Norsk institutt for luftforskning
Norwegian Institute for Air Research

Observations, lecture 9

Data Assimilation of Research Satellite Data

William Lahoz, wal@nilu.no

WAVACS Summer School, ”Water vapour in the climate system”

Cargese, Corsica, France, 14-26 September 2009

Outline

- Information
- Data assimilation: adding value
- Data assimilation of research satellite data: ozone, water vapour
- Evaluation of observations and models using data assimilation
- Overview
- Bibliography

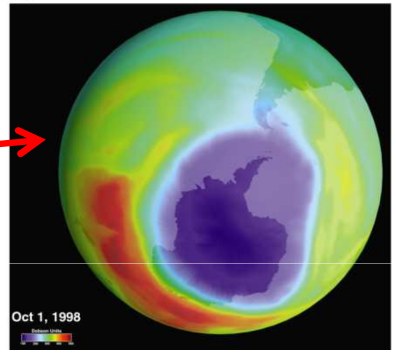
Information:

Need for information:

Main challenges to society require information for an intelligent response, including making choices on future action - examples:

- Climate change
- Impact of extreme weather
- Environmental degradation
- Ozone loss

Antarctic ozone hole (Dobson units):
Red high total column ozone; blue low total column ozone from TOMS



We can take action according to information obtained:

- Future behaviour of system of interest, future events – prediction
- Test understanding of system & adjust understanding – hypothesis testing
- Understand cause of events, change, mitigate, adjust – attribute cause & effect

Chain of information processing:

- Gather information
- Test hypotheses based on this information
- Build methods to use information – attribute cause & effect
- Use methods to make predictions

Need two ingredients:

- Means of gathering information – observations (different types)
- Methods to build on information gathered, organize information gathered – models (conceptual, numerical...)

Observations: roughly direct link with system of interest via measurements

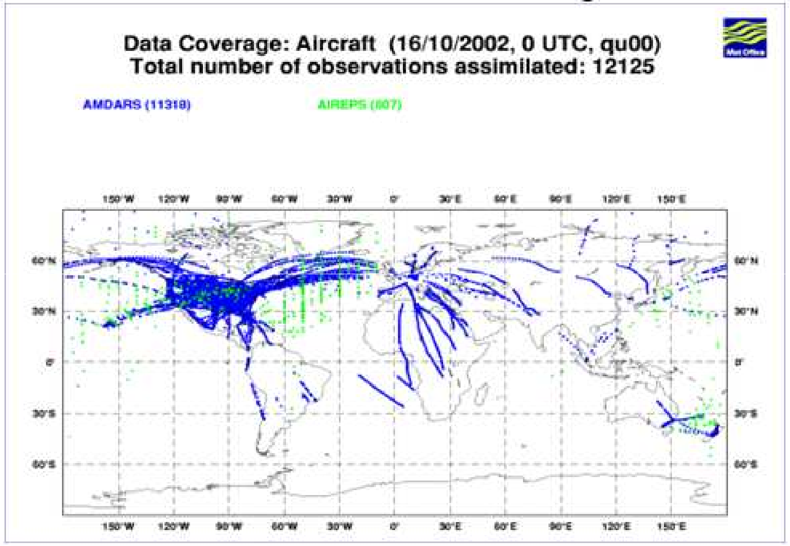
Models: roughly indirect link with system of interest – embody information received from measurements, experience & theory

Models & observations are sources of information

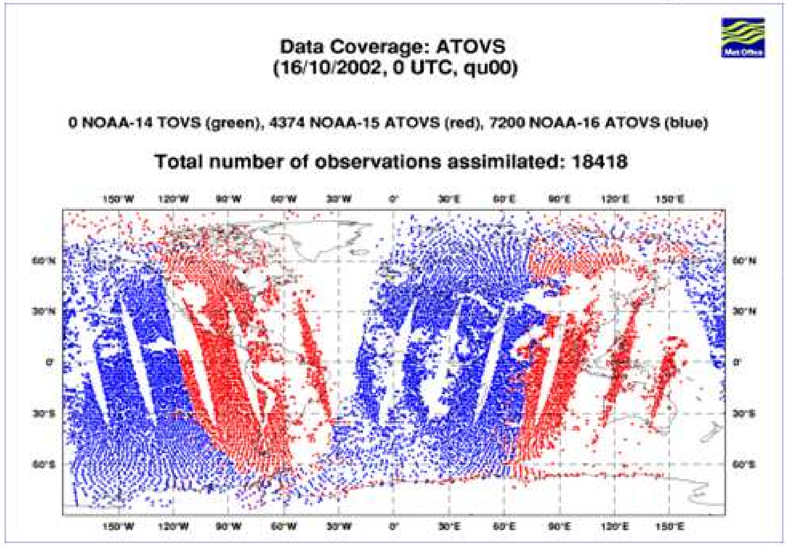
Sources of information:

Measurements: observations – different time & spatial scales

Aircraft Local coverage



ATOVS Global coverage



Understanding: embodied in models. Can be:

Qualitative: e.g. higher velocity, higher KE; quantitative: e.g. KE = $(1/2)\,mv^2$

Characteristics of information:
Observational errors:

- Random - precision
- Systematic – bias
- Representativeness – e.g. different spatial scales: sonde, satellite

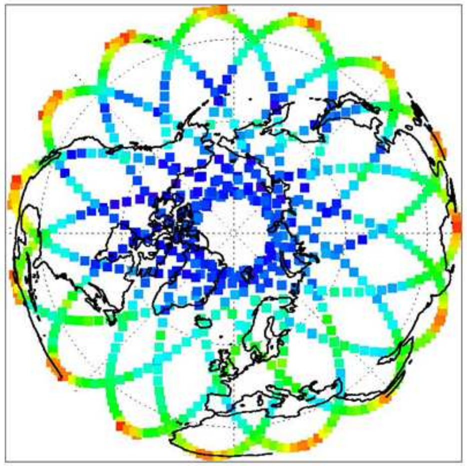
Models also have errors
- Construction of models – incomplete models
- Imperfect simulation of "real world"

Information (observations/models) has errors – need to take this into account

Observations (measurements) are discrete in space and time - information provided by observations has gaps

We would like to fill in gaps

UARS MLS ozone 10 hPa, 1 Feb 1997



Objective ways of filling in information gaps:
Algorithm attributes:

- Consistent (mathematically, physically,...) rules
- Objective (impartial principles) rules: max/min of a function,...

Algorithm: Model (propagates information in space and time)

- Linear interpolation
- Navier-Stokes equations
- Chemistry equations
- Parametrizations

Can build a hierarchy of models

Mathematics: "What combination of the observation and model information is optimal?" & estimate of the errors of the "optimal" or "best" estimate ->

"Data assimilation" (Earth Observation data/model fusion): has strong links to several mathematical disciplines, including control theory & Bayesian statistics

Mathematics:
Combine information from a model & observations plus errors

3D-Variational method (variational, minimize penalty function, J):

$$J = \frac{1}{2}[x - x^b]^T \mathbf{B}^{-1}[x - x^b] + \frac{1}{2}[y - H(x)]^T \mathbf{R}^{-1}[y - H(x)].$$

Model information

Observational information

Add time:

4D-Var

Kalman filter method (sequential):

Forecast step

$$x_n^f = \mathbf{M}_{n-1} x_{n-1}^a;$$

$$\mathbf{P}_n^f = \mathbf{M}_{n-1} \mathbf{P}_{n-1}^a \mathbf{M}_{n-1}^T + \mathbf{Q}_{n-1};$$

Run model

Analysis step

$$x_n^a = x_n^f + \mathbf{K}_n[y_n - \mathbf{H}_n x_n^f];$$

$$\mathbf{K}_n = \mathbf{P}_n^f \mathbf{H}_n^T [\mathbf{R}_n + \mathbf{H}_n \mathbf{P}_n^f \mathbf{H}_n^T]^{-1};$$

$$\mathbf{P}_n^a = [\mathbf{I} - \mathbf{K}_n \mathbf{H}_n] \mathbf{P}_n^f.$$

Combine model and observational information

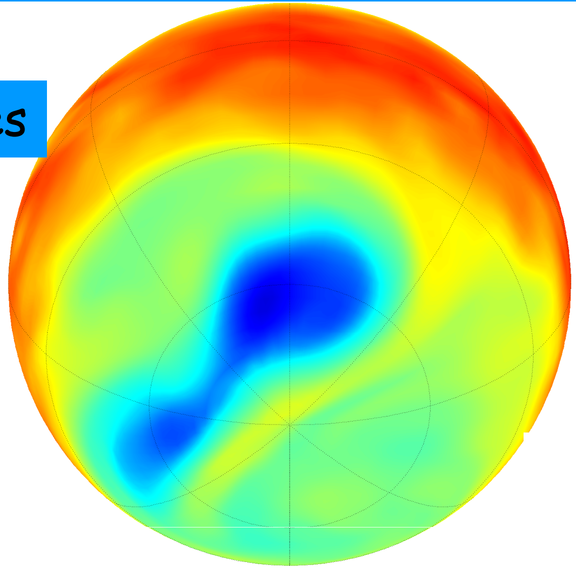
Also ensemble methods, e.g., Ensemble Kalman Filter

