contributes to the rate in each direction is shown in the boxes along the arrows. The reactions that make up each branch are indicated along with how often each reaction is the rate-limiting step (i.e. the slowest) for that branch and the mean rate of that reaction.

7. Example analysis of ATMOS data

The Atmospheric Trace Molecule Spectroscopy is an instrument that flies in the payload bay of the space shuttle and measures the concentrations of gases present in the atmosphere at altitudes between 10 and 150 kilometers. As the shuttle's orbit carries it into and out of Earth's shadow, the ATMOS instrument views the Sun as it sets or rises through the atmosphere. The spectrometer measures changes in the infrared component of sunlight as the Sun's rays pass through the atmosphere. Trace gases absorb very specific wavelengths which allows us to determine what gases are present, in what concentrations, and at what altitudes. ATMOS has flown several times.

The STS-45/ATLAS 1 mission was launched on March 24, 1992 from the Kennedy Space Center. During 8 days of operation, the ATMOS instrument made a total of 98 observations, spanning a substantial portion of the globe. The 53 measurements taken at orbital sunrise covered the mid-latitude and equitorial regions of the earth from 30°S to 30°N. The 41 sunset observations were made at 25°S to 55°S. ATMOS was only able to monitor the atmosphere down to a height of about 20 km, due to a recent eruption of Mount Pinatubo, which clouded the region below that with dust and aerosols. Figure 1 shows the equivalent PV latitude - theta vertical temperature cross-section from the UKMO analysis for 29 March 1992 and the associated variance. As just two examples of the observations made Figures 2 and 3 show all of the O_3 and $ClONO_2$ observations made during the ATLAS-1 mission by ATMOS plotted as a function of equivalent PV latitude and theta.

7.1. A detailed study of 40°S

Over this period the equivalent PV latitude for which the vertical profiles covered the largest range of altitudes, and for which the largest number of species were observed simultaneously was centered on about 40°S. For this equivalent PV latitude a detailed chemical analysis was performed on the basis of our chemical data assimilation. To perform this analysis a chemical scheme containing a total of 59 species was used. 54 species were integrated, namely: $O(^1D)$, $O(^3P)$, O_3, N, NO, NO_2, NO_3, N_2O_5, HONO, HNO_3, HO_2NO_2, CN, NCO, HCN, Cl, Cl_2, ClO, ClOO, OClO, Cl_2O_2, $ClNO_2$, $ClONO_2$, HCl, HOCl, CH_3OCl, Br, Br_2, BrO, $BrONO_2$, BrONO, HBr, HOBr, BrCl, H_2, H, OH, HO_2, H_2O_2, CH_3, CH_3O, CH_3O_2, CH_3OH, CH_3OOH, CH_3ONO_2, $CH_3O_2NO_2$, HCO, HCHO, CH_4, CH_3Br, CF_2Cl_2, CO, N_2O, CO_2, H_2O. The model contained a total of 357 reactions, 233 bimolecular reactions, 31 trimolecular reactions, 48 photolysis reactions, 41 heterogeneous reactions, and 4 unimolecular heterogeneous reactions.

Figure 4 shows some of the results from this 4D-Var chemical analyses for 29 March 1992 of ATLAS-1/ATMOS data. The vertical axis in each case is the potential temperature, θ, in K, the horizontal axis is the local solar time in hours. Overlaid on the analysis are the observations made by ATMOS. As the ATMOS instrument uses solar occultation the observations are only available at sunrise or sunset. It can be seen from Figure 4 that the 4D-Var was capable of simultaneously reproducing many of the the observations made by ATMOS.

On the left hand side of Figure 5 is shown the rate of propagation for 29 March 1992 of the NO_x, ClO_x, BrO_x, and HO_x cycles based on the 4D-Var chemical analysis of ATLAS-1/ATMOS data in an equivalent latitude band centered on 40°S. On the right hand side is shown the corresponding chain length (how many times the catalytic cycle is executed before the radical is removed) for the NO_x, ClO_x, BrO_x, and HO_x cycles.

Although the full chemical analysis of this data is beyond the scope of this article, it is very exciting to see that we are now able to extract such a level of detail from atmospheric observations.

8. Conclusion

Significant progress has been made in the development of chemical 4D-Var in terms of infra-structure and analysis tools.

Recognition

The significance of the research in this area has been recognised as pioneering by the Royal Meteorological Society, and the Royal Society who awarded Dr. Lary a Royal Society University Research Fellowship at Cambridge University, and most recently by the government of Israel in the award of the prestigious Alon Fellowship.

ACKNOWLEDGMENTS

David Lary would like to thank the European Union for its support, the Royal Society for its provision of a Royal Society University Research Fellowship and the Government of Israel for its recent provision of an Alon fellowship.

REFERENCES

Fisher, M., Lary, D. J., Lagrangian 4-dimensional Variational Data Assimilation of Chemical Species, *Q. J. R. Meteorol. Soc.*, *121*(527) Part A, 1681-1704, 1995.

World Meteorological Organisation, Scientific assessment of stratospheric ozone: 1988, WMO Global Ozone Research and Monitoring Project, Rep. *20*, 1990.

World Meteorological Organisation, Scientific assessment of stratospheric ozone: 1991, WMO Global Ozone Research and Monitoring Project, Rep. *25*, 1992.

FOUR–DIMENSIONAL DATA ASSIMILATION FOR TROPOSPHERIC OZONE ANALYSIS

H. Elbern, H. Schmidt, A. Ebel

Institute for Geophysics and Meteorology, University of Cologne, F.R.G.

ABSTRACT

The problem of analyzing the chemical state of the troposphere by observations and a chemistry–transport model (CTM) using advanced assimilation techniques is considered. The method applied is the four dimensional variational data assimilation (4D-var) which iteratively minimizes the misfit of modeled concentration levels with observations. In the present development phase the assimilation procedure is able to optimize initial values or emission rates of the CTM. It can be shown that the space-time variational approach is able to analyze and to predict observed species to a certain limit after the end of the assimilation window.

Key words: chemical data assimilation; adjoint modelling; parameter optimization; inverse modelling, ozone prediction, chemistry–transport modelling.

1. INTRODUCTION

The data assimilation procedure of exploiting observations scattered in time for the analysis of the chemical state of the atmosphere is considered for a complex mesoscale chemistry transport model. In an attempt to transfer the good success with the four–dimensional variational data assimilation technique recently made in meteorological forecasting, this method has been applied to tropospheric gas phase chemistry modelling. In this exposition initial values and emission rates are the parameters to be optimized. However, the underlying control theory admits for further parameters to be optimized. For example, emphasis may also be placed on lateral boundary values, and deposition velocities.

The objective of the present paper is to demonstrate the feasibility of the four–dimensional variational data assimilation technique (4D–var) for comprehensive gas phase chemistry transport models, and to give an account of first experiences with the analysis skill.

2. THEORY AND MODEL DESCRIPTION

2.1. Space–time variational data assimilation

Data assimilation procedures seek to find an initial model state which assures an optimal compliance between observations and an ensuing model integration. An objective measure to quantify the difference between measurements and model state is conveniently defined by a distance function as follows (Lorenc, 1986):

$$\mathcal{J}(\mathbf{x}(t)) = \frac{1}{2}(\mathbf{x}_b - \mathbf{x}(t_0))^T \mathbf{B}^{-1}(\mathbf{x}_b - \mathbf{x}(t_0)) + \frac{1}{2}\int_{t_0}^{t_N} (\hat{\mathbf{x}}(t) - \mathbf{x}(t))^T \mathbf{O}^{-1}(\hat{\mathbf{x}}(t) - \mathbf{x}(t))dt \quad (1)$$

where $\mathcal{J}$ is a scalar functional defined on the time interval $t_0 \leq t \leq t_N$ dependent on the vector valued state variable $\mathbf{x}$. The first guess or background values $\mathbf{x}_b$ are defined at $t = t_0$, and $\mathbf{B}$ is the covariance matrix of the estimated background error. The observations are denoted $\hat{\mathbf{x}}$ and the observation and representativeness errors are included in the covariance matrix $\mathbf{O}$.

We now introduce the model equation $\frac{\partial \mathbf{x}(t)}{\partial t} = \mathbf{M}(\mathbf{x}(t)) + \mathbf{e}$ with emission rates $\mathbf{e}$ as a strong constraint with Lagrange multipliers $\lambda(t)$ as an additional term for (1), and a scalar product $\langle \ , \ \rangle$, then defining a constrained distance function $\mathcal{L}(\mathbf{x}, \mathbf{e}, \lambda)$. Hence,

$$\mathcal{L}(\mathbf{x}, \mathbf{e}, \lambda) = \mathcal{J}(\mathbf{x}, \mathbf{e}) + \int_{t_0}^{t_N} \langle \lambda, \frac{\partial \mathbf{x}(t)}{\partial t} - \mathbf{M}(\mathbf{x}(t)) - \mathbf{e} \rangle dt \quad (2)$$

At the stationary point $\delta\mathcal{L} = 0$ the following holds:

$$\frac{\partial \mathcal{L}(\mathbf{x},\mathbf{e},\lambda)}{\partial \lambda} = 0, \ \frac{\partial \mathcal{L}(\mathbf{x},\mathbf{e},\lambda)}{\partial \mathbf{x}} = 0, \ \frac{\partial \mathcal{L}(\mathbf{x},\mathbf{e},\lambda)}{\partial \mathbf{e}} = 0.$$

At the stationary point of $\mathcal{L}$, after differentiation with respect to $\mathbf{x}$ and integration by parts, we find

$$-\frac{d\lambda(t)}{dt} - \mathbf{M}'^{*}\lambda(t) = \mathbf{R}^{-1}(\hat{\mathbf{x}}(t) - \mathbf{x}(t)). \quad (3)$$

The adjoint initial condition $\lambda(t_N) = 0$ is applied with $\mathbf{M}'^{*}$ being the adjoint model.

