# Ideas for adding flow-dependence to the Met Office VAR system


**Andrew Lorenc**

*Met Office, Exeter, UK*
andrew.lorenc@metoffice.gov.uk

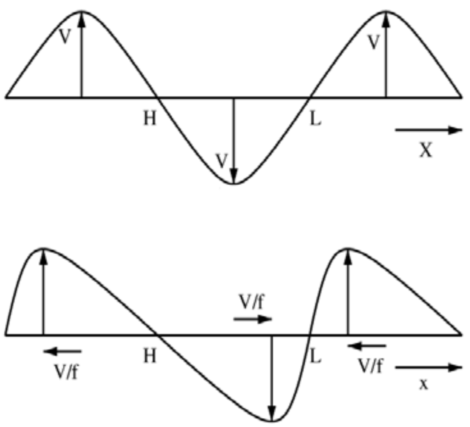

## 1. Introduction

The Met Office started developing variational assimilation (VAR) in 1993, implemented global 3D-Var in 1999 (Lorenc et al. 2000) and global 4D-Var in October 2004 (Rawlins et al. 2007), with implementations in the regional NWP system soon following. Our convective-scale NWP uses 3D-Var – 4D-Var will not be affordable for some years. In section 2 I review several attempts in this period to introduce flow-dependence. As this is a workshop paper, in section 3 I present some comments and speculations, and in section 4 I outline plans.

## 2. Ideas tried

### 2.1. Geostrophic Co-ordinate Transform

Following a suggestion by Desrozier (1997), a horizontal geostrophic coordinate transform was developed for our global VAR by Mark Dubal. The transform is a mapping, on model levels, of the VAR control variables, into geostrophic coordinates in which, it is hoped, error correlations will be more uniform. Figure 1 shows an idealized one-dimensional example.




$$x_G = x_R + \frac{1}{f} v_g$$

$$y_G = y_R - \frac{1}{f} u_g$$

*Figure 1 (TOP) Velocity pattern V appears sinusoidal in geostrophic, X, space. There is an anticyclone, centre H and a cyclone, centre L. It is assumed that errors will be similarly homogeneous in X-space. (BOTTOM) When transformed to real space, x, the anticyclone appears broader and weaker while the cyclone becomes tighter and stronger; the assumed error patterns follow this (Mark Dubal).*

In our VAR system it is implemented as an extra step in the U-transform (Lorenc et al. 2000) which maps control variable into model increments. It consists of a horizontal translation (similar in form to a semi-Lagrangian advection), after the horizontal and vertical transforms which define spatial covariances and before the parameter transform into model variables.

Trials of this scheme showed disappointing results, with impacts generally small; worse than 3D-Var alone in the northern hemisphere and better in the southern hemisphere. Possibly this result is due to approximations in our implementation: the necessary interpolations introduced additional smoothing, we used a smoothed rotational (rather than geostrophic) linearization-state wind to define the transform, on model levels (rather than isentropic surfaces). The latter choice was to ensure we had defined, regular grids in transformed space (for spatial transforms) and model space. This desire also makes the design of a limited-area version difficult.

Following a further trial (Semple 2001, and Figure 2 below) the approach was shelved in favour of error modes of the day (below).

## 2.2. Errors modes of the day

In parallel with the work above, Dale Barker extend the VAR system to use "error modes of the day" as additional degrees of freedom to fit the observed data. In order to allow some spatial variation in the use of the modes, each has a (horizontally) varying weight α which is determined variationally, at the same time as the traditional control variables. The idea was a development of Kalnay and Toth (1994), who used a separate pre-analysis step. Initial Met Office tests of the flow-dependent error structures used in this way used an error breeding system to provide a single "error mode" (Semple 2001, 2003). Figure 2 is from Semple (2001). It shows the 3D-Var response to a single observed temperature increment of 1K in a baroclinic zone. The basic 3D-Var increment field is not flow-dependent, whereas both the GC-transform and the bred modes cause increment patterns which follow the isentropic structure. These tests showed encouraging behaviour in some cases such as that shown, but overall neutral results. Considering the effort needed to develop and maintain the error breeding system, we decided in 2004 to shelve the development pending the availability of the Met Office's own ensemble system (MOGREPS), which is being developed using an ensemble transform Kalman filter (ETKF). Meanwhile the work has been continued by Dale Barker and collaborators at NCAR.

