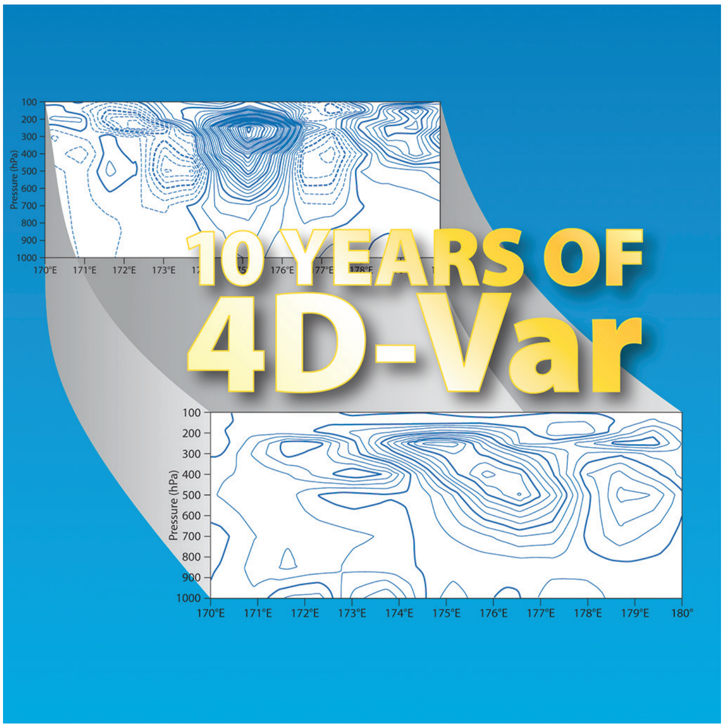
ECMWF Feature article

from Newsletter Number 115 – Spring 2008

METEOROLOGY

ECMWF's 4D-Var data assimilation system – the genesis and ten years in operations



www.ecmwf.int/en/about/news-centre/media-resources

doi:10.21957/wnmguimihe

This article appeared in the Meteorology section of ECMWF Newsletter No. 115 – Spring 2008, pp. 8-12.

ECMWF's 4D-Var data assimilation system – the genesis and ten years in operations

Erik Andersson, Jean-Noël Thépaut

One of ECMWF's biggest-ever projects was the development of the variational data assimilation system. The project culminated with the operational implementation of 4D-Var on 25 November 1997. Recently we celebrated ten years of continuous operational production of 4D-Var analyses. In that time the system has matured into a high-performing, efficient algorithm that is capable of combining the information brought by the modern array of terrestrial and satellite observations. The resulting analyses provide the starting point (initial condition) for the ECMWF forecasts. In this short article we look back at the early years of development, and try to answer the question: on what basis was the decision taken to launch such an ambitious project, at the time?

The motivation

The 4D-Var algorithm is fundamentally tightly coupled with the forecast model – see Box A. This means that the software for assimilation, physics and numerics needs to be closely integrated. This was the impetus that initiated the coding of a new forecasting system for ECMWF: the Integrated Forecast System, the IFS. The chain of events was as follows.

- IFS/Arpege development started at the end of 1987.
- The IFS model was introduced on 2 March 1994 (Cy11r7) on the Cray C90 computer.
- 3D-Var became operational on 30 January 1996 (Cy14r3).
- 3D-Var was migrated from CRAY (shared memory) to Fujitsu VPP700 (distributed memory) on 19 September 1996 (Cy15r5).
- 4D-Var became operational on 25 November 1997 (Cy18r1).

We can see that the 4D-Var development work in itself took ten years to complete. A quick look at the plans as they were presented to the Scientific Advisory Committee (SAC) back in 1988 reveals that the amount of work required had initially been vastly underestimated. At that time, the plan spoke of the IFS model becoming operational in 1990, 3D-Var in 1991, and 4D-Var in 1993. Three years later, at the SAC in 1991, 4D-Var's huge hunger for computing power as well as the need for further software, science and algorithmic developments had become more fully understood. The completion date was then more accurately scheduled for 1995/1996.