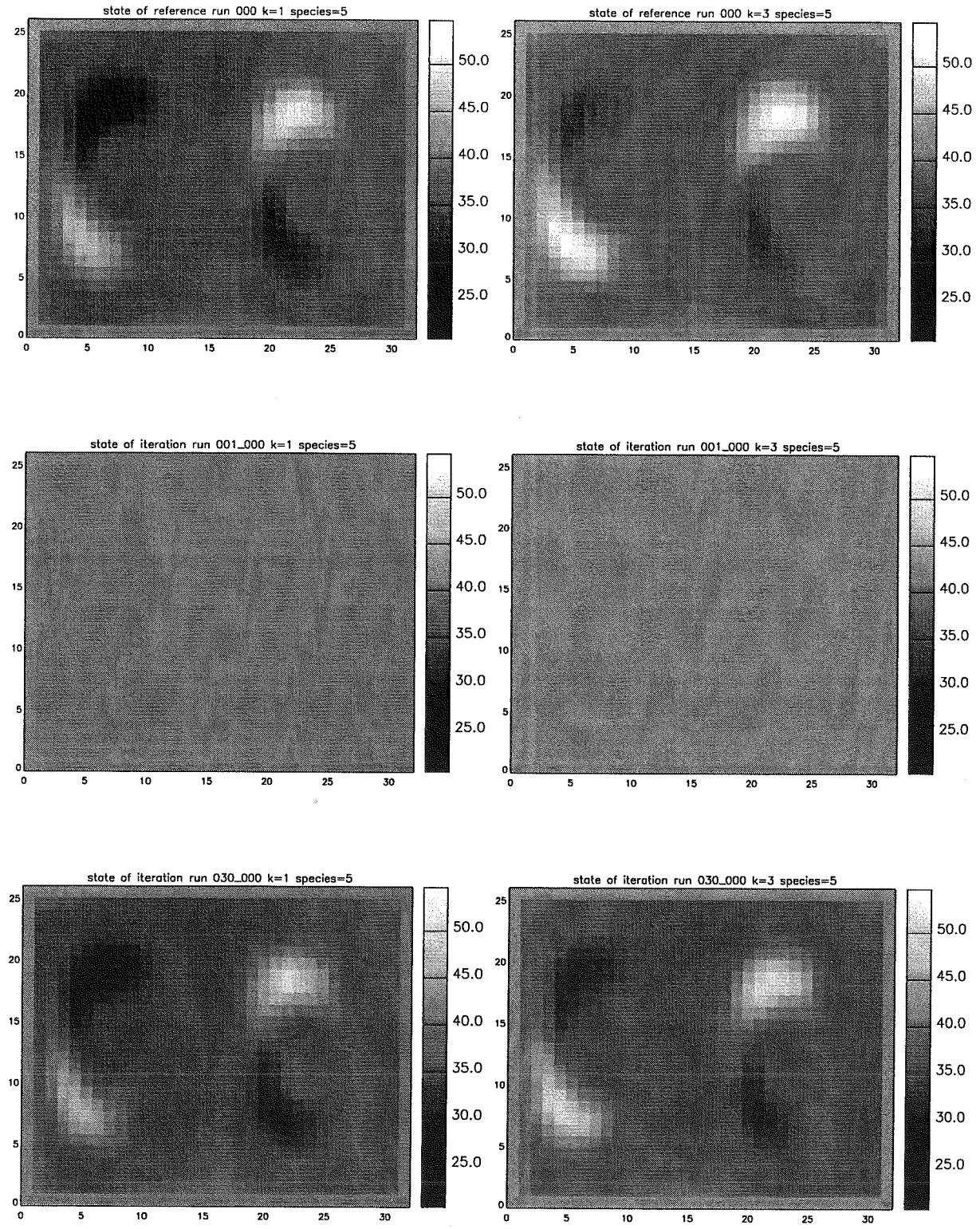


Figure 1. Identical twin experiment displaying assimilation results of surface ozone observations for the lowest level (left panel column) and third level (right column), scaled to be 230 m. Left upper panel: initial reference state ("truth") of the surface ozone and ensuing simulated concentration levels as provided as observations. Right upper panel: reference state ("truth") of the third level, scaled to be 230 m, but depicted ozone observations are not provided. Medium panels: first guesses for the lowest (left) and third level (right). Lower panels: assimilation results for the lowest (left) and third level (right).

Similarly, after derivation of $\mathcal{L}$ with respect to $\mathbf{e}$ we obtain as stationary condition

$$\mathbf{e}_b - \mathbf{e} = \int_{t_0}^{t_N} \lambda(t)dt, \qquad (4)$$

with the left hand side also giving the gradient of $\mathcal{J}$ with respect to e. For a more detailed description see for example Talagrand and Courtier (1987), Navon (1997), Elbern et al. (1997, 1998).

2.2. The Limited Area Chemistry Transport Model

The EURAD CTM2 is a comprehensive tropospheric Eulerian model operating on the mesoscale–α. A full description of the EURAD (European Air pollution Dispersion model) CTM2, which is an offspring of the Regional Acid Deposition Model RADM2 (Chang et al., 1987) may be found in Hass (1991). The chemistry transport model calculates the transport, diffusion, and gas phase transformation of about 60 chemical species with 158 reactions. For the gas phase chemistry a semi–implicit and quasi steady state approximation method (QSSA) is applied for the numerical solution of the stiff ordinary differential equation system as proposed by Hesstvedt et al. (1978). Horizontal and vertical transport is simulated by a fourth order Bott (1989) scheme.

3. IDENTICAL TWIN EXPERIMENTS

As a first test the capability of the 4D–var method is taxed within the framework of identical twin experiments. With this method artificial 'observations' are produced by a preceding reference model integration based on initial values or emission rates to be analyzed by the subsequent assimilation procedure.

3.1. Analysis of the Initial Concentrations

The skill of the 4D–var method can then be estimated by comparing the analysis with initial values withheld from the reference run. The assimilation time interval spans six hours, starting at 06:00 local time in the centre of the integration domain at midsummer conditions. The length of the assimilation interval is limited by available computing time. The wind field is defined to form two vortices, a cyclonic one in the eastern, and an anticyclonic one in the western part of the integration domain, with logarithmically increasing wind speed aloft.

In the experiment two circular emission areas are introduced, in the north–western and the south–eastern part of the integration domain. Only ozone observations of the lowest level are provided to the algorithm. The analysis result for ozone at the lowest model level is depicted in Fig. 1, left panel column. Information transfer from the observed surface layer to elevated height levels are only due to the vertical transport and diffusion operators. As a consequence of underdetermination, adaption to the set of observations is not only possible by modification of unobserved chemical species but also by mixing and transport downward of the observed tracer. From the viewpoint of minimization this can be reformulated as follows: although the minimum of the cost function may be unique in a reasonably bounded domain of the phase space, an elongated extremum reflecting the poor preconditioning of the minimization problem may severely hamper the approximation to a sufficient vicinity of the true state. The right panel column of Fig. 1 displays the analysis result for ozone at height level three, representing an elevation of about 230 m.

3.2. Emission Rate Optimization

For this experiment model configurations and meteorological conditions are the same as in the case of initial value optimization. The key difference is that the initial values are taken to be known correctly, but the emission rates are deviating from the "truth" and are to be analyzed. Fig. 2 displays an experiment where a time dependent emission scenario for NO is given for several grid points, representing for example a street network with common emission rates. The temporal variation within the time frame of the assimilation window of six hours is assumed to be composed by a constant background and a cosine. The first guess emission rates are depicted by the dashed line, while the "truth" is given by the bold line. Analysis results from the various grid points are displayed by the dotted lines. Clearly a significant approximation to the true state is obvious.

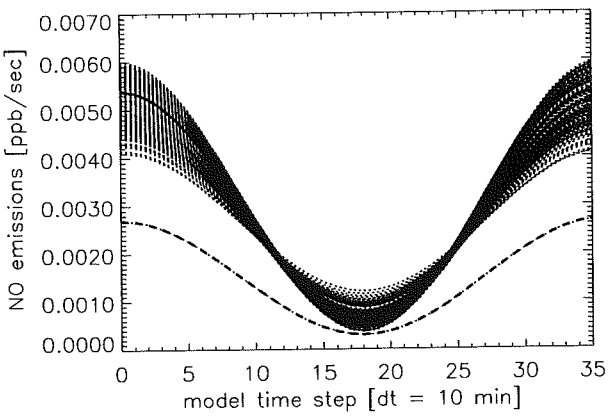


Figure 2. Temporal variation of the NO emissions during the 6 hours assimilation interval: solid line is reference time series, dash–dotted line is first guess. Dotted curves represent the analysis for each grid cell of the line sources. Units in [ppb/s].

4. REAL CASE STUDY

In this section a first application of the 4–dimensional variational technique to a real case study of an ozone episode during August 1997 is presented. The model grid configuration is given by a 77 × 67 × 15 grid structure encompassing nearly the entire European continent.

Only a limited number of about 70 measurement stations, mostly confined to central Europe, is available.