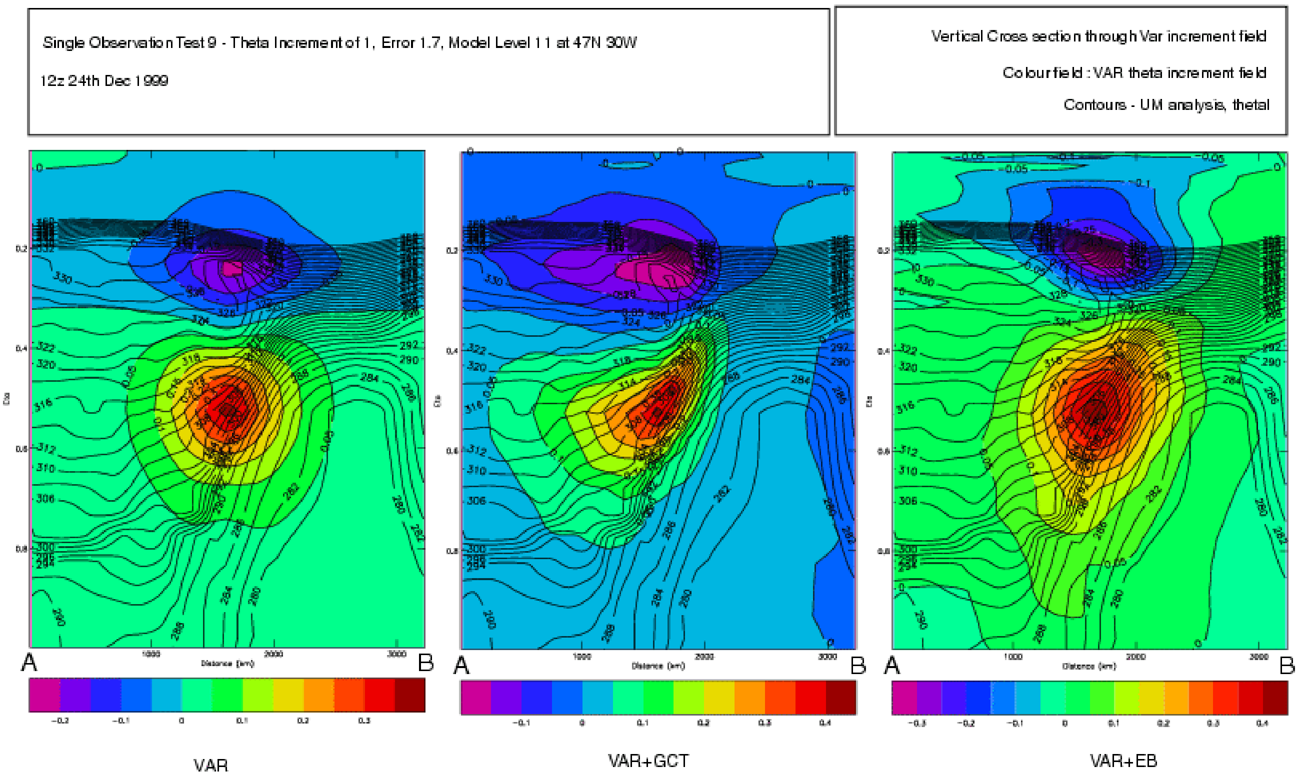



*Figure 2. The effect of the error breeding and geostrophic coordinate transform on a single observational increment of 1K inserted on model level 11, placed on a frontal zone. The basic 3D-Var increment (left) ignores the frontal structure completely, whereas the increments from 3D-Var + Geostrophic Coordinate Transform (centre) and 3D-Var + bred modes (right) shows the observation 'flowing up' the theta surface. (Semple 2001)*

Ensemble perturbations are used in a similar way to represent background error structures in ensemble Kalman filters, and the spatially varying α control variable performs the same role as localisation in such filters, allowing reasonable results with only a few members. This was pointed out by Lorenc (2003) and the result was further explored by Wang *et al.* (2007).

In the Met office implementation, the Schur product of the error mode with an α control variable field is performed in transformed parameter space. This ensures that the localisation introduced by the (smooth) variations in α do not introduce imbalance. The design allows for multiple modes, but only one has been tested. It was found that care was needed in the pre-conditioning of α. It is also desirable to reduce the weight given to the traditional control variables. Barker and collaborators have been working in this area.

## 2.3. 4D-Var

One of the often quoted advantages of 4D-Var is its flow dependence; the main reason that the work above has made slow progress has been the priority given instead to development and implementation of 4D-Var. Lorenc and Rawlins (2005) reported on a clean inter-comparison of 3D-Var and 4D-Var in a full NWP system. By inventing a new "synoptic 4D-Var" method where observation increments are all treated in 4D-Var as if they were in the middle of the time-window (Figure 3), they were able to provide a clean test of the value of the evolved flow-dependent covariances in 4D-Var compared to those used in 3D-Var with FGAT. This had a larger impact on the fields than the introduction of FGAT in 3D-Var, or the correct treatment of observation times in 4D-Var. In statistics from a one month trials these differences were on average (but not always) positive (Figure 4).