Data assimilation - adding value:

Ozone 10hPa, 12Z 23 Sep 2002

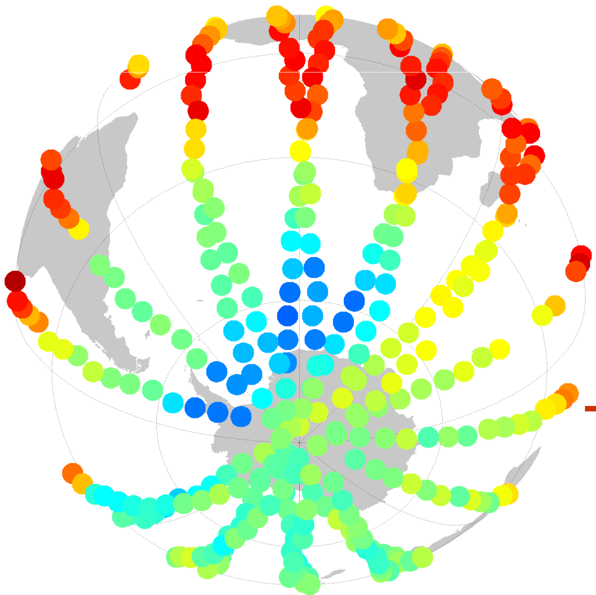
Red: high ozone
Blue: low ozone

Analyses



DA adds value to both observations and model

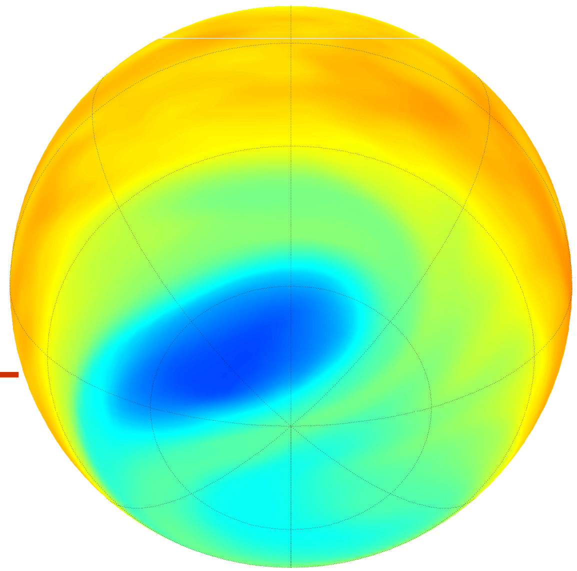
Geer et al., QJRMS ,2006
Lahoz et al.,ACP, 2007



MIPAS observations

DA

Errors



6 day forecast

Data assimilation and Numerical Weather Prediction, NWP:

Key idea: Confronting models with observations

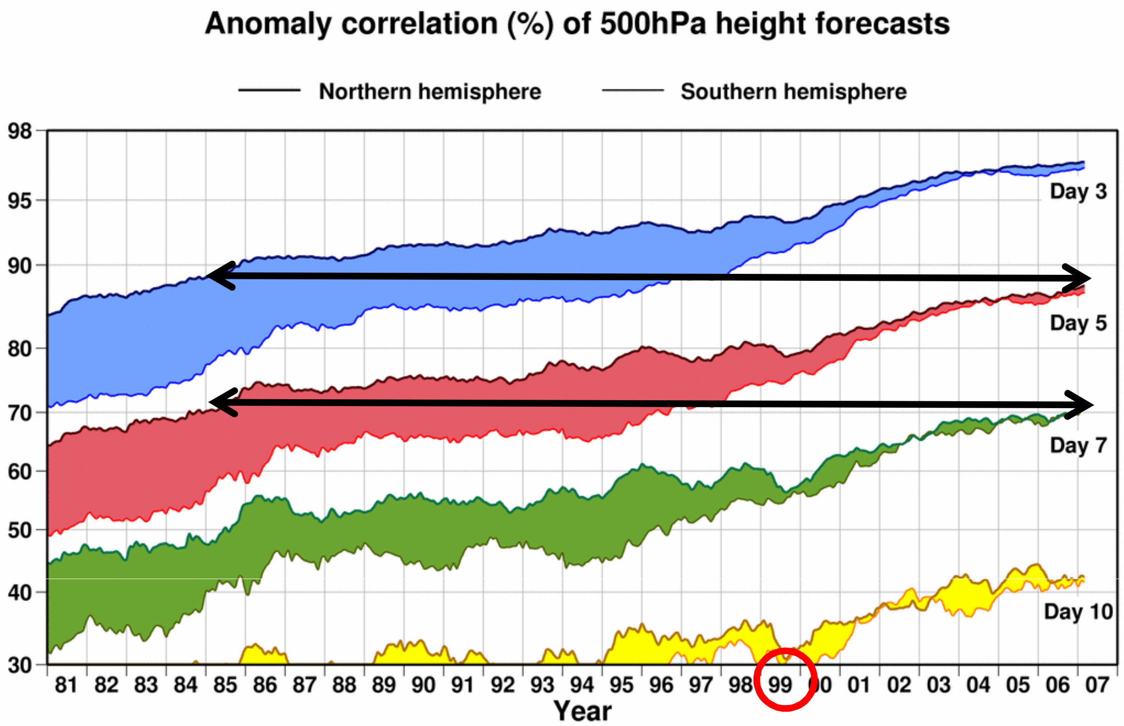
Progress in NWP has been a combination of:

- Better models: higher resolution, better processes
- Better observations: satellites
- Better use of observations: bias correction, quality-control, radiances
- Better computing power
- Data assimilation: better use of observations and models; use of 4d-variational (4d-var) approach

This has allowed observations and models to be evaluated and improved

This has allowed improvement in NWP forecasts (e.g. European Centre for Medium-Range Weather Forecasts, ECMWF)

NWP: success for data assimilation



AC coeffs, 3-, 5-, 7- & 10-day ECMWF 500 hPa ht forecasts for extra-tropical NH & SH, plotted as annual running means of archived monthly-mean scores for Jan 1980 - Nov 2006. Values plotted for a particular month are averages over that month & 11 preceding months. Colour shadings show differences in scores between two hemispheres at the forecast ranges indicated (After *Simmons & Hollingsworth, QJRMS, 2002*)

Impact of satellite observations, impact of data assimilation

Towards end of 1999: a more advanced 4D-Var developed & significant changes in the GOS mainly due to launch of 1st ATOVS instrument onboard NOAA satellites

Evaluation of observations and models:

We can apply NWP ideas to evaluating observations & models

- Observations: Do they have Gaussian errors? Are they biased?

Self-consistency
Data assimilation as a transfer standard: estimate bias

- Models:
  - Chemistry models: Chemistry-transport models (CTMs)
  - Climate models: General circulation models (GCMs)
  - Climate-chemistry models (CCMs)