There were two main motivations to launch this ambitious project. The first was general recognition that initial data (the analysis) is of critical importance for success in forecasting the medium range. However, from about 1986 there had been growing frustration with a couple of aspects of the then operational Optimum Interpolation (OI) assimilation scheme:

- It was difficult to produce a good analysis at the large horizontal scales, as OI by necessity was becoming more and more localized with higher observation density.
- There was only small impact from TOVS satellite data.

Initially (at the SAC meeting in 1986), the variational approach (*LeDimet & Talagrand,* 1986; *Lewis & Derber,* 1985) was seen as an effective means to address the first issue, and it was noted that, in addition, it had the prospect to introduce consistent use of the dynamics in assimilation (*Lorenc,* 1986, 1988). Philippe Courtier and Olivier Talagrand quickly set out to demonstrate these capabilities of 4D-Var with a vorticity equation model in 1986 and a shallow water model in 1987 (published in 1987 and 1990, respectively). These results were so significant that the ECMWF four-year plan in 1987 proclaimed:

"Advances in experimental application of 4D-Var techniques have reached a stage where it has become feasible to develop and evaluate such a scheme with a view to replacing the current intermittent OI scheme with a 4D-Var scheme in the early 1990s."

The SAC commented a little more cautiously that they *"regarded the planned long-term development of variational techniques as an excellent, albeit speculative, research project".*

 doi:10.21957/wnmguimihe

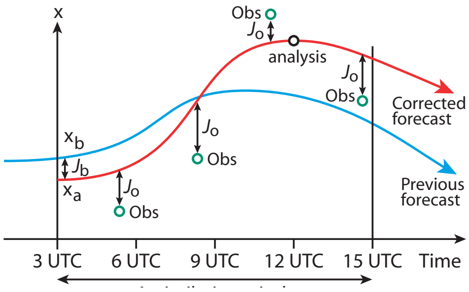
The second motivation, that variational methods would provide a solid foundation for the assimilation of satellite data, was understood one year later. New results from two separate studies had led Jean Pailleux and John Eyre to this realisation: Experimentation with so called 'physical retrievals' at the UK Met Office had demonstrated the benefit of using model fields as first guess for the retrieval of temperature and humidity profiles from the observed radiances. In addition the "PERIDOT" system at Météo-France had attempted to use radiances directly in OI. *"In March 1988 everything on this was clear in our minds"* according to Philippe Courtier. TOVS temperature and humidity retrievals were being used operationally in OI, but reports about small or sometimes negative impact forced ECMWF in May 1991 to withdraw the use of such data from the northern hemisphere and the tropics. An exasperated SAC took comfort in that *"direct use of radiances in a 3D-Var through a direct radiative transfer model and its adjoint is a natural and attractive solution".*

3D-Var and 4D-Var assimilation techniques **A**

4D-Var is a four-dimensional variational data assimilation technique. It performs a statistical interpolation in space and time between a distribution of meteorological observations and an *a priori* estimate of the model state (the background). This is done in such a way that account is taken of the dynamics and physics of the forecast model to ensure the observations are used in a meteorologically consistent way.

The basic idea behind 4D-Var is illustrated by reference to the figure. In the case illustrated here, for a single parameter x the observations are compared with a short-range forecast from a previous analysis over a twelve-hour period. The model state at the initial time is then modified (by adding the analysis increment) to achieve a statistically good compromise x_a between the fit J_b to the previous forecast x_b and the fit J_o to all the observations within the assimilation window. J_b and J_o are referred to as cost functions.

The observations are compared to model fields at the correct time, but in 3D-Var the forecast model is not involved in propagating the analysis increment to the observation times. The 3D-Var analysis increment thus refers to the central time of the interval and it does not evolve. The 3D-Var algorithm is therefore computationally less demanding and less accurate than that for 4D-Var.