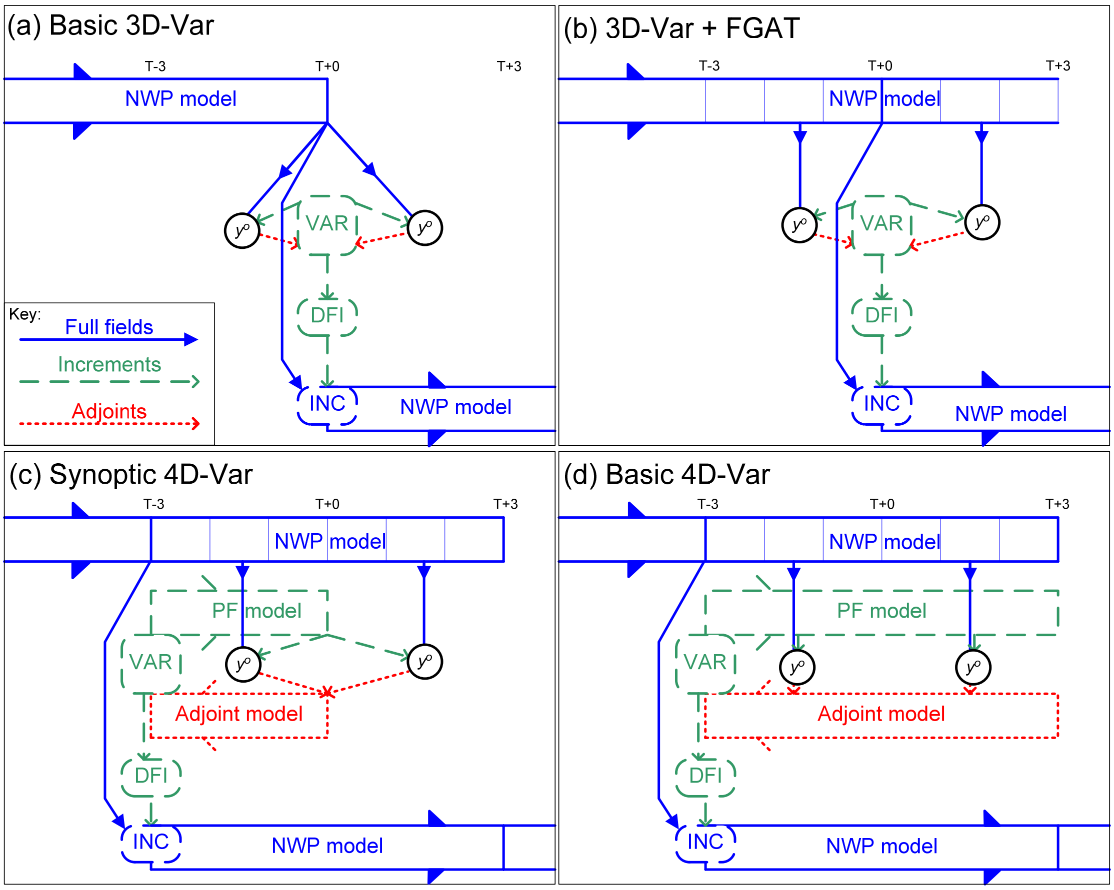



*Figure 3 The four variational assimilation schemes used. Time goes from left to right. Two observations are shown as* $y^o$*. VAR is the incremental variational minimisation, DFI is the initialisation of increments by digital filtering, and INC is the addition of increments to the full model.*

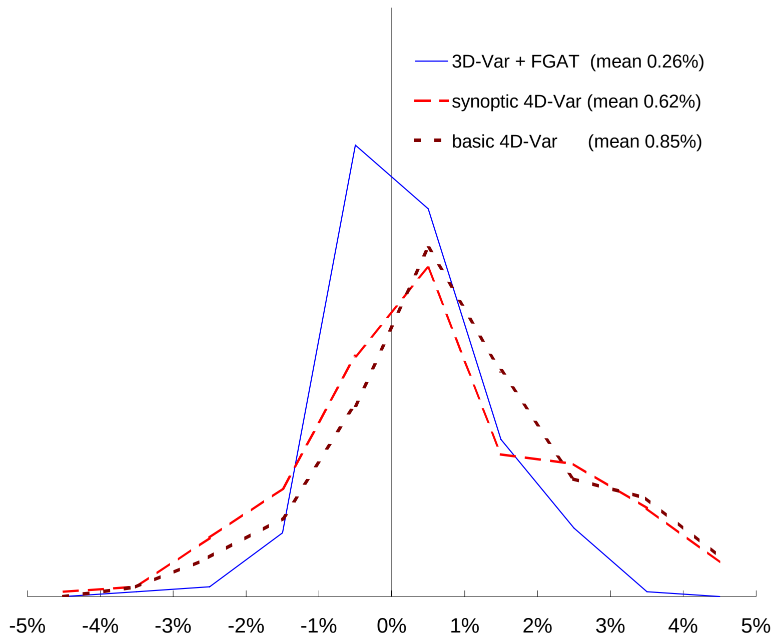



*Figure 4 Histograms of the reduction in r.m.s. verification scores against observations, relative to basic 3D-Var, for 234 fields+levels+forecast-lengths. The legend also shows the mean improvements.*

## 2.4. Cloud-topped inversions

Probably the commonest cause of forecast error in the UK is the misrepresentation of inversions and strato-cumulus layers. Such cases can have very high impact, for instance in December 2006 3 days of poor visibility at Heathrow caused over a thousand flights to be cancelled, disrupting hundreds of thousands of Christmas travel plans.

Data assimilation in such situations requires very specific flow-dependent methods. For many years we have had some success using the MOPS assimilation system (Macpherson *et al.* 1996). Satellite and surface observations are pre-processed to give a 3-dimensional cloud analysis. Included in the processing are checks that layer cloud is aligned with the forecast inversion. The model's relative humidity is then nudged over several timesteps to make the model's cloud fit the MOPS analysis. There is no vertical spreading of information. This system is awkward to combine with 4D-Var, but the combination still gives significant benefits. So Richard Renshaw has been testing the direct assimilation of the MOPS cloud analysis in VAR.

To illustrate the flow-dependent nature of errors in such situations, Figure 5 shows observation minus model (background) correlations, from global radiosondes, composited by the level in which a layer cloud top was diagnosed the background. (Strictly these are not background error correlations, since observation errors are included and mean errors are not subtracted, but they give a good understanding of errors nevertheless.) The red line is the correlation from all sondes, and the dotted line is the assumed VAR correlation (**B**). The **B** temperature correlations with level 4 are clearly too broad (left), while those for RH match that for all sondes, but are too broad in cases with layer cloud (centre). In particular, correlations across the inversion are much less than in **B**.