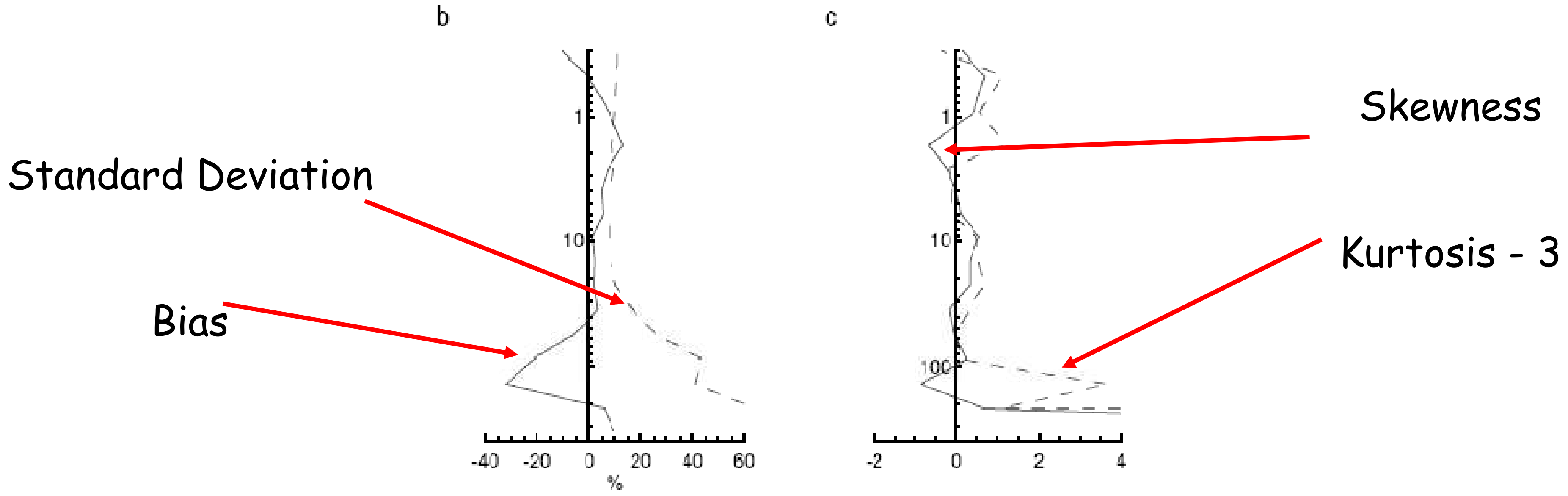
  Can extend ideas to other models: Earth System models (ESMs)

Obs quality:

DA: Self-consistency of MIPAS ozone data

Statistics: 14-28 Sep 2002

Obs (MIPAS) minus short-range Forecast (model), OmF



OmF ~ 0 in stratosphere

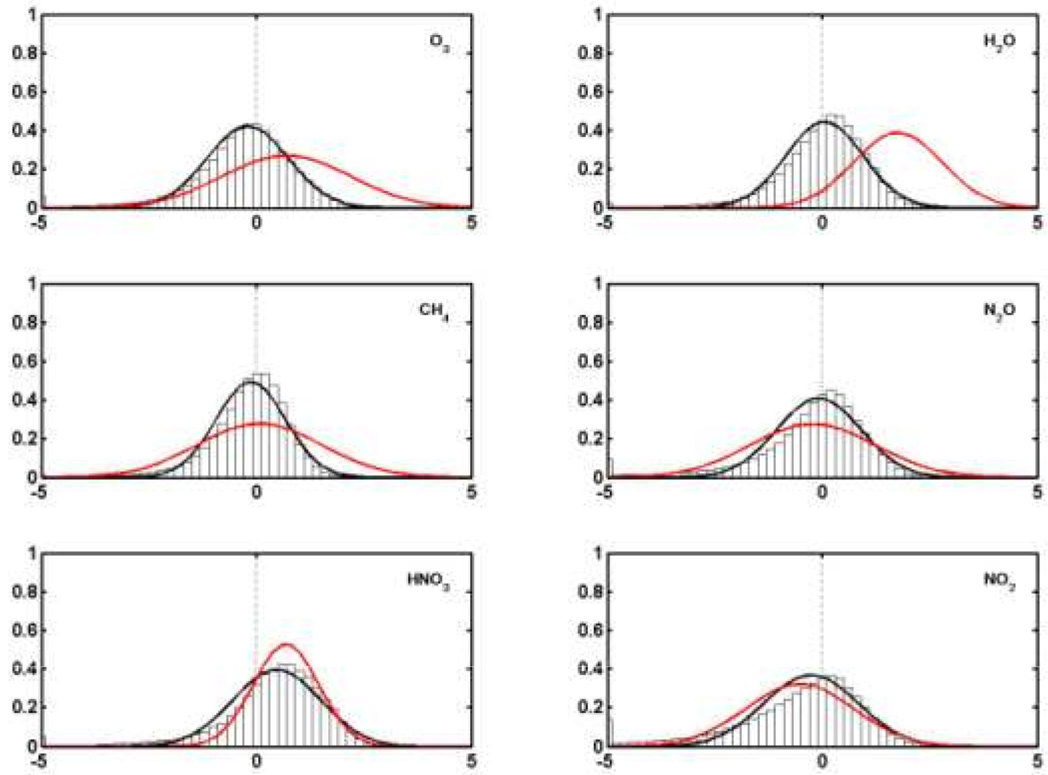
Consistent with Gaussian errors in stratosphere

Geer et al. QJRMS (2006); Struthers et al. JGR (2002)

Self-consistency & added value

OmF:

Observation minus forecast



Evaluation of analyses using histograms of OmF differences (normalized by observation error) averaged for stratosphere, globe & August 2003 for six stratospheric constituents: O_3 (top left), H_2O (top right), CH_4 (middle left), N_2O (middle right), HNO_3 (bottom left) and NO_2 (bottom right). Constituent observations rom ESA MIPAS off-line retrievals. Frequency of histograms normalized to lie between 0 and 1. Black line is Gaussian fit to histograms; red line is Gaussian fit from model run without assimilation.

Results support assumption of Gaussian errors in observations & forecast, & show analyses are closer to observations than simulations from model run without assimilation. Experiments performed at BIRA-IASB. With permission from *Lahoz et al., ACP, 2007*

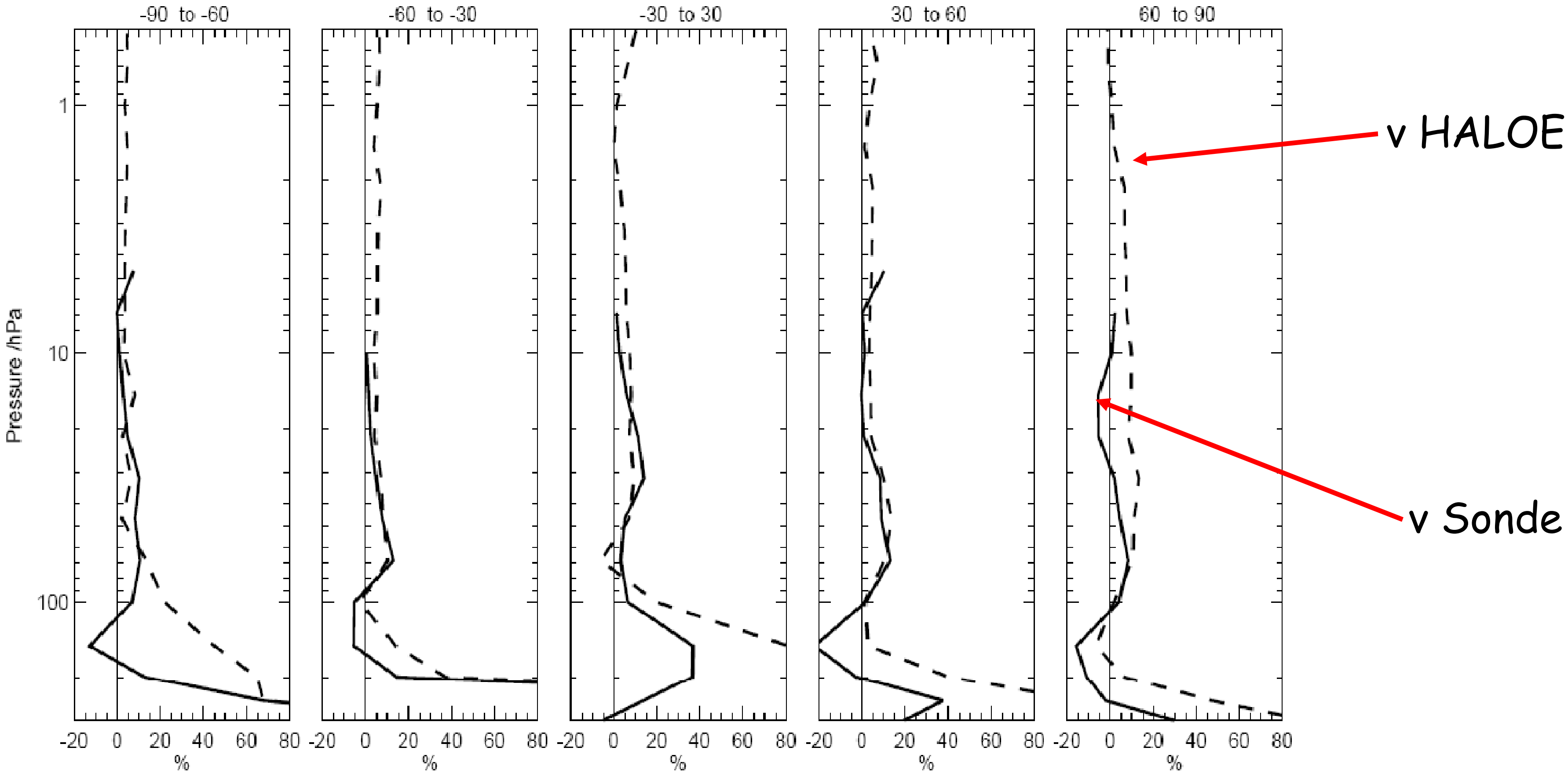
Obs bias:

DA: Evaluation of MIPAS ozone using independent data

BASCOE used as "interpolating" analysis

Statistics: 18 Aug – 30 Nov, (Obs1-Analysis) – (Obs2-Analysis):

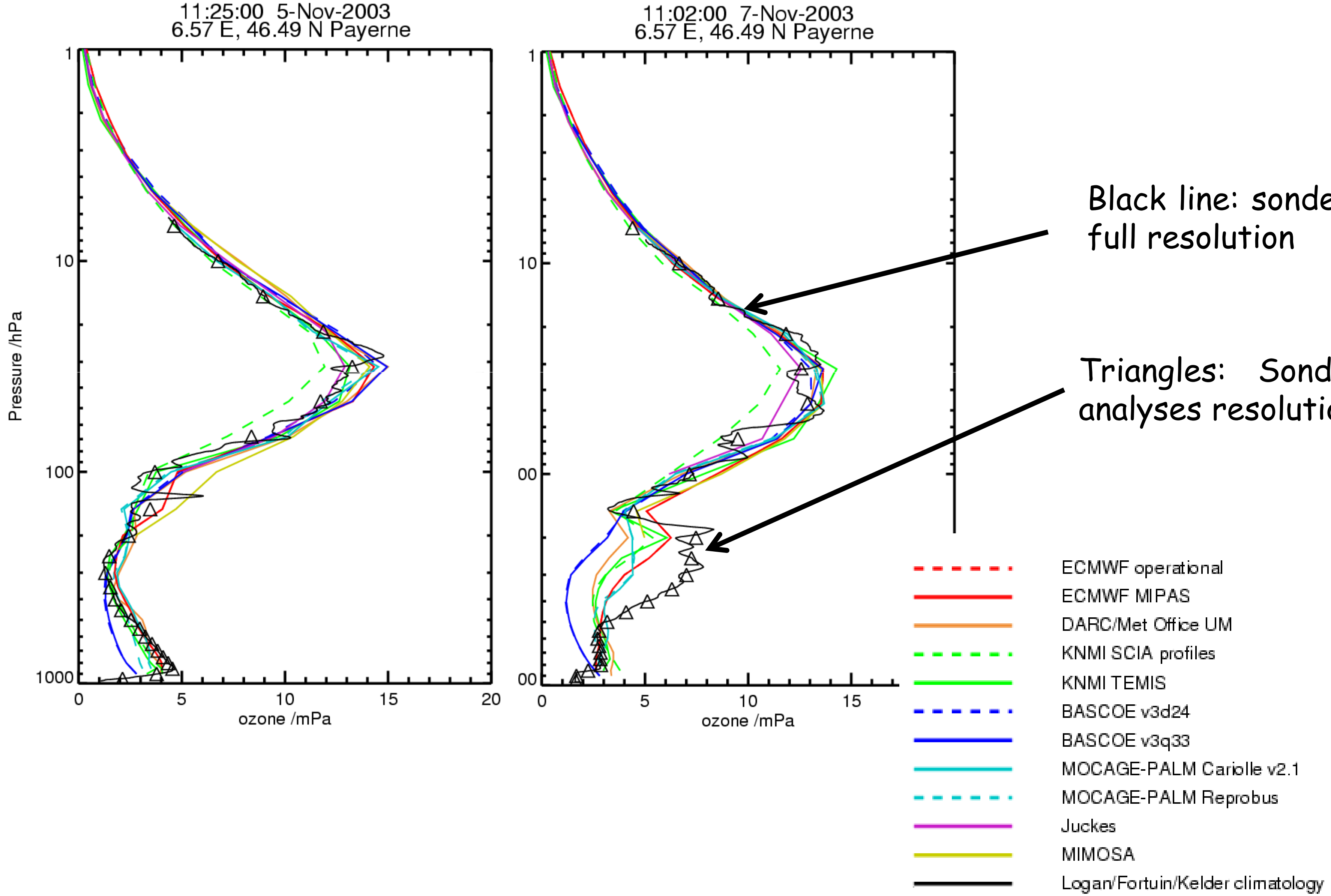
Geer et al. ACP (2006); Lahoz et al. ACP (2007), QJRMS (2006, 2007)



Bias in MIPAS ozone generally positive: ~5% - ~10% -> feedback to MIPAS team

Mid latitude Upper Trop/Lower Strat: Payerne ozonesonde profiles
(*Geer et al., ACP, 2006*)

Take account of different resolution of matched datasets



Black line: sonde at full resolution

Triangles: Sonde at analyses resolution

Model set-up:

Accuracy of combined ozone information (obs/model)

ASSET project

Geer et al., ACP, 2007
Lahoz et al., ACP, 2007a, b

Good performance in stratosphere: Within 5-10% of HALOE instrument

Complexity of chemistry: Parametrization v comprehensive (e.g. ECMWF v BASCOE)

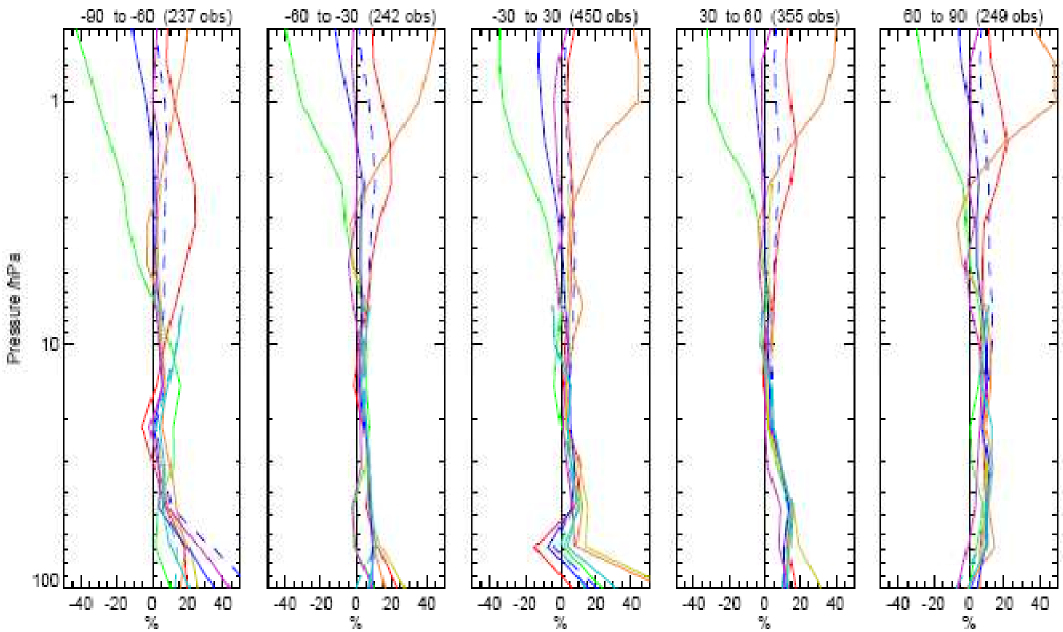


Fig. 10. Mean of analysis minus HALOE differences, normalized by climatology, for the period 18 August–30 November 2003. See Fig. 9 for colour key. The numbers in brackets indicate the HALOE/analysis coincidences within each latitude bin. Units: percent. These data are used to evaluate the performance of the ozone analyses. Based on Geer et al. (2006).