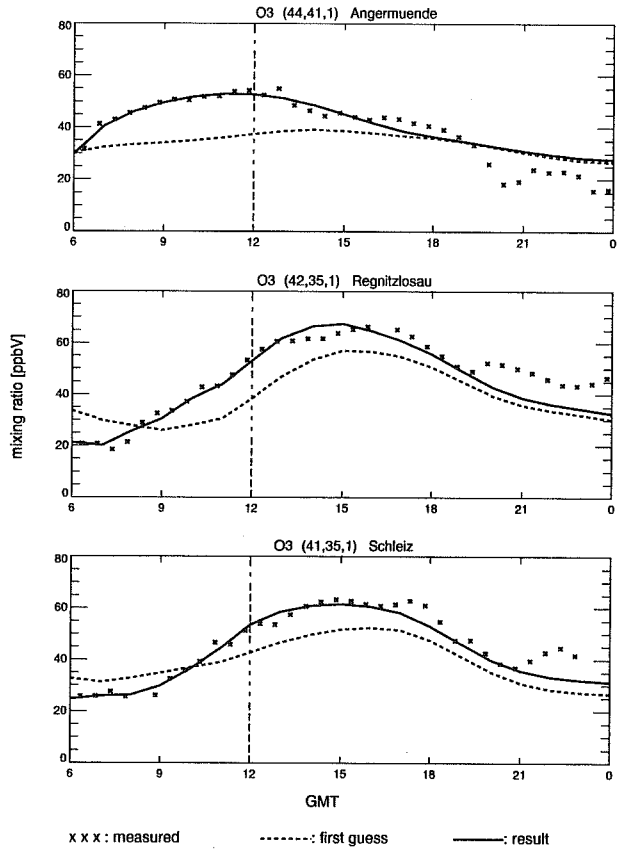


Figure 3. Initial value optimization: Time series for ozone concentrations for the measurement sites Angermünde (top), Regnitzlosau (middle), and Schleiz (bottom) at 5.8.97 . Observations are indicated by dots. Model result with first initial values indicated by dotted lines, results after initial value optimization indicated by bold lines. The assimilation window is the 6 hours interval 6 - 12 GMT, display of later hours for performance check only.

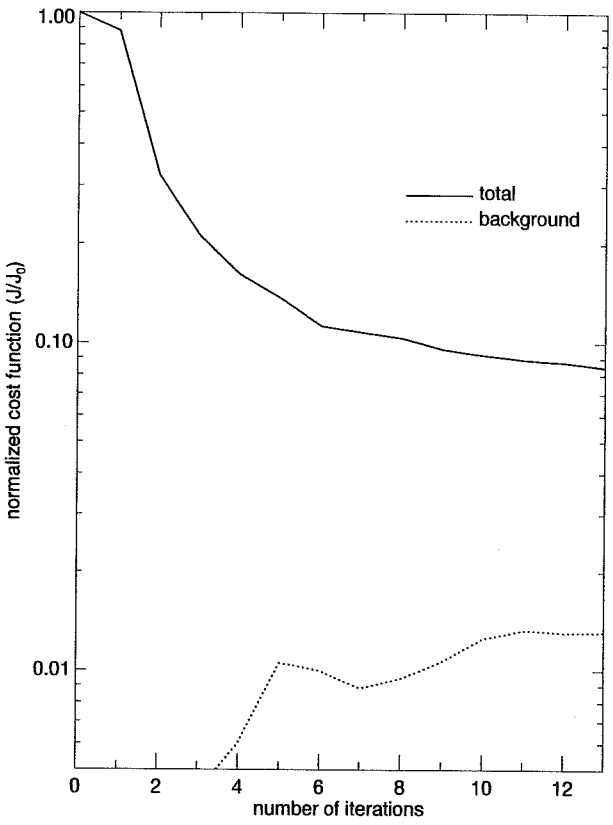


Figure 4. Evolution of the normalized cost function (1) as a function of iterations. The background fraction is indicated by dotted line.

4.1. Analysis of the Initial Concentrations

In Fig. 3 measurements of three surface observation stations are presented, along with the first guess based model simulation and the final analysis results. A significant performance improvement can be claimed not only during the assimilation interval from 6 to 12 GMT, but also for later hours. The favourable performance degrades however, presumably due to the missing radius of influence to be associated with each observation. Clearly concentration levels of air masses windward to the observation locations cannot be modified by the analysis and exhibit their unaltered ozone level when reaching the observed grid point. This effect can be progressively observed with integration time beyond the assimilation interval. Fig. 4 shows the decrease of the total cost function with respect to the number of iterations. The increasing background portion as given by the dotted line indicates the movement of the initial state's position from the background to the analysis.

4.2. Emission Rate Optimization

At some locations it was obvious that the emission inventory of the adjacent area was obviously erroneous. Nineteen species are simulated to be emitted, mostly nitrogen oxides and hydrocarbons. The analysis procedure is rerun with the option of emission optimization activated. Fig. 5 shows observations, first guess model performance and analysis performance again for ozone for three measurement sites. The assimilation window was defined to be 14 hours to span the principal part of the diurnal emission cycle. Again a significant improvement can be claimed after modification of the emission rates.

5. CONCLUSIONS

It is shown that the 4D–var data assimilation procedure is able to markedly improve the model performance. This can be expected to be further the case, when a proper definition of radius of influence is implemented for each observation. For the future both optimization procedures, initial value and emission rate optimization must be combined to avoid spurious artificial modifications caused by over-adaption. A sound statistical basis must be established to ensure an expedient weighting for both components.

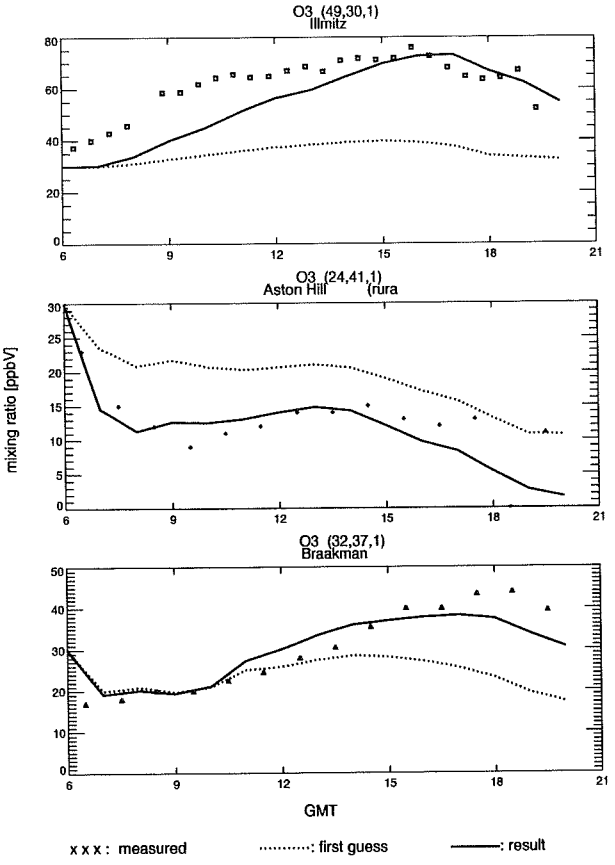


Figure 5. Emission rate optimization: Time series for ozone concentrations for the measurement sites Illmitz (top), Aston Hill (middle), and Braakman (bottom) at 5.8.97 . Observations are indicated by dots. Model result with first guess emissions indicated by dotted lines, results after emission optimization indicated by bold lines. Duration of assimilation window is 14 hours.

ACKNOWLEDGMENTS

We are grateful to those organizations who provided data records of observations: AIRPARIF Paris, IRM–KMI Belgium, RIVM The Netherlands, UBA Germany, UBA Austria, BUWAL Switzerland, NERI Denmark, NAQIA United Kingdom. Computational support was provided by the Central Institute for Applied Mathematics (ZAM) of the Research Centre Jülich, Germany.

REFERENCES

Bott, A. 1989, A positive definite advection scheme obtained by nonlinear renormalization of advective fluxes. *Mon. Wea. Rev.*, **117**, 1006–1015.

Chang, J. S., R. A. Brost, I. S. A. Isaksen, S. Madronich, P. Middleton, W. R. Stockwell, and C. J. Walcek 1987, A three-dimensional Eulerian acid deposition model: physical concepts and formulation, *J. Geophys. Res.*, **92**, 14,618–14,700.

Elbern H., H. Schmidt, and A. Ebel 1997, Variational data assimilation for tropospheric chemistry modeling. *J. Geophys. R.*, **102**, D13, 15,967–15,985.

Elbern H., H. Schmidt, O. Talagrand, and A. Ebel 1998, 4D–variational data assimilation with an adjoint air quality model for emission analysis, *Procedures of the APMS meeting*, Champs–sur–Marne.

Hass, H. 1991, Description of the EURAD Chemistry–Transport–Model Version 2 (CTM2), *Mitteilungen aus dem Institut für Geophysik und Meteorologie der Universität zu Köln*, A. Ebel, F.M. Neubauer, P. Speth, eds. **Nr. 83**, 100 pp.

Hesstvedt, E., Hov, Ø., Isaksen, I.S.A. 1978, Quasi–steady state approximation in air pollution modeling: Comparison of two numerical schemes for oxidant prediction. *Int. J. Chem. Kinet.*, **10**, 971–994.

Lorenc, A. C. 1986, Analysis methods for numerical weather predictions, *Q. J. R. Met. Soc.*, **112**, 1177–1194.

Navon, I. M. 1997, Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography. Dynamics of Atmospheres and Oceans, **27**, 55–79.

Talagrand, O., and P. Courtier 1987, Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory, *Q. J. R. Meteorol. Soc.*, **113**, 1311-1328.