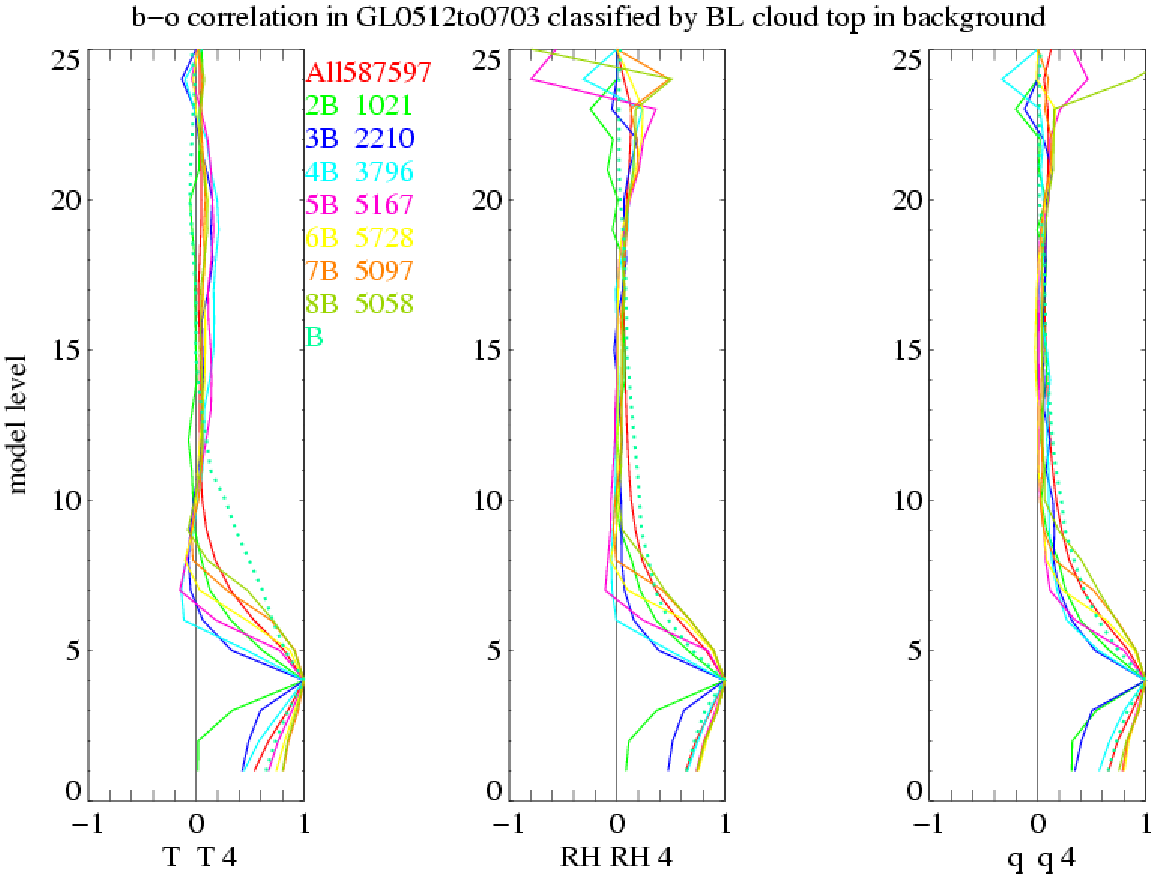



*Figure 5 Vertical correlation of o-b differences between 15 months of radiosondes and the 6hr background forecast. Correlations of T, RH and q are shown with model level 4 (320m above ground). The numbers in each class are shown in the legend, for instance the light blue curves show correlations from 3796 cases where the background had a diagnosed cloud top at level 4. The dotted line sows the assumed VAR correlation.*

Using **B** correlations such as that shown dotted in the middle graph of Figure 5, it is not surprising that the increments due to cloud data are spread too far in the vertical by VAR. This is illustrated in Figure 6 right, with the increments from the old, operational, cloud nudging scheme on the left. Forecasts from the cloud in VAR assimilation were significantly worse.