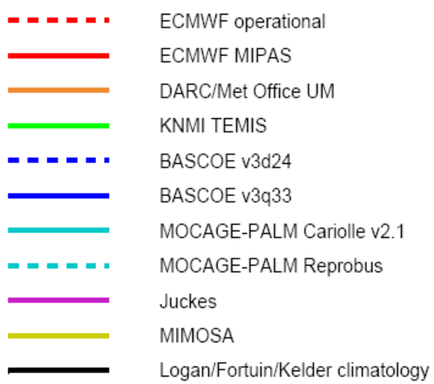
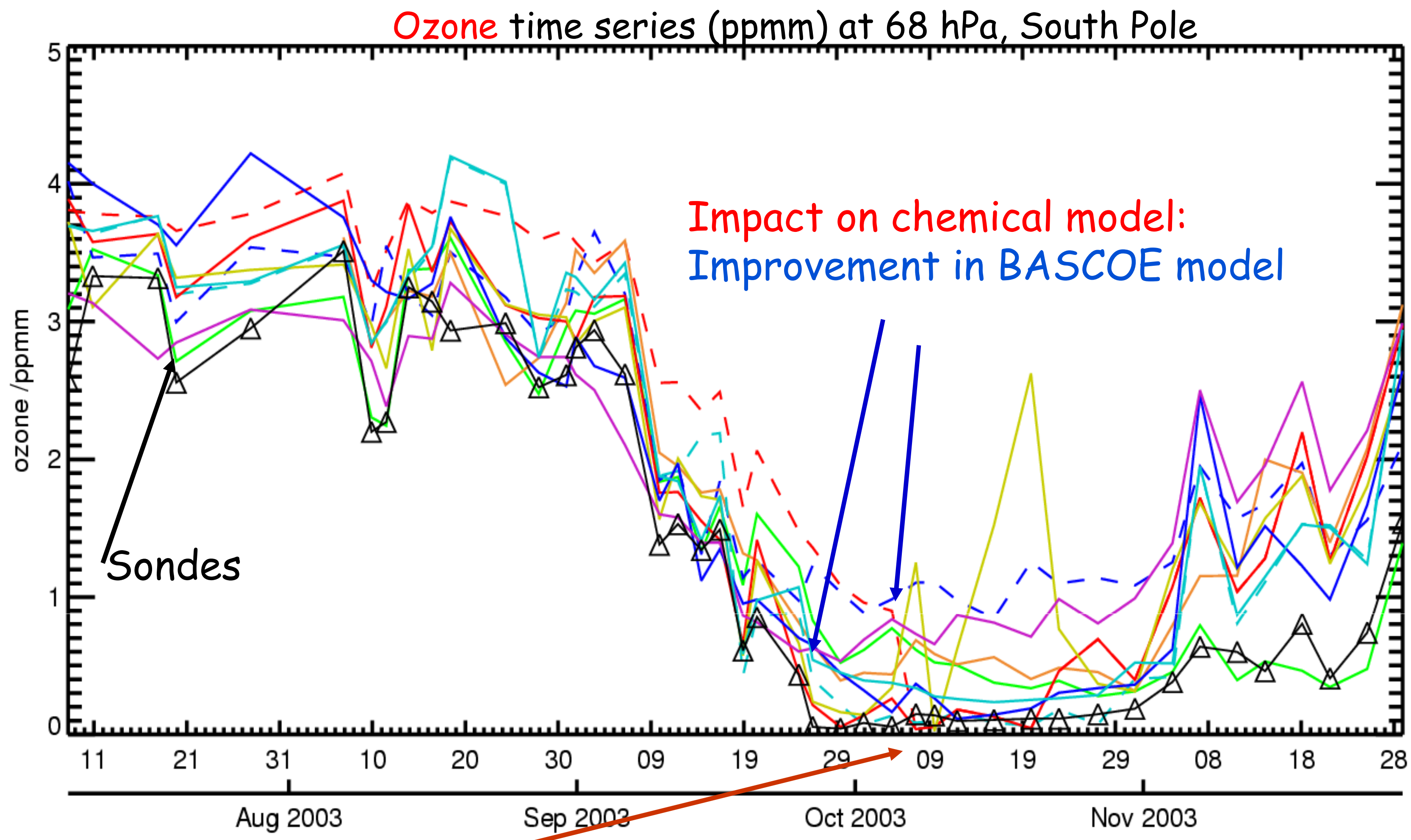
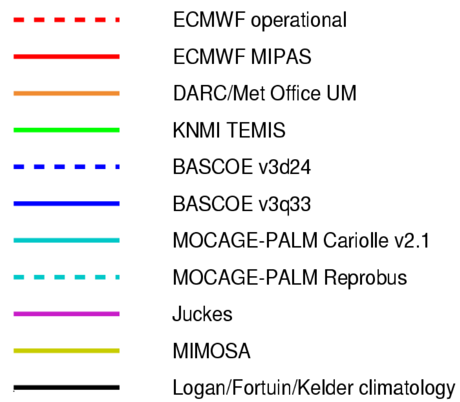


Fig. 9. Colour key used in Figs. 10–11.



Impact of new chemical observations:
Operational ECMWF assimilates MIPAS ozone



Geer et al. ACP, 2006,2007
Lahoz et al. ACP, 2007a, b

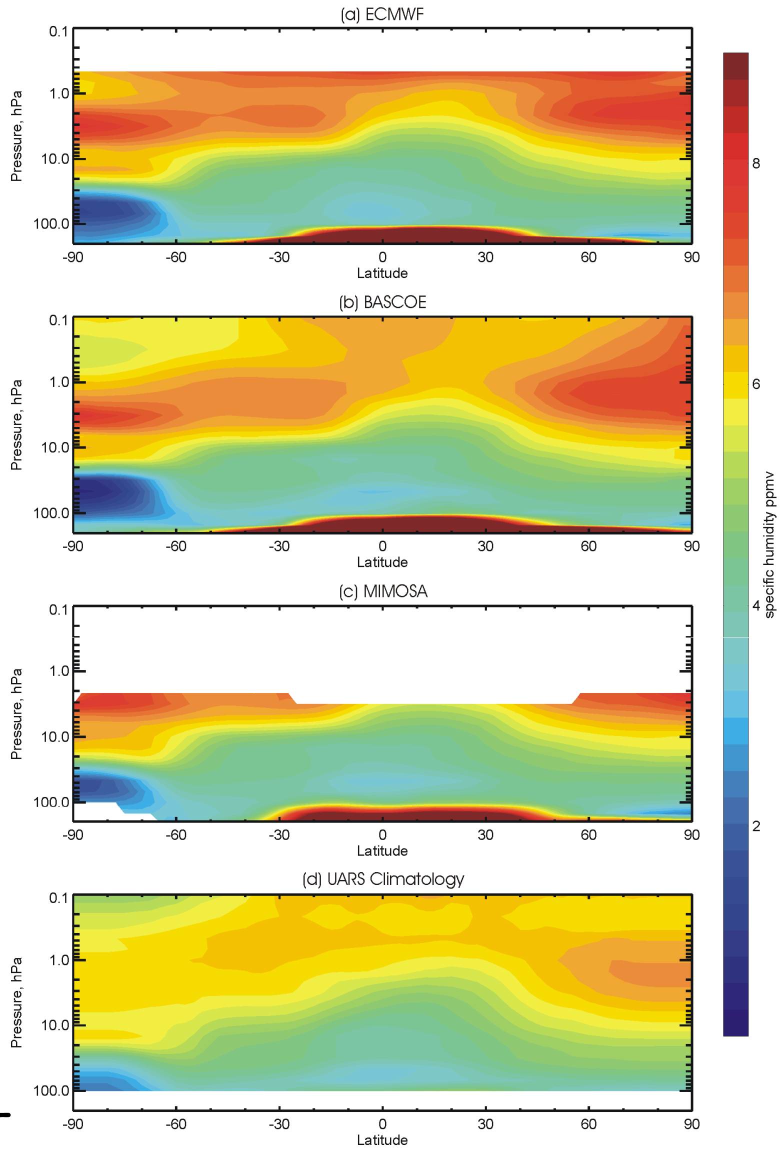
Accuracy of combined water vapour (WV) information (obs/model)

ASSET project

Lahoz et al., ACP, 2007a, b
Thornton et al., ACP, 2009

Main features of stratospheric WV captured:

- Tropical WV minimum,
- SH polar vortex WV minimum
- Brewer-Dobson circulation
- Mesosphere: analyses wetter than UARS clim & reflect wet bias of MIPAS obs



Monthly zonal mean specific humidity analyses, Sep 2003: (a) ECMWF, (b) BASCOE, (c) MIMOSA; (d) UARS clim
MIPAS WV profiles assimilated in ECMWF, BASCOE & MIMOSA analyses.
Blue: relatively low specific humidity values
Red: relatively high specific humidity values. Units: ppmv.