Towards an operational system

The group of scientists involved in coding the IFS had gradually grown from the initial two pioneers (Mats Hamrud and Philippe Courtier) to about ten. David Burridge (who was Head of Research) and Jean Pailleux (Head of Data Assimilation Section) had been enthusiastic supporters right from the outset. Tony Hollingsworth (Head of Data Division) had been a little more reluctant at first but became deeply involved following a set of convincing results obtained in 1991–1992 (e.g. *Rabier et al.,* 1993; *Thépaut et al.,* 1993; *Andersson et al.,* 1994). These studies had demonstrated that 4D-Var could (a) generate well-balanced fields, (b) create flow-dependent, vertically sloping corrections from single-level observations, and (c) induce wind-field information from a time-sequence of humidity-sensitive satellite radiance observations. The 'variational team' of developers were now busy building all the elements needed for a full-scale 4D-Var which included:

- A forecast model and its adjoint.
- The observation operators linking the observed variables to the model quantities; code to compute the observation cost function J_o and its gradient.
- The first-guess operator, to incorporate information from recent analyses; code to compute the first-guess cost function J_b and its gradient.
- Balance operators to ensure the appropriate relationship between mass and wind.
- General minimisation algorithm, to seek the analysis as the minimum of the cost function $J_o + J_b$.
- A suitable solution algorithm that can take advantage of the computing power available on multi-processor computing platforms.

doi:10.21957/wnmguimihe

The computing cost of 4D-Var was always a concern. It became clear that it would be prohibitively expensive, even taking into account the planned computer upgrade in 1996, to solve the full system. It was clear that significant cost-saving devices had to be developed. The breakthrough came by adopting the so-called incremental approach (published in 1994 by Courtier et al. who was influenced by the approach of John Derber from NOAA/NCEP, Washington), which enabled very significant trade-off opportunities between cost-savings and accuracy. Once the Fujitsu computer became available in September 1996, the performance of incremental 4D-Var was thoroughly evaluated and was soon shown to outperform 3D-Var (*Rabier et al.,* 1997; 1998a).

Before this, the variational system in the form of 3D-Var needed to prove its worth with respect to the operational, and by this stage finely tuned, OI scheme. At first there was no immediate time pressure, but as the arrival of the Fujitsu computer was drawing nearer, a decision had to be made whether to (a) invest considerable manpower to adapt the OI codes to this distributed memory parallel computer or (b) rely on the already parallelized 3D-Var codes and take a bet on its eventual performance. It was decided not to migrate OI. Shortly thereafter it was (fortunately) possible to show that 3D-Var could outperform OI by using additional new data such as scatterometer near-surface winds and satellite radiance data for the first time, and by producing less noisy analyses (*Courtier et al.,* 1998; *Rabier et al.,* 1998b; *Andersson et al.,* 1998). Having ensured these results, a great deal of technical work still remained to facilitate the operational implementation of 4D-Var. This included developing efficient observation processing, code parallelisation, optimisation of codes and algorithms and design of UNIX scripts in close collaboration between the Research and Operations Departments.

In *ECMWF Newsletter No. 78* there was an article by François Bouttier and Florence Rabier to proudly announce the successful implementation of 4D-Var on 25 November 1997. Here are some quotations from that article.

"This is the first ever operational application of the 4D-Var technique successfully applied to a high-resolution assimilation and forecast system."

"This was made possible by more than 10 years of scientific and technical developments in and around ECMWF's IFS as well as the availability of a powerful new computer system organized around a Fujitsu VPP700 with 116 processors."

"This is an impressive yet young assimilation system that offers an exceptional scope for future improvements."

In the article some results were presented from the extensive pre-operational validation of 4D-Var. Figure 1 shows the impact on the 500 hPa geopotential height of going from 3D-Var to 4D-Var with everything else being kept identical. These results show that there is an improvement in forecast accuracy when using 4D-Var to provide the initial conditions (analyses) for the forecasts.

The promise of 'exceptional scope for future improvements' was not an exaggeration. The list of accomplishments that have actually materialized in the past ten years is too long and varied to summarize here. However, the articles that have appeared in the *ECMWF Newsletter* give a clear indication of the way 4D-Var has developed – see Box B.

A visual impression of the progress can be seen by comparing the plots shown in Figure 2. It shows the magnitude of analysis increments, i.e. the amount of work done by the assimilation system, in terms of 500 hPa height for the month of October in four different years: 1994 (OI), 1997 (3D-Var), 1998 (4D-Var) and 2007 (recent 4D-Var). The panels show a clear reduction in the magnitude of increments, which can be interpreted as a reduction in errors (increased accuracy) in the data assimilation process. This reflects improvements in model accuracy, assimilation methods, and the vast increase of available satellite data.