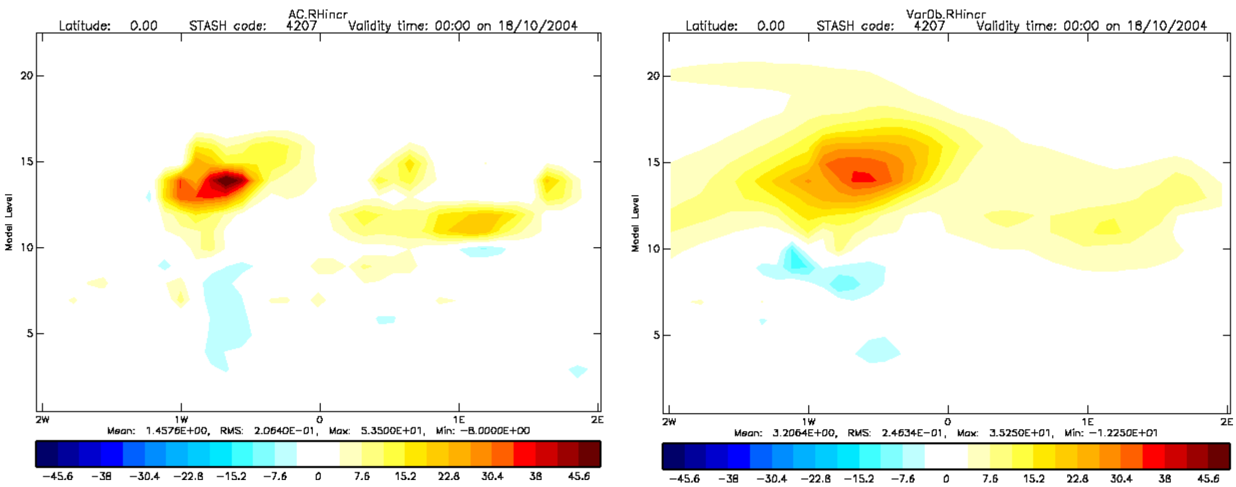



*Figure 6 RH increments from the assimilation of MOPS cloud analysis in the mesoscale model: LEFT from the AC nudging scheme, RIGHT from 3D-Var. (Richard Renshaw)*

It is not easy to modify the vertical covariance in VAR without damaging results in other cases, so a more pragmatic solution was found specifically for cloud data. In order to reduce the extrapolation of increments into cloud-free levels, the background value is inserted as a "pseudo-observation" at all levels observed to be cloud free. (No pseudo-observations are generated at unobserved levels.) This is illustrated schematically in Figure 7.

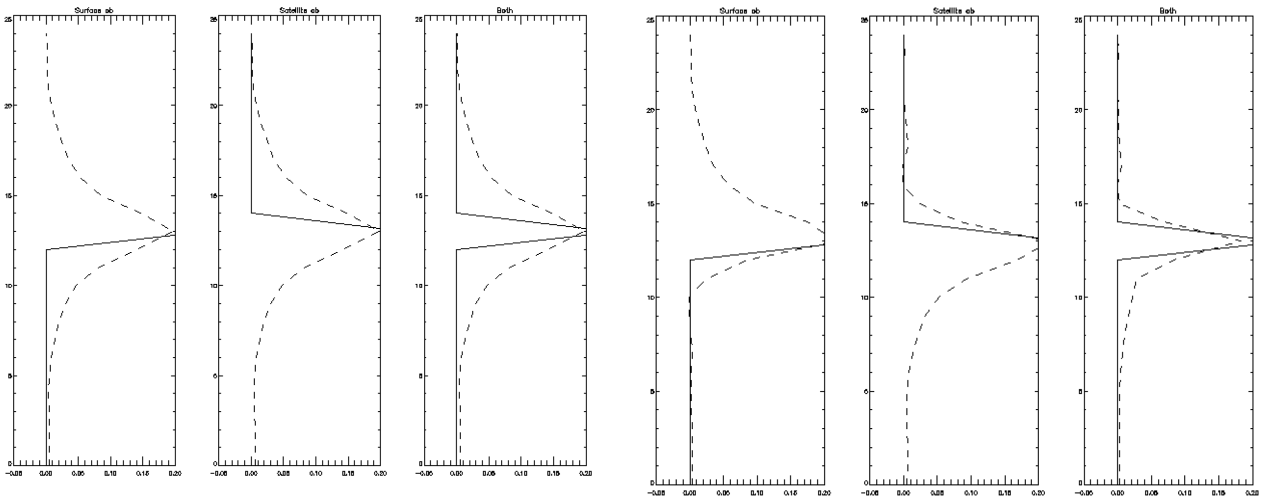


*Figure 7 The left 3 graphs show the effect of standard assimilation of a cloud layer observed from below, above and both. The dotted line shows the VAR increment – the same in all 3 cases. The right 3 graphs show the same cases in the new scheme. Where cloud is observed to be absent, the background (solid line) value is used as a pseudo-observation. This reduces the spreading of increments from the cloud observation into layers observed to be cloud-free, as shown by the dotted lines. (Richard Renshaw)*

This revised cloud in VAR scheme now matches the performance of the operational VAR plus cloud nudging, both beating runs with no use of cloud data. This is naturally most evident in cloud amount forecasts. Figure 8 illustrates this with scores from a mesoscale model 3D-Var trial. The cloud in VAR results are better than the old nudging scheme for the first 24 hours.