Water vapour analysed fields, 68 hPa, 21 Sep 2003, 1200 UT
Various data assimilation systems

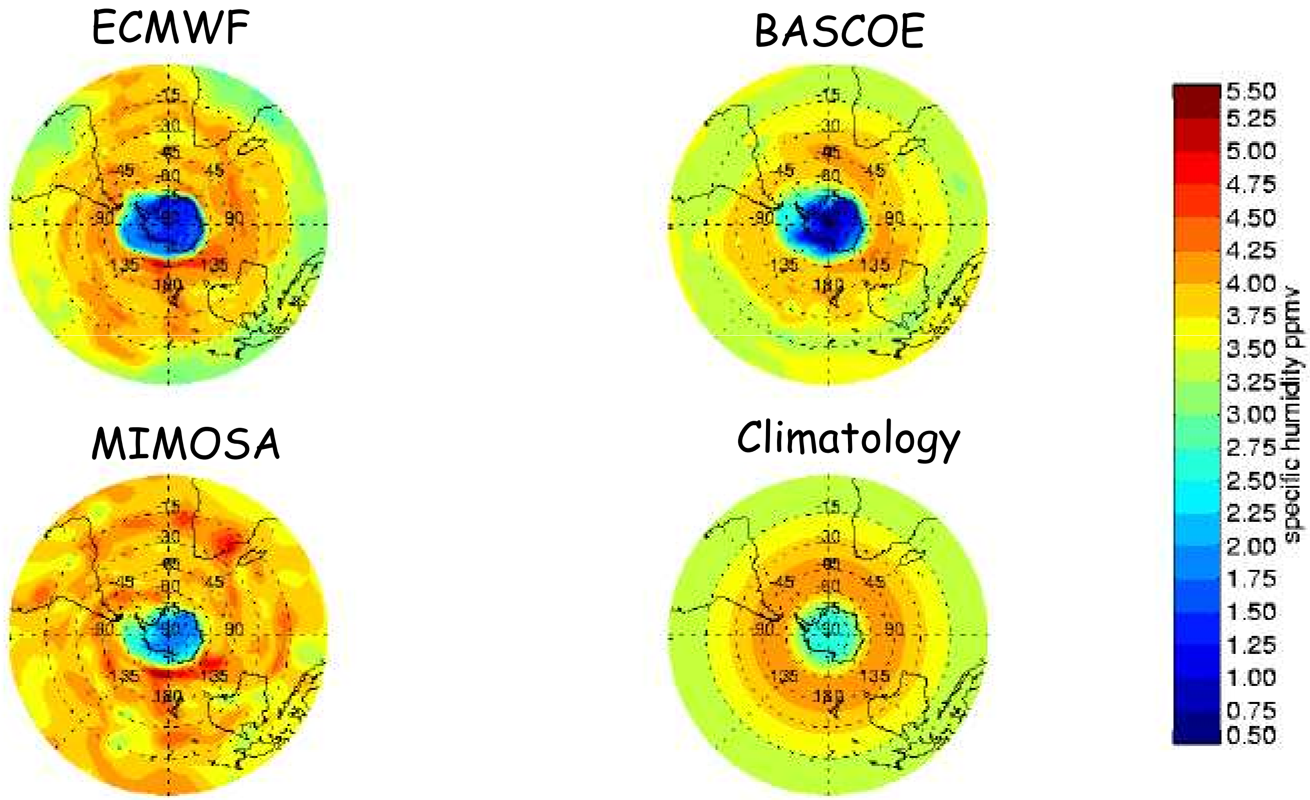


Fig. 8. Polar stereographic projection of the specific humidity field (ppmv) for the southern hemisphere on 21 September 2003 12:00 UT at 68 hPa, for ECMWF (top left), BASCOE (top right), MIMOSA (bottom left) and climatology (bottom right).

Cal-val of water vapour analyses

Comparison with independent data:

HALOE (black triangles)

(1) ECMWF & Met Office (GCM)
ECMWF (Red): humidity control variable
Normalized RH, reducing to normalized specific humidity in strat
Met Office (Light blue): normalized specific humidity

(2) BASCOE (Green) & MIMOSA (Dark blue) (CTM)

Specific humidity control variable

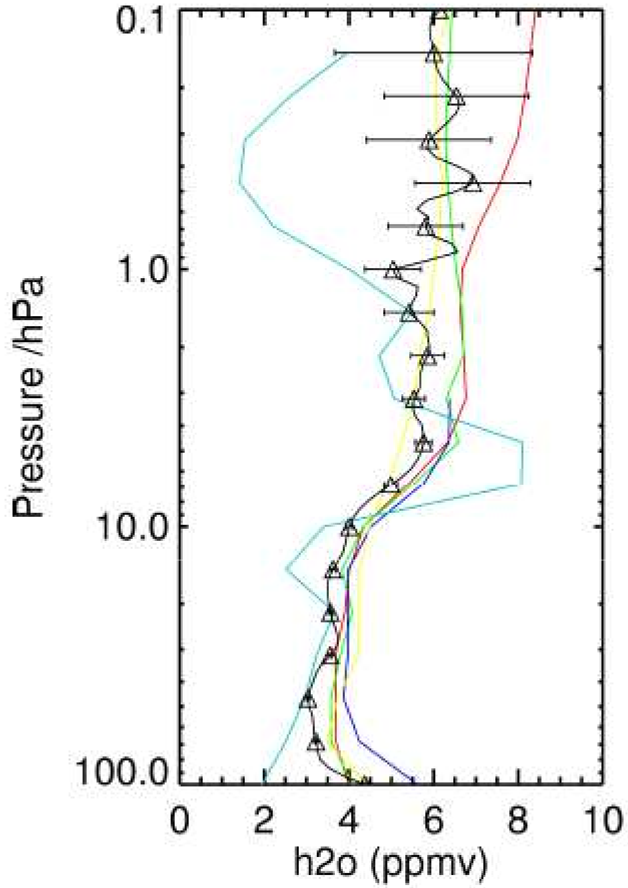


Fig. 2. Comparison of a HALOE profile at a resolution of 30 points per decade (black line), HALOE values (black triangles) and error bars on common grid levels with the four analyses, ECMWF (Red), BASCOE (Green), MIMOSA (Dark Blue) and the Met Office (Light Blue), and the UARS climatology (Yellow) on 2 September 2003 at 04:19, 6.45° E, 65.73° N.

Water vapour analyses

Analyses minus Obs (AmO) diagnostics

Obs:

MIPAS (self-consistency)

HALOE, SAGE II, POAMIII (independent)

MIPAS:
AmO within 5% (10-1 hPa)

HALOE:
AmO within 20% (10-1 hPa)

Positive bias in analyses

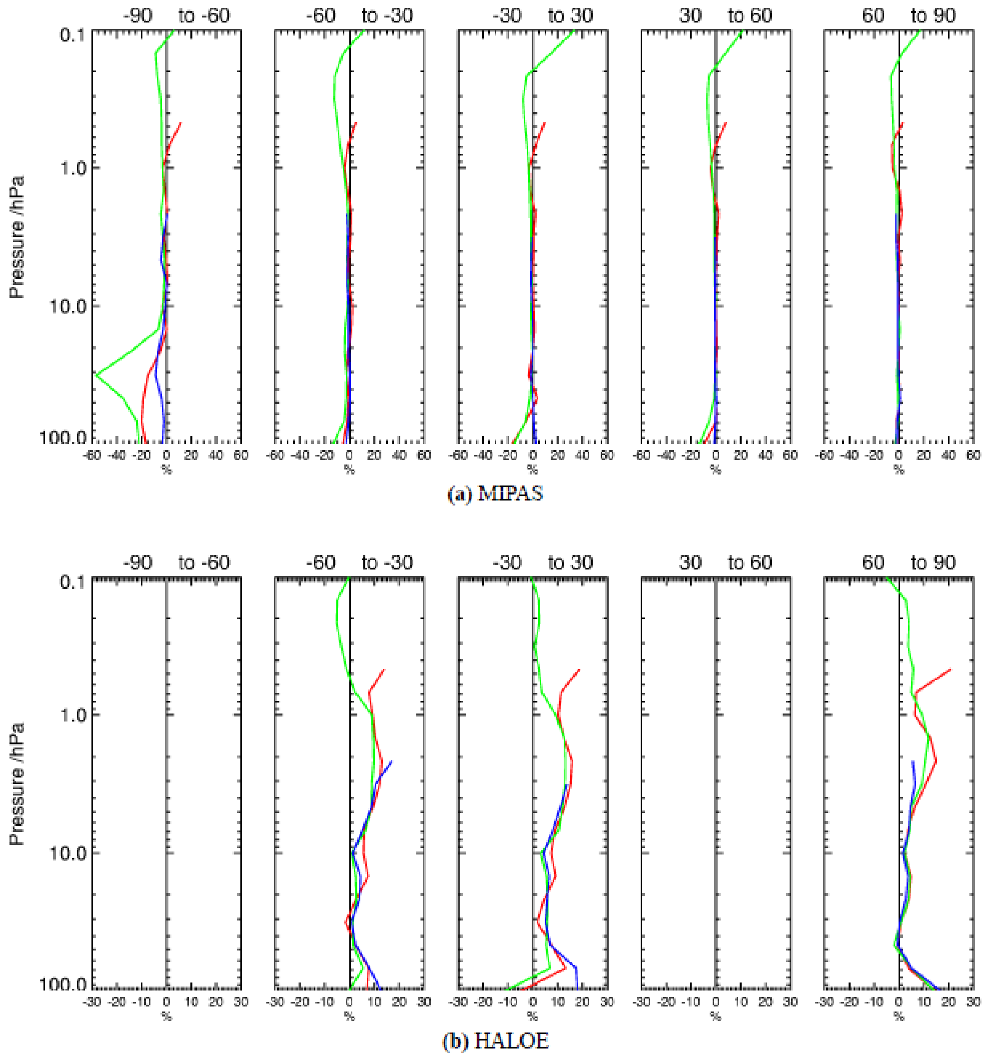


Fig. 5. Mean of (Analysis–Observations) water vapour mixing ratio, normalised by climatology (in percent) over the intercomparison period, for ECMWF (Red), BASCOE (Green) and MIMOSA (Blue), for the five different latitude bins. For rows **(a)** to **(d)**, the analyses are compared with MIPAS, HALOE, SAGE II and POAM III data, respectively. If there are not any satellite profiles available, the graphs are blank.