What next?

Satellite data with high density and increasing computer power make it possible to explore the benefits of higher-resolution assimilation on a global scale. The performance of T1279 (15 km) assimilation with analysis increments at up to T399 (50 km) resolution will be assessed within the coming year.

The potential benefits of longer-window 4D-Var assimilation, beyond the current 12 hours, are being evaluated within a simulation test-bed with a quasi-geostrophic model. This tests the performance of an extension to 4D-Var that accounts for model error – the so-called weak-constraint 4D-Var.

The use of data assimilation ensembles is another important area for current research. It is planned that in the near future this will facilitate every analysis field being provided with a reliable estimate of its accuracy (uncertainty). Apart from its inherent value, this will also be valuable as an input to the Ensemble Prediction System (EPS) whose aim it is to predict the reliability of today's forecast. It is envisaged that 5 to 10 lower-resolution 4D-Var assimilations will be run in parallel, each one with differently perturbed observation inputs.

 doi:10.21957/wnmguimihe

The information from the ensemble will also be used in the assimilation system itself to give varying weights to observations depending on today's situation. This will be particularly beneficial for the analysis of intense and extreme weather events.

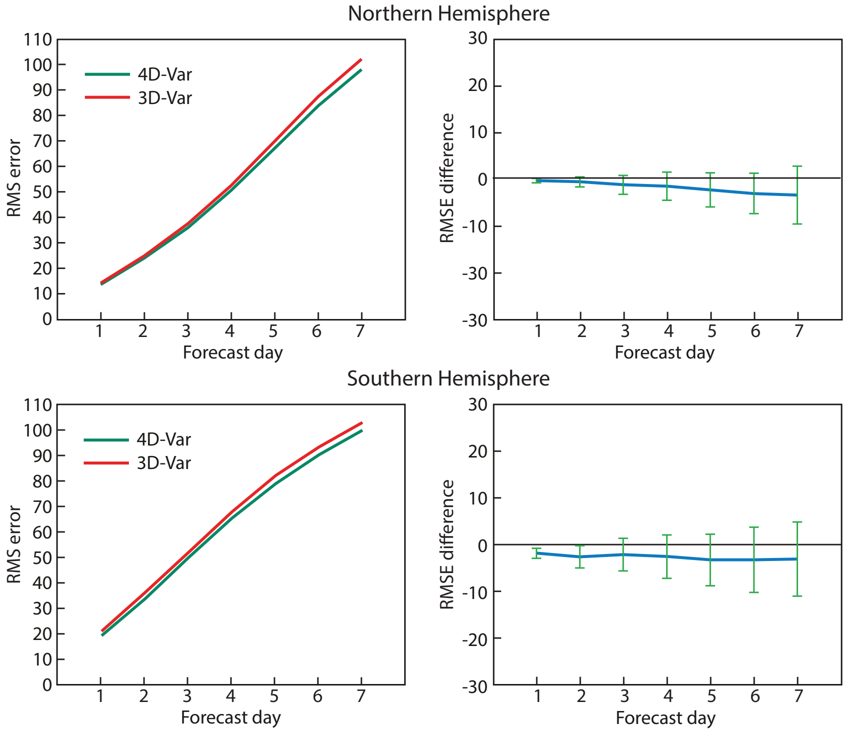


Figure 1 Comparison of RMS error for a set of forecasts started from 3D-Var and 4D-Var analyses (left) and the corresponding difference in the RMS error between the two sets of forecasts with 95% confidence intervals (right) for the northern hemisphere extratropics (top) and southern hemisphere extratropics (bottom). (From *Bouttier & Rabier,* 1998, *ECMWF Newsletter No. 78*).