A COMPUTATIONALLY EFFICIENT EXTENDED KALMAN FILTER FOR CHEMISTRY-TRANSPORT MODELS

X.-F. Zhang, S. M. Polavarapu, P. A. Makar

Atmospheric Environment Service, 4905 Dufferin Street, Downsview, Ontario, Canada, M3H 5T4

ABSTRACT

The Reduced-Rank Square Root Filter (RRSQRT-EKF) of Verlaan and Heemink (1995) is implemented for a nonlinear chemistry model in order to evaluate the potential efficiency of this algorithm when applied to a full chemistry-transport model. Two experiments using simulated observations are performed: one with erroneous emission rates and a second with model error. In both cases, the RRSQRT-EKF was able to provide a gain in efficiency with no loss in accuracy over the conventional Kalman filter. Also, when data generated by a complex mechanism was assimilated by a model with a simplified mechanism, the observed species were well recovered. Information was obtained about some (but not all) unobserved species.

Key words: data assimilation; Kalman filter.

1. INTRODUCTION

Recently, Fisher and Lary (1995) have demonstrated that data assimilation may be used to improve the estimation of chemical species. While this study concerned stratospheric ozone estimation, Elbern et al. (1997) have demonstrated that Four-Dimensional Variational Assimilation (4DVAR) can be beneficial to the problem of tropospheric ozone estimation. In this work, we investigate the applicability of a computationally efficient Kalman filter to the problem of tropospheric ozone estimation.

The tremendous cost of the Kalman filter is due to the necessity of integrating the forecast error covariance matrix in time. For a matrix of dimension $N \times N$, this requires $2N$ model integrations. Since N is typically 10^5 or 10^6 for meteorological applications, a Kalman filter is not feasible. While there are many methods of approximating the Kalman filter to reduce its cost, we investigate the efficiency of the Reduced-Rank Square Root Kalman Filter (RRSQRT-KF) proposed by Verlaan and Heemink (1995) for applications involving chemistry-transport models. The RRSQRT-KF has been shown to be as much as several orders of magnitude cheaper than the conventional algorithm for the transport problem alone. To evaluate its possible efficiency for the full chemistry-transport problem, we first consider its applicability to the chemistry problem alone. Thus, our model will be a nonlinear chemistry box model, at a single point in space.

2. THE CHEMISTRY MODEL

The chemistry model used in this study is a very simplified model that represents some of the main processes involved in tropospheric ozone formation. The reaction set, indicated in Table 1, involves only 22 of hundreds of possible reactions. The most important simplifications made are the omission of nighttime chemistry and the omission of all the complexity of the hydrocarbon interactions. The latter processes are represented here by the presence of a single, non-methane hydrocarbon, denoted HYD.

Table 1. A Simple Tropospheric Reaction Set

No.	Reaction	
1	NO_2	$\rightarrow NO + O(^3P)$
2	$O(^3P) + O_2 + M$	$\rightarrow O_3 + M$
3	$O(^3P) + NO_2$	$\rightarrow NO + O_2$
4	$NO + O_3$	$\rightarrow NO_2 + O_2$
5	$NO + NO + O_2$	$\rightarrow 2\ NO_2$
6	O_3	$\rightarrow O(^3P) + O_2$
7	O_3	$\rightarrow O(^1D) + O_2$
8	$O(^1D) + H_2O$	$\rightarrow 2\ OH$
9	$O(^1D) + M$	$\rightarrow O(^3P) + M$
10	$O_3 + OH$	$\rightarrow HO_2 + O_2$
11	$NO + HO_2$	$\rightarrow NO_2 + OH$
12	$O_3 + HO_2$	$\rightarrow OH + 2\ O_2$
13	$HO_2 + HO_2$	$\rightarrow H_2O_2 + O_2$
14	$HO_2 + HO_2 + M$	$\rightarrow H_2O_2 + M + O_2$
15	$HO_2 + HO_2 + H_2O$	$\rightarrow H_2O_2 + H_2O + O_2$
16	$HO_2 + HO_2 + H_2O$	$\rightarrow H_2O_2 + H_2O + O_2$
17	H_2O_2	$\rightarrow 2\ OH$
18	$H_2O_2 + OH$	$\rightarrow HO_2 + H_2O$
19	$CO + OH$	$\rightarrow HO_2$
20	$CH_4 + OH$	$\rightarrow RO_2 + H_2O$
21	$HYD + OH$	$\rightarrow RO_2$
22	$RO_2 + NO$	$\rightarrow NO_2 + HO_2$

The use of such a simplified model was motivated by two reasons. Firstly, we can attempt to address the implications of using a simplified chemistry model for data assimilation applications when the "true" system is known to be much more complex. This is important because four-dimensional data assimilation schemes are very expensive so that it may be necessary on computational grounds to use simplified chemistry schemes. Secondly, a simplified system that is representative of the nonlinearity of more realistic models will allow our results to be more easily interpreted.

Chemical integrations of the system of eight coupled differential equations resulting from Table 1 were performed using a second-order Euler backward scheme. A fixed chemistry time step of two minutes was employed and emissions of NO were introduced at each model time step. The reaction rates were held fixed over 1 hour time steps. No deposition was included.

3. THE DATA ASSIMILATION SCHEMES

Because the model is nonlinear, we use a first-order Extended Kalman filter (or EKF). Additionally, we will perform some experiments where the data is generated by running a more complex and complete model of tropospheric ozone. Thus the simple model will have considerable model error which will be correlated in time. To avoid making an assumption of white (in time) model error, we first augment the model state by the model error vector. Then the model error is itself modelled as a Markov process:

$$\begin{aligned} \boldsymbol{x}^f_{k+1} &= \Phi_k \boldsymbol{x}^a_k + \boldsymbol{w}_k + \boldsymbol{u}_k \\ \boldsymbol{w}_{k+1} &= \alpha \boldsymbol{w}_k + \eta. \end{aligned}$$

Here $\boldsymbol{x}$ is the state variable with dimension n, Φ is the nonlinear chemistry model and $\boldsymbol{w}$ is the model error vector with dimension q. $\boldsymbol{u}$ is an n-vector of emissions. Both n and q are 8. η is a white noise and α is taken to be 0.2. By defining a new state variable, $\boldsymbol{X}^a_k = [\boldsymbol{x}^a_k, \boldsymbol{w}_k]^{\mathrm{T}}$ of dimension $n+q=16$, a new model equation which is forced by white model error is found.

Assuming that the measurement error covariance matrix R (of dimension m) is a diagonal matrix, the update procedure of the RRSQRT algorithm for time step $k+1$ is summarised as follows:

$$\begin{aligned} \boldsymbol{x}^f_{k+1} &= \Phi_k \boldsymbol{x}^a_k + \boldsymbol{u}_k \\ L^f_{k+1} &= [\Phi_k \widetilde{L}^a_k, Q^{1/2}] \\ U_{k+1} D_{k+1} U^{\mathrm{T}}_{k+1} &= (L^f_{k+1})^{\mathrm{T}} L^f_{k+1} \\ \widetilde{L}^f_{k+1} &= [L^f_{k+1} U_{k+1}]_{n\times s} \end{aligned} \quad (1)$$

where Q is the model error covariance matrix and $\widetilde{L}^a_k$ is a $n \times s$ square root matrix of P^a_k. The columns of U_{k+1} contain the eigenvectors of $L_{k+1}L^{\mathrm{T}}_{k+1}$, and D_{k+1} is a diagonal matrix with the eigenvalues of $L_{k+1}L^{\mathrm{T}}_{k+1}$ as its elements. The analysis equations are

$$\begin{aligned} \psi_k &= (\widetilde{L}^f_k)^{\mathrm{T}} H^{\mathrm{T}}_k \\ \gamma_k &= (\psi^{\mathrm{T}}_k \psi_k + R)^{-1} \\ K_k &= \widetilde{L}^f_k \psi_k \gamma_k \\ \widetilde{L}^a_k &= \widetilde{L}^f_k - K_k \psi^{\mathrm{T}}_k [1 + (\gamma_k R)^{1/2}]^{-1} \\ \boldsymbol{x}^a_k &= \boldsymbol{x}^f_k + K_k [\boldsymbol{z}_k - H_k \boldsymbol{x}^f_k] \end{aligned} \quad (2)$$

Here $\boldsymbol{z}$ is the m-vector of observations and H is a linear interpolation operator mapping state variables to observed variables. It should be noted that the implementation of equation (2) needs to be performed m times for m uncorrelated measurements at each time step. The RRSQRT-EKF is started from the initial conditions: $\boldsymbol{x}^a_0 = \boldsymbol{x}_0$ and $P^f_0 = Q$.

In order to evaluate the efficiency and accuracy of the RRSQRT-EKF results, we also implement a 4DVAR scheme since it has been used by both Fisher and Lary (1995) and Elbern et al. (1997). The cost function employed is as follows:

$$\begin{aligned} J &= \sum_{k=0}^{N} [\boldsymbol{z}_k - \mathrm{H}(\boldsymbol{x}_k)]^{\mathrm{T}} R_k{}^{-1} [\boldsymbol{z}_k - \mathrm{H}(\boldsymbol{x}_k)] \\ &+ [\boldsymbol{x}_0 - \boldsymbol{x}^f_0]^{\mathrm{T}} (P^f_0)^{-1} [\boldsymbol{x}_0 - \boldsymbol{x}^f_0]. \end{aligned} \quad (3)$$

The interpolation operator H is both linear and time invariant in this work. At the start of the assimilation period, $P^f_0 = Q$ and $\boldsymbol{x}^f_0 = \boldsymbol{x}_0$.