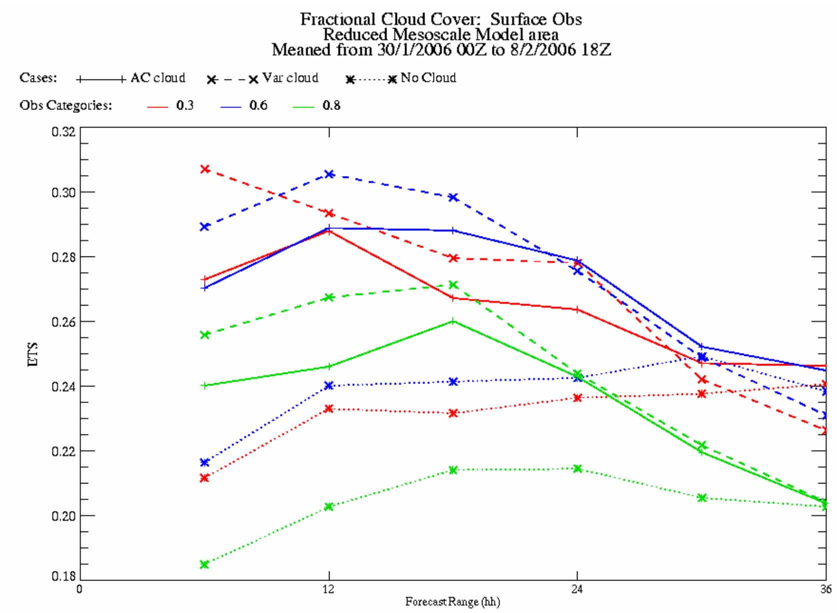



*Figure 8. Equitable Treat Scores, verified against surface observation, for three levels of cloud cover: 0.3 (red), 0.6 (blue) and 0.8 (green), from 3D-Var UK mesocale model forecasts with cloud nudging (solid), cloud in VAR (dashed) and no cloud (dotted). (Richard Renshaw)*

Preliminary results with 4D-Var are also encouraging; not only are the cloud forecasts still improved, but this is achieved by increments to fields other than humidity (Figure 9). It seems that 4D-Var is adjusting the dynamics to fit the cloud data.

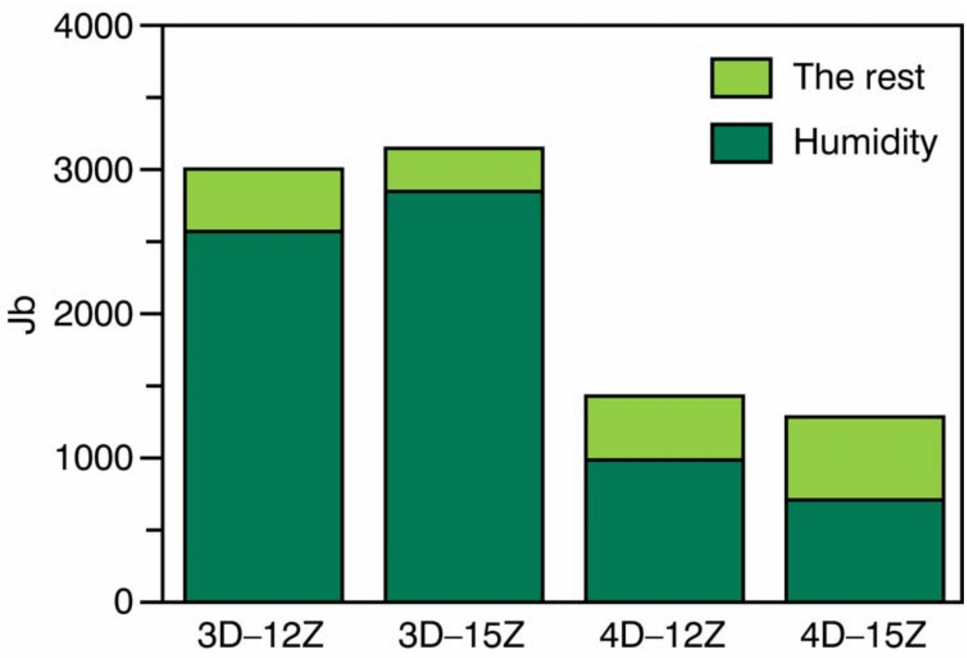



*Figure 9 Relative size of humidity and other increments. Each column shows the size of Jb at the end of the VAR minimisation. This measures the size of the analysis increments, using as norm* $B^{-1}$*, where B is the background error covariance. The lower (dark green) portion shows the humidity increment, the upper light green portion all other variables. The left two columns show results from 3D-Var for 12Z and 15Z, the right two columns results from 4D-Var. (Richard Renshaw)*

# 3. Comments

## 3.1. Get the basics right first

Our experience is that VAR results are still very sensitive to changes in the covariance model: changes which are smaller than inaccuracies which we believe to be present. Testing of augmented flow-dependent methods in these circumstances is difficult; impact might be due to bye-product changes in the mean statistics rather than true flow dependence.

## 3.2. Analysis Resolution

Some forecaster complaints, that the data assimilation has not fitted observations, do not take account of the limitations of model resolution. One way to reduce misperception is to display observations filtered to the model grid-length. An example is in Figure 10, from the set of global soundings with cloud top at level 4 used in Figure 5. However even displayed like this, although drawing towards the observation at most levels, the assimilation fails to correct the observed inversion level. In other examples (e.g. Semple 2006) the covariances are blamed for the assimilation not fitting observed detail. In assimilation schemes like VAR which generate their covariances from forecast differences (from the NMC method) or ensemble perturbations, it is only possible to fit structures which are (common) in the learning set of perturbations. Thus it is not the grid-length which determines the best possible increment resolution, but the scale of structures spontaneously generated by the model.