Key articles about variational data assimilation techniques and data impact that have appeared in the ECMWF Newsletter **B**

- Operational implementation of 4D-Var, *Winter 1997/98, 78, 2–5*
- Recent improvements to 4D-Var, *Autumn 1998, 81, 2–7*
- Raw TOVS/ATOVS radiances in the 4D-Var system, *Spring 1999, 83, 2–7*
- Assimilation of meteorological data for commercial aircraft, *Autumn 2002, 95, 9–14*
- Assimilation of high-resolution satellite data, *Spring 2003, 97, 6–12*
- ERA-40: ECMWF's 45-year reanalysis of the global atmosphere and surface conditions 1957–2002, *Summer/Autumn 2004, 101, 2–21*
- CO2 from space: estimating atmospheric CO2 within the ECMWF data assimilation system, *Summer 2005, 104, 14–18*
- New observations in the ECMWF assimilation system: satellite limb measurements, *Autumn 2005, 105, 13–17*
- "Wavelet Jb" – A new way to model the statistics of background errors, *Winter 2005/06, 106, 23–28*
- A variational approach to satellite bias correction, *Spring 2006, 107, 18–23*
- Analysis and forecast impact of humidity observations, *Autumn 2006, 109, 11–15*
- Assimilation of cloud and rain observations from space, *Winter 2006/07, 110, 12–19*
- Operational assimilation of GPS radio occultation measurements at ECMWF, *Spring 2007, 111, 6–11*
- Evaluation of the impact of the space component of the Global Observing System through Observing System Experiments, *Autumn 2007, 113, 16–28*

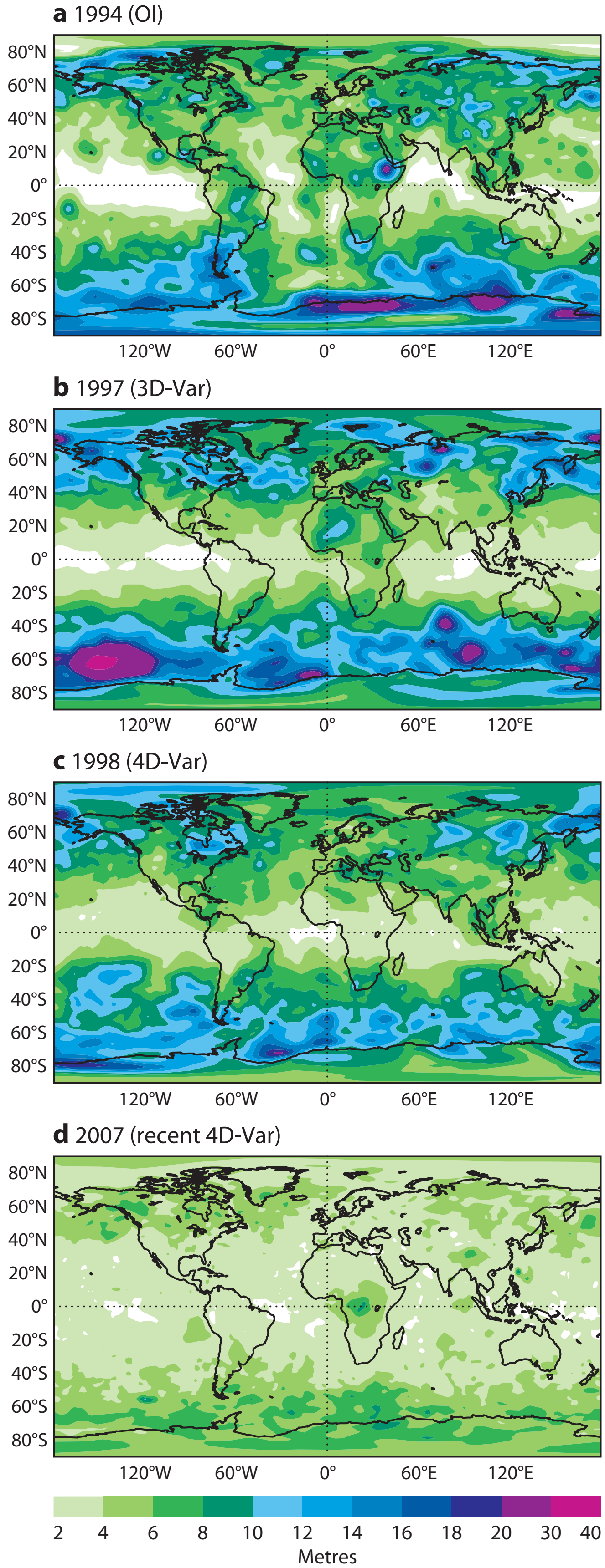


Figure 2 RMS of analysis increments in terms of 500 hPa geopotential height in ECMWF's operational assimilation system for the month of October for (a) 1994 (OI), (b) 1997 (3D-Var), (c) 1998 (4D-Var) and (d) 2007 (recent 4D-Var). Where observations are available, these results reflect the assimilation accuracy which depends on the accuracy of the observations, the forecast model and the assimilation system.