SAGE II:
AmO within 20% (10-1 hPa)

Positive bias in analyses

POAM III:
AmO within 20% (10-1 hPa)

Negative bias in analyses

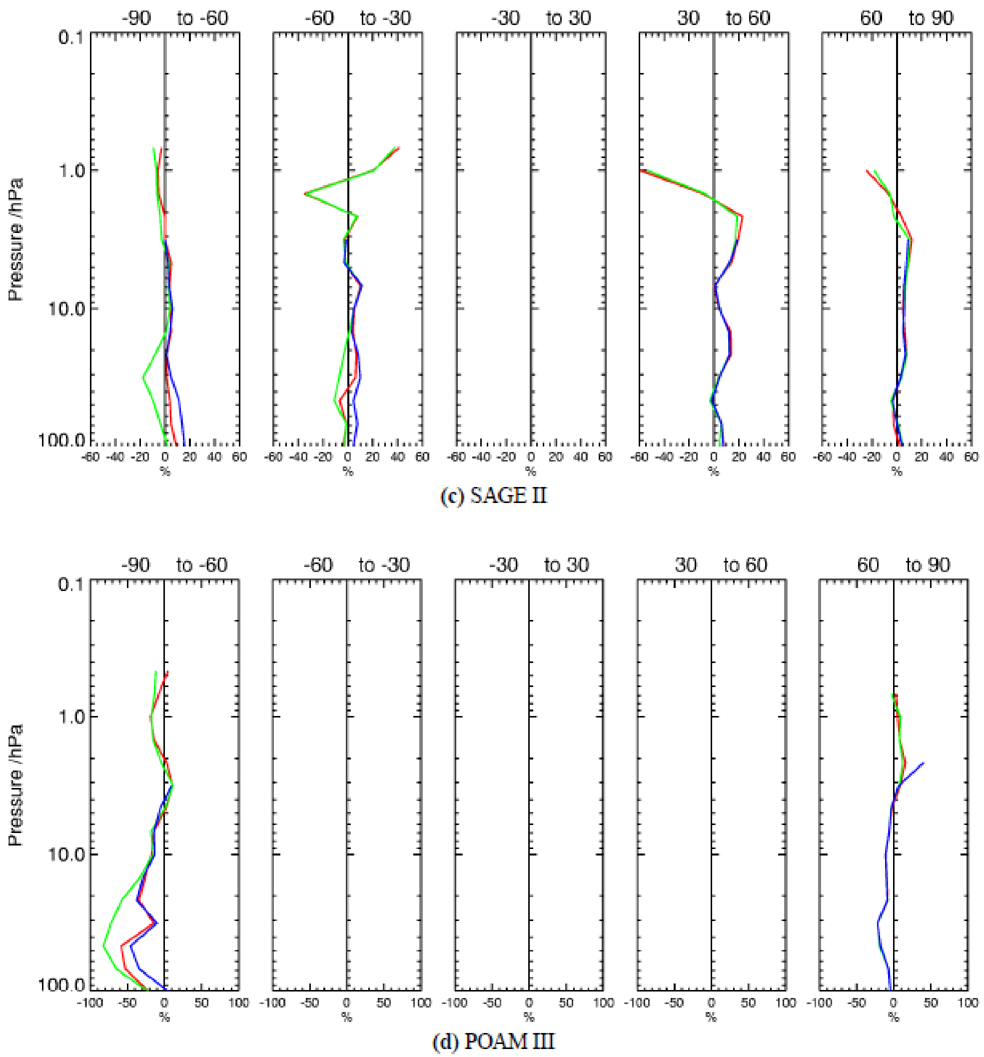


Fig. 5. Continued.

Water vapour control variable

Choice:

- Relative humidity, RH
- Normalized RH
- Normalized specific humidity

Aim:
Control variable with desirable properties:

- Usable in troposphere & stratosphere
- Approx Gaussian background errors (**B**)
- Temp & humidity increments decoupled
- Realistic vertical correlations

Tests at Met Office

The Met Office has investigated the impact of varying the control variable in the assimilation of MIPAS humidity data.

To achieve this, the Met Office have combined the ideas of Dee and da Silva (2003) and Hólm et al. (2002), and defined a normalized relative humidity variable. The impact of the normalization is to produce a considerably better conditioned background error covariance matrix and consequently the minimization in the 3D-Var algorithm is much faster. The removal of the influence of temperature increments reduces spurious upper stratospheric increments.

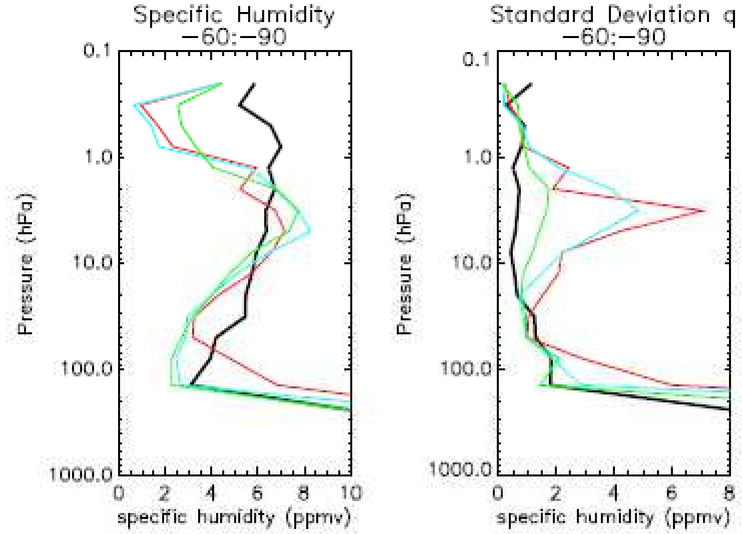


Fig. 2. Met Office mean (left-hand plot) and standard deviation (right-hand plot) of specific humidity profiles for 25 September 2002 over the 60° S–90° S region. Black: MIPAS observations; red: analyses using a relative humidity (RH) control variable; blue: analyses using a normalized RH control variable; green: analyses using a normalized specific humidity control variable. Units: ppmv.

Water Vapour control variable: details of Met Office work

(1) What is the control variable (why choice?)
(2) How to calculate background error covariance

A possible way forward is to reconsider the way in which the humidity background error covariance matrix is calculated, rather than to scale existing covariances, as we have attempted here. A recent study by Jackson et al. (2008), has shown that the 3D-Var analyses on which the NMC covariance calculation was based, suffer from a lack of dynamical balance between the mass and wind fields. Spurious gravity waves are generated to restore this balance and can be seen in a 24-h forecast and consequently are present in the NMC error covariance matrix. This may explain the unrealistic vertical correlations in the humidity covariance matrix reported here. Of course, the presence of a spurious gravity wave signal in the error covariances may have an adverse effect on all analysis variables, but this effect may be greatest for humidity because of the lack of suitable stratospheric humidity observations to constrain the analyses on which the NMC method is based.

Solutions to this problem include the use of better-balanced analyses in the NMC calculation. For example, at the Met Office, 3D-Var analyses have recently been superseded by 4D-Var analyses, which are in much better dynamical balance and give rise to NMC covariances which contain a reduced spurious gravity wave signal. Other techniques of calculating error covariances may also be more effective at removing spurious gravity waves. Such techniques include the method described by Polavarapu et al. (2005b) (the so-called Canadian Quick covariances) and the use of ensembles. Ensembles are used to calculate background error covariances at ECMWF. This may explain why the ECMWF stratospheric humidity analyses presented here are much more accurate than the corresponding Met Office analyses, even though the humidity control variable used at both institutes is very similar. It is not easy to apply the ECMWF covariances to the Met Office DA scheme, due to the different model formulations, however the use of ensembles to generate covariances is being further developed at the Met Office.