### 3.3. Even perfect covariances are not sufficient

Where there are physical structure like inversions, or fronts, or cyclones, or convective cells, which tend to have position errors but otherwise similar structures, then error distributions in normal coordinates are non-

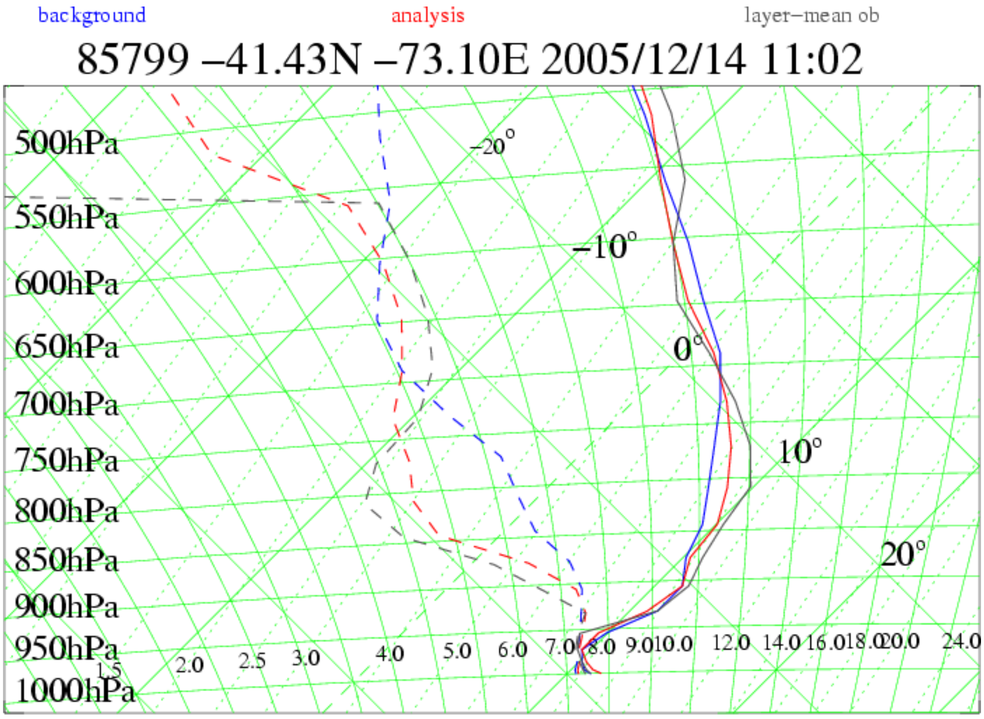



*Figure 10 Typical tephigram sounding from the set used to calculate the correlations shown in figure 5.*

Gaussian. Most of our theory of VAR, which relies on finding a least-squares best fit, using covariances to characterize PDFs, breaks down. This was discussed for convective scales in Lorenc and Payne (2007). In our example of section 2.4, one symptom is that the "best estimate" mean of the PDF, given a particular background state, is not the background state! If the background state has a sharp inversion, the "best" estimate smoothes it. So for our composites of cases with background inversion at a particular level these is a systematic bias, even though the model over all cases is nearly unbiased (Figure 11). If we composite by inversion level in the sounding rather than the background, the biases have opposite sign (not shown). Although no doubt our current covariances could be improved to handle such cases, it may well be that our linear, covariance-based methods are fundamentally flawed, so that even the best possible covariances would not give ideal results.

## 4. Plans

Because of the sentiments expressed in section 3.1 our firm plans for flow-dependent developments are few:

Evaluate the MOGREPS ensemble to see whether its perturbations at T+6 are consistent with what we know about background error covariances. If necessary improve the ensemble generation (localised ETKF, with model perturbations). Generate improved covariances (replacing the NMC method). Only then go on to look at errors of the day as described in section 2.2.

Implement the "MOPS in VAR" scheme of section 2.4 in operation regional 4D-Var and UK 3D-Var schemes, replacing the MOPS cloud nudging. Extend the scheme for other 1D-Var cloud retrievals from IR satellite sounders.

Other ideas are not yet fully planned or resourced:

Revise our covariance model to allow more flexible horizontal variation of statistics, probably using a wavelet based approach (Bannister 2007). Use the spread of the MOGREPS ensemble to modulate the average variances.

Institute a vertical transform as well as, or instead of the horizontal GC transform (section 2.1). The grid should be chosen by an optimisation process constrained to keep the same vertical domain, but within that domain to seek to equalise the spacing in isentropic coordinates.