4. THE EXPERIMENTS

Two experiments are performed. In the first, the simple model is used to generate the "truth" and an emission of NO is included. However, the data assimilation schemes will assume an incorrect emission rate for NO. This experiment addresses the fact that emissions are often a dominant factor in tropospheric pollutant estimation, and they are not well known.

In the second experiment, a more complete mechanism of tropospheric ozone processes is run to generate the "truth". The simple 8-species model is then used to assimilate the observations of 3 of the 8 species. This experiment attempts to mimic the situation where a simplified chemistry model is used for data assimilation while the true system is much more complex.

4.1. Experiment 1: Incorrect Emissions

In this experiment, the simple model (based on Table 1) will be run to generate a "truth". Emissions of NO of 0.6 ppbv/hour are assumed during this simulation. However, the RRSQRT-EKF and 4DVAR assimilation schemes will be run with an incorrect emission rate of 1.2 ppbv/hr (or an error of 100%). The initial conditions for the data assimilation are indicated in the first column of Table 2. The initial conditions used for the truth appear in the second column. The background state has values of similar orders of magnitude to the true initial condition for all species with the exception of HO_2 and RO_2. The observation error covariance matrix, R, includes errors of representativeness and is taken to be diagonal

with elements given by the squares of the third column of Table 2. Finally, the model error covariance matrix, Q, is also assumed diagonal with standard deviations given by the last column of Table 2.

Table 2. The initial conditions and error statistics for Experiment 1. The units are ppbv.

Species	Background	Truth	Obs error	Model error
NO_2	1.0	0.5	10	30
NO	0.2	0.3	10	30
O_3	50.0	80.0	30	90
CO	220.0	200.0	100	3
H_2O_2	1.0	2.0	1	3
HYD	1.0	0.5	1	3
HO_2	1.0E-02	1.0E-19	1	3
RO_2	1.0E-03	1.0E-19	1	3

The results of the assimilations are shown in Fig. 1. The rank of the RRSQRT-EKF solution shown is 8 and the 4DVAR solution used an assimilation period of 3 hours. In this figure, we see that without data assimilation, the background solution of NO_2 produces unrealistically large estimates due to the use of an incorrect emission rate. However, with data assimilation, both the RRSQRT-EKF and 4DVAR solutions correct for the error in the emission rate and produce the true solution for NO_2, NO and O_3. For the case of H_2O_2, the true solution is approached after about 24 hours. For HO_2 and RO_2 (not shown), both assimilation schemes produce good estimates but both have some difficulty reproducing the peak at hour 6. The 4DVAR solution is closer to the peak but noisy. The RRSQRT-EKF solution underestimates the peak. Finally, the rate of decay of CO is well obtained by both assimilation schemes but there is no adjustment for the initially incorrect value of CO. This is because in this simple system with no CO emissions, the CO evolves according to:

$$[\mathrm{CO}] = [\mathrm{CO}]_0 \exp\left(-\mathrm{k}\int_0^t [\mathrm{OH}]\mathrm{dt}\right). \qquad (4)$$

The OH solution is approximately described by the HO_2 evolution and from Fig. 1 we see that it is close to the truth. We can also note that the destruction rate of CO is close to the true destruction rate (left middle panel). However, observations of NO, NO_2 and O_3 do not help to determine the initial value of CO. Thus, the simple 8-species model is completely determined only with additional information: at least one good measurement of CO. The solution for HYD (not shown) behaves similarly to (4) and thus the same reasoning applies. Some information about HYD is also needed to completely determine the system.

Overall, from Fig. 1 we see that with data assimilation, the true solution is well approximated (with the exceptions noted above). Since an incorrect emission rate is the same as having an incorrect model forcing (or model error), it is interesting to note that the 4DVAR solution is as good as the RRSQRT-EKF solution although the 4DVAR algorithm used does not

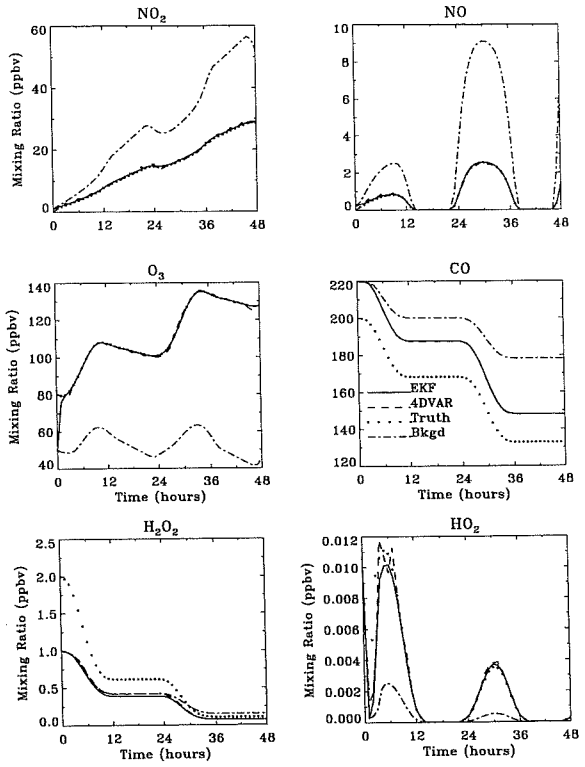


Figure 1. Experiment 1. Results of using three measurements (O_3,NO,NO_2), and an NO emission of unknown magnitude. There is no model error.

account for the existence of model error. The reason that the 4DVAR scheme performed well in this experiment is that a 3 hour assimilation period was employed. The growth of model error can be controlled by reducing the assimilation period. A discontinuity in the solution is then obtained after every assimilation period, but with a 3 hour assimilation length, the discontinuity is barely noticeable in Fig. 1.

In Fig. 2 we note the impact of reducing the rank of the square root of the forecast error covariance matrix on the solutions. Only the species for which the greatest impact was felt are shown. The rank 8 and rank 5 solutions are indistinguishable. However, reducing the rank from 5 to 4 results in noticeable changes to the estimation of the NO_2 rate, and the peaks in NO, HO_2 and RO_2. Therefore, even in this simple system, it is possible to gain some efficiency by reducing the the rank from 8 to 5 with little effect on the accuracy of the solution.

4.2. Experiment 2: Model error

In this experiment, the truth is obtained by running a much more complete model of the processes involved in tropospheric ozone formation. This model is the gas-phase reaction mechanism of ADOM (Venkatram and Karamchandani, 1988), solved using Gear's method (Gear, 1971), and includes 114 reactions and 47 species so that the simple model (Table 1) is not expected to provide a good approximation to it.

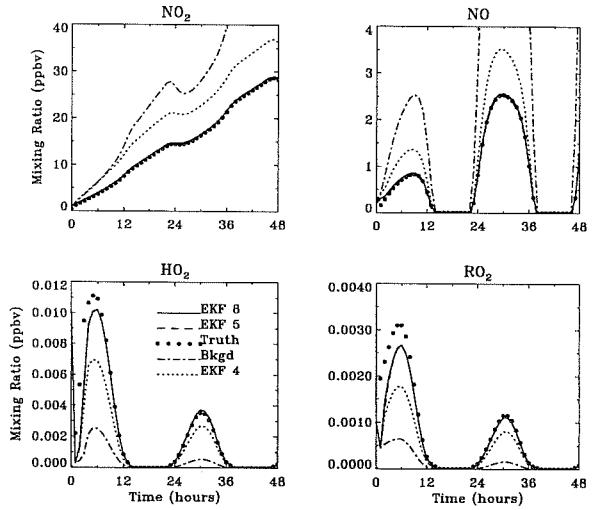


Figure 2. Experiment 1. Comparing the RRSQRT-EKF solutions with reduced ranks of 8, 5 and 4. The truth and background are also shown.

The ADOM run was initialised with the hydrocarbon mixing ratio of 0.5 ppbv distributed equally between the ADOM mechanism's 7 anthropogenic unoxygenated hydrocarbons; ethane, propane, higher alkanes, ethene, higher alkenes, toluene and higher aromatics. The truth in this case becomes a sum of hydrocarbon mixing ratios, to be compared to the single hydrocarbon of the simple model. For the latter, the reaction rate for HYD was set to the value for ethene, intermediate in reactivity between the ADOM species listed. No emissions are included.

In this experiment, the simple model will actually start from the true initial condition but because its processes are greatly simplified (as noted earlier) it quickly deviates from the ADOM model simulation which is taken as the truth (see columns 1 and 2 of Table 3). The observation error covariance matrix is identical to that used in Experiment 1. Since the model error is considerable, in this experiment, the model error standard deviations are assumed 10 times worse than in the previous experiment. Also, a cross-covariance between CO and O_3 is allowed. (Since the square root of Q and not Q itself is specified in the code, the presence of off-diagonal elements in $Q^{1/2}$ causes the model error standard deviation for CO to be greater than 10 times the value in Table 2.)