Climate models:

Recent NWP-based ideas to evaluate climate models:

- CAPT initiative – improve parametrizations in GCMs (Phillips et al. 2004) – requires accurate NWP analyses; systematic error can be largely attributed to parametrization deficiencies

- Seamless prediction – fundamental physical/dynamical processes common to both weather & seasonal forecasts, & climate-change timescales (*Palmer et al. , BAMS, 2008*)

    - Proposal: probabilistic validation of models at timescales where validation data exist (e.g. daily, seasonal,...) can be used to calibrate climate-change probabilities at longer timescales.
    - Need for calibration reflects a need for model improvement

- Estimating climate model parameters (e.g. gravity wave drag; early days)

- Uncertainty analyses & ensemble experiments (early days)

Observations must constrain parameters of interest

Overview

- Helpful to regard observations & models as sources of information
- Data assimilation invaluable for studying polar stratosphere:
    - Fills gaps between observations (need a model)
    - Allows use of heterogeneous measurements
    - Makes sense of observations (multiple, heterogeneous)
- Data assimilation can add value to observations & models, compared to information that each can supply on their own
- Data assimilation allows evaluation of models/observations
- Data assimilation underpins evaluation of impact of current observation types using OSEs (observing system experiments), and the future global observing system using OSSEs (observing system simulation experiments)

Crucial for setting up Global Observing System (GOS)

– see observations lecture 14

Bibliography:

Buehner, M., et al., 2008: Inter-comparison of 4D-Var and EnKF systems for operational deterministic NWP. WWRP/THORPEX Workshop on 4D-VAR and Ensemble Kalman Filter intercomparisons, Bs. As., Argentina, Nov 2008.

Courtier, P., J.-N. Thépaut & A. Hollingsworth, 1994: A strategy for operational implementation of 4D-var, using an incremental approach. Q. J. R. Meteorol. Soc., 120, 1367-1387.

Daley, R., 1991: Atmospheric Data Analysis, R. Daley, Cambridge University Press.

Durran, D.R., 1999: Numerical Methods for Wave Equations in Geophysical Fluid Dynamics. Spring-Verlag, New York.

Elbern, H., et al., 2007: Emission rate and chemical state estimation by 4-dimensional variational inversion. Atmos. Chem. Phys., 7, 3749–3769, 2007.

Errera, Q. & D. Fonteyn, 2001: Four-dimensional variational chemical data assimilation of CRISTA stratospheric measurements. J. Geophys. Res., 106, 12253-12265.

Errera, Q., et al., 2008: 4D-Var assimilation of MIPAS chemical observations: Ozone and nitrogen dioxide analyses. Atmos. Chem. Phys., 8, 6169-6187.

Eskes, H.J., 2003: Stratospheric ozone: satellite observations, data assimilation and forecasts. Paper presented in Workshop on Recent developments in data assimilation for atmosphere and ocean. ECMWF, Reading, UK. http://www.ecmwf.int.

Eskes, H.J, et al., 2003: Assimilation of GOME total-ozone satellite observations in a three-dimensional tracer-transport model. Q. J. R. Meteorol. Soc., 129, 1663-1681.
Geer, A.J., et al., 2006:Assimilation of stratospheric ozone from MIPAS into a global general circulation model: the September 2002 vortex split. Q. J. R. Meteorol. Soc., 132, 231-257.
Geer, A.J., et al., 2006: The ASSET intercomparison of ozone analyses: method and first results. Atmos. Chem. Phys., 6, 5445-5474.
Geer, A.J., et al., 2007: Evaluation of linear ozone photochemistry parametrizations in a stratosphere-troposphere data assimilation system. Atmos. Chem. Phys., 7, 939-959.
Grewe, V. & Sausen, R., 2009: Comment on "Quantitative performance metrics for stratospheric-resolving chemistry-climate models" by Waugh and Eyring. Atmos. Chem. Phys. Discuss., 9, 14141-14164.
Hollinsgworth, A., et al., 2008: The Global Earth-system Monitoring using Satellite and in-situ data (GEMS) Project: Towards a monitoring and forecasting system for atmospheric composition. Bull. Amer. Meteorol. Soc., doi:10.1175/2008BAMS2355.1.
IGACO, 2004: The Changing Atmosphere. An Integrated Global Atmospheric Chemistry Observation theme for the IGOS partnership. ESA SP-1282, Report GAW No. 159 (WMO TD No. 1235), September 2004; Implementation up-date, December 2004. Available from: http://www.igospartners.org/docsTHEM.htm.
Kalnay, E., 2003: Atmospheric Modeling, Data Assimilation and Predictability. CUP, Cambridge.

Khattatov, B.V., et al., 1999: Assimilation of photochemically active species and a case analysis of UARS data. J. Geophys. Res., 104, 18715-18737.
Lahoz, W.A. & Errera, 2009: Constituent Assimilation. Data Assimilation: Making sense of observations, Eds. W.A. Lahoz, B. Khattatov and R. Ménard, Springer.
Lahoz, W.A., et al., 2005: An Observing System Simulation Experiment to evaluate the scientific merit of wind and ozone measurements from the future SWIFT instrument. Q. J. R. Meteorol. Soc., 131, 503-523.
Lahoz, W.A., et al., 2007: The Assimilation of Envisat data (ASSET) project. Atmos. Chem. Phys., 7, 1773-1796.
Lahoz, W.A., et al., 2007: Data assimilation of stratospheric constituents: A review. Atmos. Chem. Phys., 7, 5745-5773.
Lahoz, W.A., Khattatov, B. & Ménard, 2009: Data assimilation for dummies. Data Assimilation: Making sense of observations, Eds. W.A. Lahoz, B. Khattatov and R. Ménard, Springer.
Lorenc, A.C., 2003: The potential of the ensemble Kalman filter for NWP: A comparison with 4D-Var. Q. J. R. Meteorol. Soc., 129, 3183-3204.
Masutani, M., et al., 2009: Observing System Simulation Experiments. Data Assimilation: Making sense of observations, Eds. W.A. Lahoz, B. Khattatov and R. Ménard, Springer.
NATO ASI, 2003: Data Assimilation for the Earth System, Eds. R. Swinbank, V. Shutyaev and W.A. Lahoz. Kluwer.

Palmer, T.N., et al., 2008: Towards seamless prediction. Bull. Amer. Meteorol. Soc., DOI:10.1175/BAMS-89-4-459.

Phillips, T.J., et al., 2004: Evaluating parameterizations in general circulation models: Climate simulation meets weather prediction. Bull. Amer. Meteorol. Soc., DOI:10.1175/BAMS-85-12.

Pulido, M. & Thuburn, J., 2008: The seasonal cycle of gravity wave drag in the middle atmosphere. J. Climate, 21, 4664-4679.

Rodgers, C.D., 2001: Inverse Methods for Atmospheric Sounding. World Scientific.

Rood, R.B., 2009: The role of the model in the data assimilation system. Data Assimilation: Making sense of observations, Eds. W.A. Lahoz, B. Khattatov and R. Ménard, Springer.

Schoeberl, M. R., et al., 2003: A comparison of the lower stratospheric age spectra derived from a general circulation model and two data assimilation systems. J. Geophys. Res., 108, 4113, doi:10.1029/2002JD002652.

Simmons, A.J. & Hollingsworth, A., 2002: Some aspects of the improvement in skill of numerical weather prediction. Q. J. R. Meteorol. Soc. 128, 647-677.

Simmons, A.J. & Hollingsworth, A., 2002: Some aspects of the improvement in skill of numerical weather prediction. Q. J. R. Meteorol. Soc. 128, 647-677.
Struthers, H., et al., 2002: Assimilation of ozone profiles and total column measurements into a global General Circulation Model. J. Geophys. Res., 107, 10.1029/2001JD000957.
Swinbank, R. & O'Neill, A., 1994: A stratosphere-troposphere data assimilation system. Mon. Weather Rev., 122, 686-702.
Tarantola, A., 1987: Inverse Problem Theory, A. Tarantola, Elsevier.
Thornton, H., et al., 2009: The ASSET intercomparison of stratosphere and lower mesosphere humidity analyses. Atmos. Chem. Phys., 9, 995-1016.
Waugh, D.W. & Eyring, V., 2008: Quantitative performance metrics for stratospheric-resolving chemistry-climate models. Atmos. Chem. Phys, 8, 5699-5713.