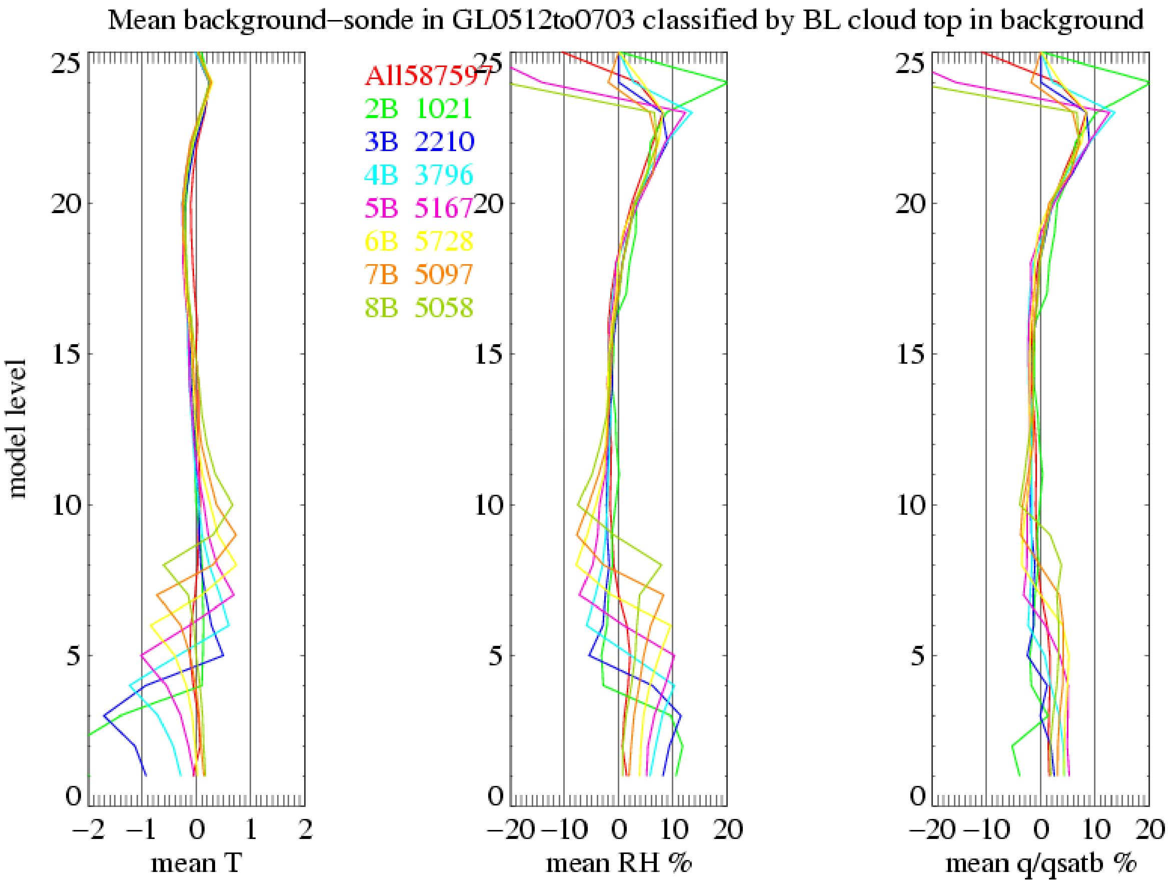



*Figure 11 As Figure 5 for mean background minus observation.*

## References


Bannister, Ross N. 2007: Can wavelets improve the representation of forecast error covariances in variational data assimilation? *Mon. Wea. Rev.*, **135**, 387-408.

Desroziers, Gerard 1997: A coordinate change for data assimilation in spherical geometry of frontal structures *Mon. Wea. Rev.*, **125**, 3030-3039

Kalnay, E., and Z. Toth, 1994: Removing growing errors in the analysis. *Preprints, 10th Conf. on Numerical Weather Prediction, Portland, OR, Amer. Meteor. Soc.*, 212–215.

Lorenc, A. C., S. P. Ballard, R. S. Bell, N. B. Ingleby, P. L. F. Andrews, D. M. Barker, J. R. Bray, A. M. Clayton, T. Dalby, D. Li, T. J. Payne and F. W. Saunders. 2000: The Met. Office Global 3-Dimensional Variational Data Assimilation Scheme. *Quart. J. Roy. Met. Soc.*, **126**, 2991-3012.

Lorenc, A. C., 2003: The potential of the Ensemble Kalman filter for NWP - a comparison with 4D-Var. *Quart. J. Roy. Met. Soc.,* **129**, 3183-3203.

Lorenc, Andrew C. and F. Rawlins 2005: Why does 4D-Var beat 3D-Var? *Quart. J. Roy. Met. Soc.*, **131**, 3247-3257.

Lorenc, Andrew C. and Tim Payne 2007: 4D-Var and the Butterfly Effect: Statistical four-dimensional data assimilation for a wide range of scales. *Quart. J. Roy. Met. Soc.*,**133**, 607-614.

Macpherson, B., Bruce J. Wright, William H. Hand, and Adam J. Maycock, 1996: The impact of MOPS moisture data in the UK Meteorological Office mesoscale data assimilation scheme. *Mon. Wea. Rev.*, **124**, 1746-1766

Rawlins, F., S. P. Ballard, K. J. Bovis, A. M. Clayton, DingMin Li, G. W. Inverarity, A. C. Lorenc and T. J. Payne 2007: The Met Ofiice Global 4-Dimensional Data Assimilation System. *Quart. J. Roy. Met. Soc.*, **133**, 347-362.

Semple, Adrian T. 2001: A Meteorological Assessment of the Geostrophic Co-ordinate Transform and Error Breeding System when used in 3D-Var. Met Office FRTR357.

Semple, Adrian T. 2003: An Error Breeding System for the Met Office 'New Dynamics'. Met Office FRTR413.

http://www.metoffice.gov.uk/research/nwp/publications/papers/technical_reports/2003/FRTR413/FRTR413.pdf

Semple, Adrian T. 2006: Impact of the Aberporth Radiosonde Ascent on a Waving Frontal Zone, 18th April 2006. Met Office FTTR490.

http://www.metoffice.gov.uk/research/nwp/publications/papers/technical_reports/2006/FRTR490/FRTR490.pdf

Wang, Xuguang, Chris Snyder and Thomas M. Hamill, 2007: On the Theoretical Equivalence of Differently Proposed Ensemble–3DVAR Hybrid Analysis Schemes. *Mon. Wea. Rev.*, **135**, 222-227.