The data assimilation results are compared to the truth and the background in Fig. 3. In the truth, both NO_2 and NO decrease in time. However, the simple model must conserve nitrogen between these two species. Thus the background (dash-dot curve) estimation of NO_2 and NO is vastly different from the truth although both models started from the same initial conditions. Because the model error is considerably worse than the observation error (Table 3), the RRSQRT-EKF solution for the observed species, NO_2, NO and O_3, closely approximates the truth. The 4DVAR solution uses an assimilation period of 6 hours and apart from the first 12 hours, the solution is a vast improvement over the background, for the observed species. For CO, the RRSQRT-EKF solution is much better than the 4DVAR solution due to the prescribed cross-covariance of CO and O_3 (Table 3). For H_2O_2 and HO_2, both the 4DVAR and RRSQRT-EKF solutions provide some information and offer considerable improvement over the background. The HYD estimation, however, is not well obtained by either scheme. Finally, the RRSQRT-EKF solution for RO_2 is vastly overestimated on the second day. The 4DVAR solution for RO_2 is much better but is due to the overestimation of NO_2 and NO on the second day. The performance of 4DVAR is limited by the fact that the algorithm assumed that the model was perfect. In this example, the model is very far from perfect, but the scheme did provide some useful information about some species. The EKF which does account for model error performed better for the observed species and CO but also had difficulties in estimating HYD and RO_2. This is because the model error was assumed to be a Markov process whereas the actual model error is deterministic and not well represented by the chosen stochastic process.

Table 3. The initial conditions and error statistics for Experiment 2. The units are ppbv. A cross-covariance of 1.5×10^5 $ppbv^2$ between CO and O_3 was also used.

Species	Background	Truth	Obs error	Model error
NO_2	0.5	0.5	10	300
NO	0.3	0.3	10	300
O_3	80.0	80.0	30	900
CO	200.0	200.0	100	167
H_2O_2	2.0	2.0	1	30
HYD	0.5	0.5	1	30
HO_2	1.0E-19	1.0E-19	1	30
RO_2	1.0E-19	1.0E-19	1	30

In summary, we see that for the observed species, both schemes provide reasonable estimates. For some of the unobserved species (CO, H_2O_2, HO_2) some improvement over the background is obtained but for some species (HYD and RO_2), the estimate is not very good. Clearly the simple model is a very poor approximation to the complex model (from which the data was obtained). Nevertheless, data assimilation with the simple model was able to improve knowledge of some species. Thus, if the goal of the data assimilation is simply to obtain estimates of ozone, then it is possible that a simplified chemistry model may be sufficient (although we do NOT advocate the use of this particular chemistry model). However, if the details of the hydrocarbons and radicals are of interest, it seems unlikely that a simplified chemistry model would be of value.

The rank of the EKF solution shown in Fig. 3 is 8. It is interesting to note that a smaller rank of 4 would provide virtually the same solution. The only difference is that the NO solution is more noisy during the first 12 hours of integration (not shown).

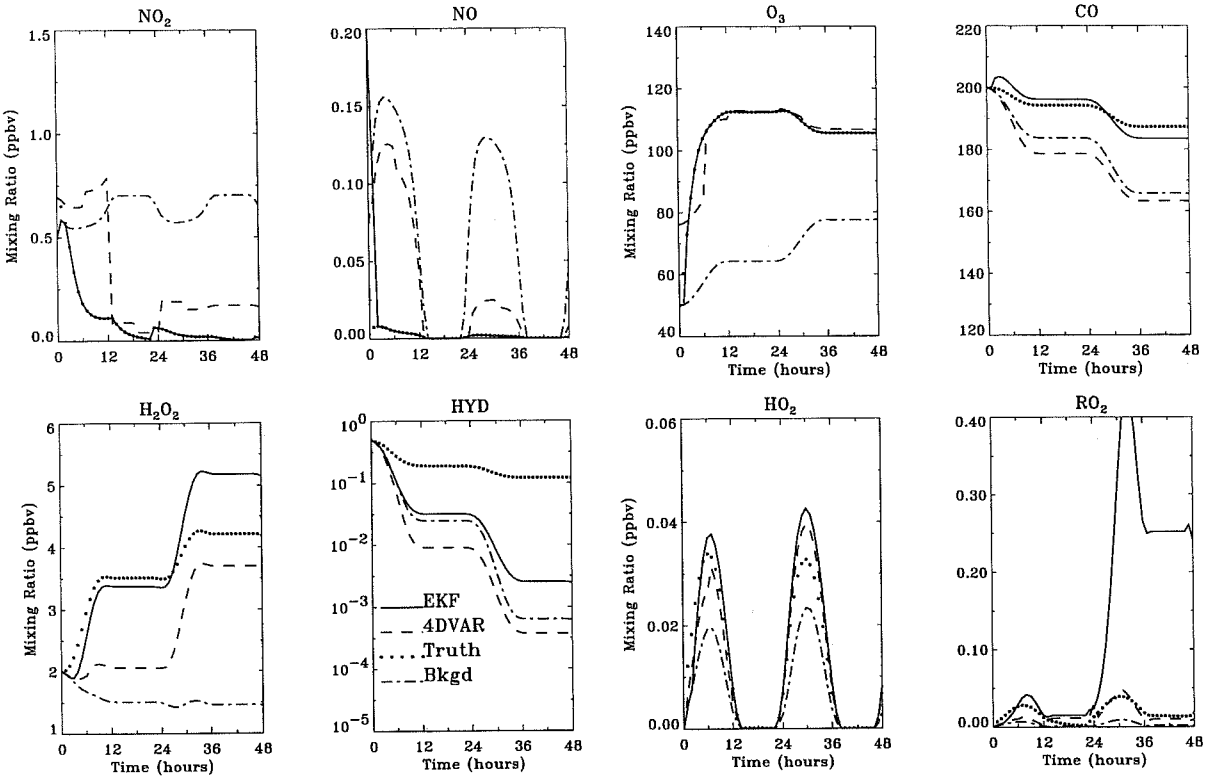


Figure 3. Experiment 3. Results of using three measurements (O_3,NO,NO_2). The truth is generated from a run of the ADOM model. The model has significant errors. There are no emissions.

5. SUMMARY

Two four-dimensional data assimilation schemes were implemented for a nonlinear chemistry box model. The model involves only 8 prognostic species and is a simplified representation of some of the processes involved in tropospheric ozone formation. The data assimilation schemes considered are a computationally efficient Kalman filter, the Reduced-Rank Square Root Extended Kalman Filter (RRSQRT-EKF) of Verlaan and Heemink (1995), and Four-Dimensional Variational Assimilation (4DVAR). In the first experiment, the truth was generated using emissions of NO of 0.6 ppbv/hr while the data assimilation schemes assumed a rate of 1.2 ppbv/hr. Nevertheless, both schemes were able to recover the true solution for most species. The initial error of CO and HYD were not corrected although the decay rate of both species was well obtained. In a second experiment, a more complex model was run to generate the truth. The simple model, however, was used for the data assimilation. No emissions were used in truth or model. The RRSQRT-EKF was able to recover the truth for the observed species and some information was obtained about some unobserved species: CO, H_2O_2 and HO_2. However, for HYD and RO_2, limited information was obtained. Therefore, we conclude that a simplified model may be of use for data assimilation applications if only ozone is of interest. If other species, particularly the hydrocarbons, are desired, it seems unlikely that a simplified chemistry model may be useful for estimation or prediction purposes.

In both experiments, the RRSQRT-EKF algorithm was shown to be more efficient than the conventional Kalman filter. The solution obtained with a reduced rank was almost identical to the full rank solution. Therefore this algorithm may be useful for applications involving chemistry-transport models. A rank reduction of several orders of magnitude is possible for the transport problem alone (Verlaan and Heemink 1995). We demonstrate here that further reduction is possible for the chemistry part, dependent on the particular chemistry model used.

REFERENCES

Elbern H., Schmidt, H. , and A. Ebel. (1997) Variational Data Assimilation for Tropospheric Chemistry Modeling., *Journal of Geophysical Research*, **102, D13**. 15,967-15,985.

Fisher, M., and D.J. Lary. (1995) Lagrangian four-dimensional variational data assimilation of chemical species. *Q.J.R. Meteorol. Soc.* **121**, 1681-1704.

Gear, C.W. (1971) *Numerical Initial Value Problems in Ordinary Differential Equations.* Prentice-Hall, NJ.

Venkatram, A., and Karamchandani, P. K. (1988) Testing a comprehensive acid deposition model. *Atmospheric Environment*,**22**, 737-747.

Verlaan, M., and A.W. Heemink. (1995) Data assimilation Schemes for Non-linear Shallow Water Flow Models. *Proceedings of the Second International Symposium on Assimilation of Observations*, Tokyo, Japan, 247-252.

DATA ASSIMILATION FOR TROPOSPHERIC OZONE PREDICTION PROBLEMS USING KALMAN FILTERING

A.W. Heemink[1], A.J. Segers[1], M. Verlaan[1]

[1]Delft University of Technology,
Faculty of Information Technology and Systems,
Department of Technical Mathematics and Informatics
Mekelweg 4, 2628 CD Delft, The Netherlands

P.H.J. Builtjes[2], M. van Loon[2]

[2]TNO, Institute of Environmental Sciences,
Postbus 342, 7300 AH Apeldoorn, The Netherlands

ABSTRACT

Large scale numerical air pollution models are of virtual importance for predicting air pollution concentrations and for reconstructing pollution emissions. Since these models are far from perfect, accurate results can only be obtained by integrating the model results with the concentration measurements that are available, both from satellites as from ground stations. Assimilating data into a numerical air pollution model is however, a procedure that required a huge amount of computer resources. Recently a number of efficient data assimilation algorithms based on Kalman filtering have been developed: the Ensemble Kalman filter and the Reduced Rank Square Root filter. The later algorithms has now been implemented in the LOTOS system for tropospheric ozone simulation studies of TNO. The resulting data assimilation system has been applied to identify the ozone generating mechanisms in the European region.

Key words: tropospheric ozone, data assimilation, Kalman filtering.