doi:10.21957/wnmguimihe

Further Reading

Andersson, E., J. Pailleux, J.-N. Thépaut, J. Eyre, A.P. McNally, G. Kelly & **P. Courtier,** 1994: Use of cloud-cleared radiances in three/four-dimensional variational data assimilation. *Q. J. R. Meteorol. Soc.,* **120,** 627–653.

Andersson, E., J. Haseler, P. Undén, P. Courtier, G. Kelly, D. Vasiljevic´, C. Brankovic´, C. Cardinali, C. Gaffard, A. Hollingsworth, C. Jakob, P. Janssen, E. Klinker, A. Lanzinger, M. Miller, F. Rabier, A. Simmons, B. Strauss, J.-N. Thépaut & **P. Viterbo,** 1998: The ECMWF implementation of three dimensional variational assimilation (3D-Var). III: Experimental results. *Q. J. R. Meteorol. Soc.,* **124,** 1831–1860.

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,** 1367–1388.

Courtier, P. & **O. Talagrand,** 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. II: Numerical results. *Q. J. R. Meteorol. Soc.,* **113,** 1329–1347.

Courtier, P. & **O. Talagrand,** 1990: Variational assimilation of meteorological observations with the direct and adjoint shallow-water equations. *Tellus,* **42,** 531–549.

Courtier, P., E. Andersson, W. Heckley, J. Pailleux, D. Vasiljevic´, M. Hamrud, A. Hollingsworth, F. Rabier & **M. Fisher,** 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Q. J. R. Meteorol. Soc.,* **124,** 1783–1807.

LeDimet, F. & **O. Talagrand,** 1986. Variational algorithms for analysis and assimilation of meteorological observations. Tellus, 38A, 97–110.

Lewis, J.M. & **J.C. Derber,** 1985: The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus,* **37,** 309–327.

Lorenc, A.C., 1986: Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.,* **112,** 1177–1194.

Lorenc, A.C., 1988: Optimal nonlinear objective analysis. *Q. J. R. Meteorol. Soc.,* **114,** 205–240.

Rabier, F.P. Courtier, J. Pailleux, O. Talagrand & **D. Vasiljevic´,** 1993: A comparison between four-dimensional variational assimilation and simplified sequential assimilation relying on three-dimensional variational analysis. *Q. J. R. Meteorol. Soc.,* **119,** 845–880.

Rabier, F., J-F. Mahfouf, M. Fisher, H. Järvinen, A.J. Simmons, E. Andersson, F. Bouttier, P. Courtier, M. Hamrud, J. Haseler, A. Hollingsworth, L. Isaksen, E. Klinker, S. Saarinen, C. Temperton, J.-N. Thépaut, P. Undén & **D. Vasiljevic´,** 1997: Recent experimentation on 4D-Var and first results from a Simplified Kalman Filter. *ECMWF Tech. Memo No. 240.*

Rabier, F., J.-N. Thépaut & **P. Courtier,** 1998a: Extended assimilation and forecast experiments with a four-dimensional variational assimilation system. *Q. J. R. Meteorol. Soc.,* **124,** 1861–1887.

Rabier, F., A. McNally, E. Andersson, P. Courtier, P. Undén, J. Eyre, A. Hollingsworth & **F. Bouttier,** 1998b: The ECMWF implementation of three dimensional variational assimilation (3D-Var). II: Structure functions. *Q. J. R. Meteorol. Soc.,* **124,** 1809–1829.

Talagrand, O. & **P. Courtier,** 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Q. J. R. Meteorol. Soc.,* **113,** 1311–1328.

Thépaut, J.-N., R. Hoffman & **P. Courtier,** 1993: Interactions of dynamics and observations in a four-dimensional variational assimilation. *Mon. Wea. Rev.,* **121,** 3393–3414.