1. INTRODUCTION

For densely populated, industrialized countries, environmental modeling and simulation of pollution reduction scenarios is becoming more and more important in view of the growing awareness of damaging effects. Reduction and control of pollution is in general an expensive procedure, leading to high costs for society and industry. It is therefore necessary to determine, as accurately as possible, critical levels and to reduce and control pollution optimally so as to minimize costs.

To accomplish predictions of transport and exchange of chemical and biochemical constituents, accurate, three-dimensional (3D) mathematical simulation models must be used. These large-scale numerical models are based on the advection-diffusion equation. They are, however, far from perfect. Errors are introduced by fluctuations in the meteorological input or by poorly known parameters in the model. Furthermore, considerable uncertainty is associated with the open boundary conditions. Since measurements generally are more accurate then model predictions but limited to a small spatial scale only, it is appealing to combine them with the model results. This is a typical form of *data assimilation*: the incorporation of concentration measurements into a numerical model to improve the forecasts and to reconstruct pollution generating mechanisms. In this project we concentrate our attention to the prediction of tropospheric ozone.

The mechanisms which lead to the formation of tropospheric ozone differ considerably over Europe. At least five clearly different areas can be distinguished: Northern Europe (Scandinavia). Northwest and Central Europe, Southern Europe, and two areas over the sea, the North Sea with parts of the Atlantic, and the Mediterranean basin. The difference in ozone patterns and behavior are caused by a number of often interrelated phenomena. Different dynamical and physical properties, like the differences in temperature, cloud cover and land-sea circulation, differences in surfaces leading to differences in dry deposition, differences in anthropogenic precursor emissions and especially in biogene VOC-emissions, all this leads to differences in ozone patterns over Europe. Not only ground level ozone values will be different, also vertical profiles will be different and subsequently also vertical fluxes of ozone between the planetary boundary layer and the free troposphere, and the free troposphere and the stratosphere. The detailed investigation into these patterns and differences is severely hampered by the limited amount of observations available. Although the amount of observation of especially ozone has increased over the last years, the coverage is still limited to mostly the north-western part of the European continent. There is a severe lack of experimental data of ozone at ground level outside the north-western part of Europe, and hardly any data over the North Sea and Atlantic, and the Mediterranean Sea. Although at some loca-

tions vertical soundings are performed on a daily basis, our knowledge about the vertical profiles of ozone in the troposphere over Europe is very scarce. These available experimental data is insufficient to create a full 3-D data set of ozone in the troposphere over Europe. The three-dimensional Eulerian grid model LOTOS which calculate ozone patterns over Europe is available at TNO and has been used to study the controlling phenomena of ozone over the last decade. Combining these models with the ozone observations from ground level, vertical soundings and satellites by using data assimilation will lead to a coherent and complete full 3-D ozone data set over Europe. Such a data set will enable the differences of ozone patterns over Europe to be studied. Especially by performing budget studies the distinction between the influence of long range transport and of local ozone production can be revealed. Also vertical ozone fluxes over the troposphere and from the stratosphere can be analyzed and determined in detail.

2. DATA ASSIMILATION

Existing data assimilation schemes were developed mainly for numerical weather prediction. The most commonly used data assimilation technique in numerical weather prediction is optimal interpolation. This however, is not an accurate method because the correction produced by optimal interpolation is produced independently from the underlying numerical model, and is therefore not consistent with it.

Date assimilation schemes can also be developed by employing Kalman filtering (Ghil et al. 1981). In order to use a Kalman filter for assimilating data into a numerical transport model, this model is embedded into a stochastic environment by introducing a system noise process. In that way it is possible to account for the inaccuracies of the underlying deterministic system. By using a Kalman filter, the information provided by the resulting stochastic, dynamic model and the (noisy) state of the system. With a Kalman filter, unlike optimal interpolation, the statistics of the introduced noise are determined by using the stochastic extension of the model. Therefore the correction produced by this filter is consistent with the stochastic model.

In the last decenium Kalman filtering has gained acceptance as a powerful tool for data assimilation (Ghil et al. 1981), especially for linear or weakly nonlinear problems. The first applications of Kalman filtering for predicting air pollution were reported by Desalu et al. 1974, Koda & Seinfield 1978 and Fronza et al. 1979. The standard Kalman filter implementation imposes a very large burden on the computer for both memory and computation times. In order to obtain a computationally efficient filter, simplifications have to be introduced. In recent literature a number of new sub-optimal scheme's for solving large-scale filtering problems has been proposed (Cohn & Todling 1995, Verlaan & Heemink 1995, Evensen 1994, Fukumori & Malanotte-Rizzoli 1995).

Another approach to data assimilation which possesses many of the desirable features of Kalman filtering is the adjoint method, based on optimal control theory (see Courtier & Talagrand 1990). Here an unknown control function is introduced into the numerical model. Using the data available, this control function is identified by minimizing a cost function that compares the difference between the model results and the data. In order to obtain a computationally efficient procedure, the minimization is performed by using a gradient based algorithm where the gradient is determined by solving the adjoint problem. The adjoint method is more suitable for nonlinear data assimilation problems than Kalman filtering. However, a disadvantage of the adjoint approach is that it requires the implementation of the adjoint model. This is often very large programming effort compared to the implementation of the Kalman filter.

3. THE RIFTOZ PROJECT

One of the main contributions of Delft University of Technology and TNO to the RIFTOZ project is to develop a data assimilation scheme for the large scale numerical transport chemistry model LOTOS of TNO. Since the LOTOS model is still under development, the data assimilation method has to be flexible and model independent. Therefore the data assimilation is based on Kalman filtering. The Reduced Rank Square Root (RRSQRT) filter implementation does not require a tangent linear model nor an adjoint model. As a result it is relatively easy to implement and completely model independent. (Verlaan & Heemink 1995)

In the first year of the project the (first order) RRSQRT filter has been implemented for a test problem. Experiments were performed with simulated data to gain insight into the values of the various parameters of the data assimilation scheme. Attention has been concentrated on the performance of the filter with respect to the strong nonlinearity of the chemistry model. Results have been published by Van Loon and Heemink (1997).

In the second year of the project the RRSQRT filter has been improved by including second order terms in the algorithm. This new approach was evaluated in detail by using the test model with simulated data (see Segers et al. 1998). Furthermore the filter implementation was also coupled with the LOTOS model.

In the final phase of the project the filter technique has been applied on the LOTOS model. First, similar experiments as which were done with the test model has be applied with LOTOS in order to judge the system performance. These tests gave an indication of the required computing power. Second, the actual assimilation of real data with LOTOS calculations has been done by using the CRAY T3E of Delft University of Technology.

4. RESULTS WITH THE LOTOS SYSTEM

An important aspect of the RRSQRT filter is the computational efficiency. Here the dominating parameter is the number of modes used. The computational burden increase approximately linearly with this number. At the other hand also the accuracy of

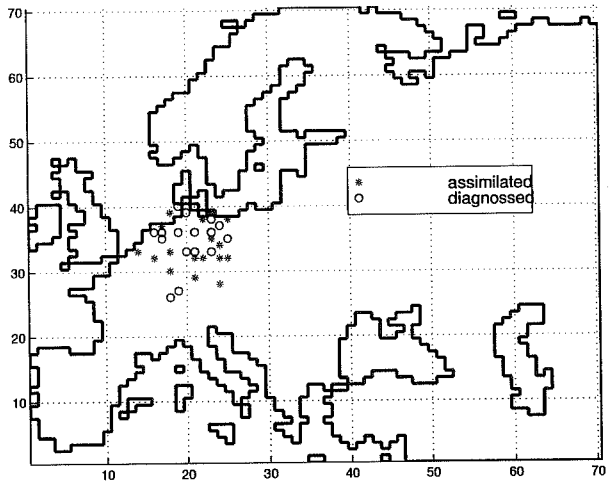


Figure 1. LOTOS grid and measurement sites.

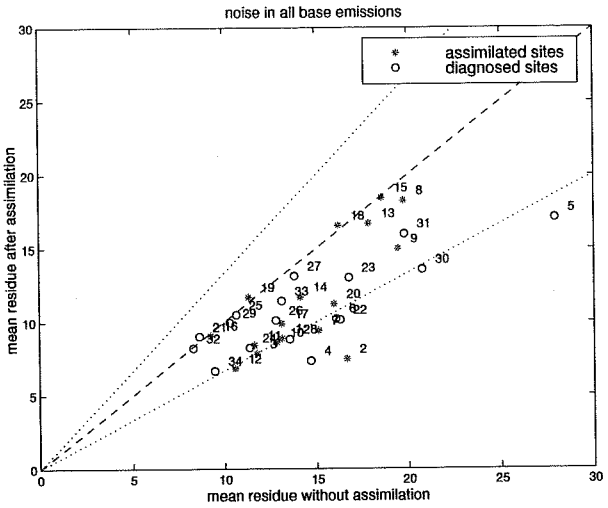


Figure 2. Mean residues during the assimilation period versus the mean residues from a background run (a single LOTOS run without assimilation). The dashed line denotes equality, while the dotted lines denote a difference of ±30%.

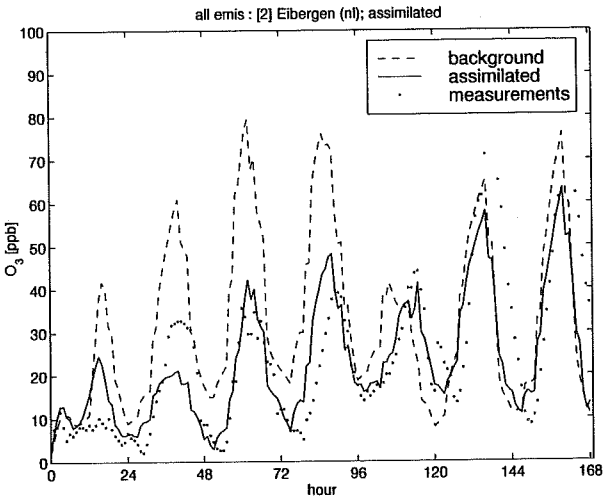


Figure 3. Ozone concentrations measured and calculated at the assimilated site Eibergen.

the data assimilation scheme increases with the number of modes. The experiments have shown that approximately 50 modes seems to be a good choice for atmospheric chemistry models. As a result the total computational burden will be 50-60 times the computational afford required for the underlying LOTOS model. Here we note that the filter implementation contains a lot of parallelism that can be exploited to improve the performance of the filter implementation on the CRAY T3E of Delft University of Technology considerably.

To find a useful specification of uncertainties in the LOTOS model, a sequence of assimilation experiments has been performed. The available set of measurements consists of hourly measured ground level values of ozone, measured at 34 sites in Germany and The Netherlands. The data from 17 of these sites are assimilated, while the other are only diagnosed (figure 1). The first week of august 1997 was taken as assimilation period.

One of the assimilation experiments made use of uncertainties specified for different emissions. The LOTOS model recognizes four types of emitted pollutants (VOC,NO_x,SO_x,and CO), emitted from 5 different antropogenic source categories; VOC is also emitted from 3 biogene sources. These emissions were supposed to vary from hour to hour with a standard deviation of 25% around their deterministic value. Spatial fluctuations in the variations are neglected. Because the measurement sites used for assimilation are located in in a rather small area, the later is not problematic. In fact, the chosen specification reflects the case of systematic error in the (modeling of the) emission data bases.

If all 23 emissions are specified to be uncertain in the way described, the filter is able to decrease the residue (difference between calculations and measurement data) with a maximum of about 30% in comparison with a non filtered simulation (figure 2). The improvement of the results is best shown by the decreased residue of the diagnosed sites, which have not been used in the assimilation process, but are only influenced by the assimilation of other data.

Figure 3 shows the ozone concentrations as calculated for the assimilated site Eibergen in The Netherlands. Without assimilation, the model overestimates the ozone concentrations, especially during the first days of the selected period. If however the available data is assimilated, the mean concentration almost perfectly follows the data after an initialization period of two days.

Investigation of the actual variations in the emissions used by the filter to reduce the residues, resulted in a selection of 8 emission data bases of the 23 in total, for which the chosen specification of uncertainty is most useful in an assimilation procedure. All of these selected emissions had an antropogenic source; this reflects the fact that biogene sources have a minor impact on the ozone concentrations in Western Europe where the measurements sites involved in this study are located.

A suitable way to judge the performance of the

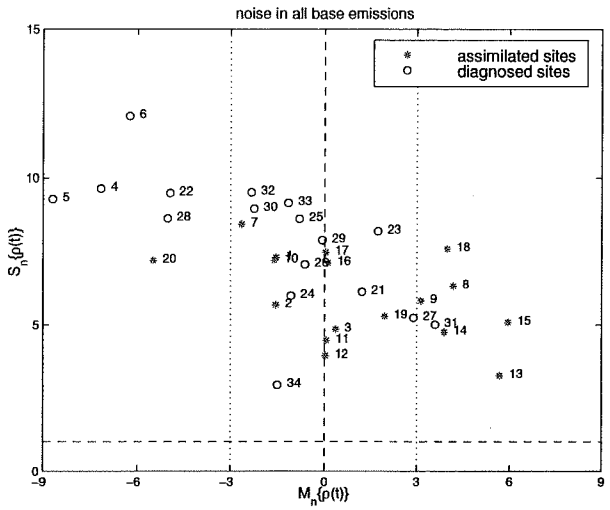


Figure 4. Sample mean M_n and sample standard deviation S_n of the ratio ρ between the actual residue and its expected standard deviation taken over the assimilation period, for each available site. The ratio has a theoretical distribution of N(0,1). The actual sample means are sometimes rather close to zero, but a standard deviation of one is never reached.

Kalman filter, is to compare the actual residues with their expected standard deviation. The later can be extracted from the covariance matrix of the state, and is in fact an impression of how accurate the filter guesses that the residue is. The ratio ρ between the residue and its guessed standard deviation is in theory N(0,1) distributed. In practice however, the spread of the ratios has a standard deviation larger than one, because the actual residues are often much larger than a few times their expected standard deviation. This is also the case for the previous described filter experiment (figure 4). The underestimation of the standard deviation of the residues indicates that the noise specification chosen here does not account for all uncertainties in the model, and that other specifications of uncertainty should be examined.

In order to extend the noise specification and to increase the filter performance, the impact of uncertainties in the upper boundary conditions was examined. Instead of using deterministic aloft concentrations, a new boundary condition was formed using total ozone columns measured with the GOME satellite instrument. Because of the large height of the columns in comparison with the height of the LOTOS grid (about 2 km), they could not serve as measurements directly. Debruyn et al. (1998) calculated daily values for the total ozone column in the first 9 km of the troposphere from the raw GOME data, with a standard deviation of 40%-50%. These columns served to calculate mean aloft concentration; a grid of stochastic variations (forced by 16 parameters) was added to the new aloft concentrations in order to represent the error.

Assimilation of data using a selected number of uncertain emissions in combination with the new boundary condition, did however not resulted in an additional decrease of the residues. Comparison of the new model output with the output of a model using the original boundary conditions showed that there is hardly any difference in calculated ground level concentrations. Similar, specification of uncertainties in the aloft concentrations did not have a significant impact on the results in case of an assimilation experiment. Even the ratio between residue and expected standard deviation was not improved through the increased amount of available uncertainties. These results show that the impact of the upper boundary on ground level concentrations is minor.

In spite of the minor impact of the upper boundary, the large amount of GOME data which is available makes the idea of using it still interesting. Use of the GOME data as measurements will be complicated through the limited height of the LOTOS grid, and the large error present in the profiles. The operational Kalman filter is however a suitable tool to decide what the minimum accuracy of the profiles should be in order to have a significant impact in a data assimilation procedure, and this will be subject of further research. Besides, the increased availability of accurate measurements in the vertical by means of balloon soundings makes it possible to examine the impact of assimilation of satellite data on the accuracy of LOTOS calculations at higher grid layers.

Systematic research of the impact of other (groups of) stochastic parameters on the assimilation results should result in a minimal stochastic extension of the LOTOS model for use in data assimilation. Increase of the number of uncertain parameters improves the results, but also implies an increased demand on computation capacity. Continued experiment should therefore point out uncertainties responsible for the majority of the residue.

5. CONCLUSIONS

The results up to now indicate that the assimilation of data into the LOTOS model with a Kalman filter procedure is feasible. The nonlinearities do not seem to cause serious problems and the computational burden of the algorithm is large but not too large. For the final computations, however, a very powerful computer was required (CRAY T3E). The filter results with the real data show a good performance of the system. In a number of measurement locations that have not been used in the assimilation procedure, the model results with assimilated data are 10%–30% more accurate then the original model results. The use of GOME data does not seem to have have a significant impact on the results in the current setup.

ACKNOWLEDGMENTS

This work has been carried out with financial support of the European Community as a part of the research project RIFTOZ.

REFERENCES

Cohn, S.E., Todling, R., Approximate Kalman filters for unstable dynamics, In *Second Int. symposium*

on Assimilation of observations in Meteorology and Oceanography, pages 241–246, Tokyo, WMO, 1995.

Courtier, P., Talagrand, O., Variational assimilation of meteorological observations with direct and adjoint shallow water equations, *Tellus*, 42A:531–549, 1990.

Desalu, A.A., Gould, L.A., Schweppe, F.C., Dynamic estimation of air pollution, *IEEE Trans. Autom. Control*, AC-19:904–910, 1974.

Debruyn, W., Brems, E., Lissens, G., RIFTOZ: Regional Differences in Tropospheric Ozone in Europe, *technical report*, 1998/TAP/R/049, 1998

Evensen, E., Sequential data assimilation with a nonlinear QG model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, 99:10143–10162, 1994.

Fronza, G., Spirito, A., Tonielli, A., Real-time forecast of air pollution episodes in the Venetian region. Part 2: the Kalman predictor, *Appl. Math. Modelling*, 3:409–415, 1979.

Fukumori, I., Malanotte-Rizzoli, P., An approximate Kalman filter for ocean data assimilation; an example with an idealized Gulf Stream model, *J Geophys, Res*, 100:6777–6793, 1995.

Ghil, M., Cohn, S.E., Tavantzis, J., Bube, K., Isaacson, E., Application of estimation theory to numerical weather prediction, In M. Ghil L. Bengtsson and E. Källén, editors, *Dynamic Meteorology: Data Assimilation Methods*, New York, Springer-Verlag, 1981.

Koda, M., Seinfield J.H., Estimation of urban air pollution, *Automatic*, 14:583–595, 1978.

Segers, A.J., Heemink, A.W., Verlaan, M., Van Loon, M.,. Nonlinear Kalman Filters for Atmospheric Chemistry Models, Submitted to *American Geophysical Union*, 1998.

Verlaan, M., Heemink, A.W.,. Data assimilation schemes for non-linear shallow water flow models, In *Proceedings of the Second International Symposium on Assimilation of Observations*, pages 247–252, Tokyo, Japan, WMO, 1995.