TECHNICAL MEMORANDUM

680

Flow-dependent, geographically varying background error covariances for 1D-VAR applications in MTG-IRS L2 Processing

Elías Valur Hólm and Tomas Kral

Research Department

December 20, 2012

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen terme

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/publications/

Contact: library@ecmwf.int

Abstract

Satellite retrievals and other 1-dimensional variational data assimilation (1D-VAR) applications do need background error statistics of the atmospheric variables used, in particular temperature, water vapour and ozone. These error statistics are often provided as single global climatological profiles. In this report we describe a self-contained program package that gives more accurate error statistics that closely follow the errors used in the ECMWF analysis system. It has been possible to make the 1D-VAR errors simpler than those used by the 3D ECMWF analysis because horizontal correlations and wind errors are not needed.

The error statistics consist of vertical error correlation matrices and three-dimensional error variances. In addition, a variable transform converts the model variables (temperature and mixing ratios of water vapour and ozone) to variables with more Gaussian and less correlated error statistics. Gaussian error statistics make the background errors more robust and removing cross-variable correlations simplifies the vertical correlation matrices.

The error correlation matrices are derived from ECMWF ensemble data assimilation statistics, and are an average over a couple of seasons. Each variable has a separate set of vertical error correlation matrices at approximately 625 km by 625 km resolution, equally distributed all over the globe. Investigation of the geographic, diurnal, and seasonal variability of the correlation matrices indicate that for averages over several days it is only the geographic variation that matters, with very little variation between seasons or between different times of day. The error correlations are provided as a constant file which is only updated infrequently, following major changes in the ECMWF system.

The background error standard deviations are also derived from the ECMWF ensemble data assimilation system, and are the same as those used by the ECMWF analysis. They are retrieved from ECMWF as 3-dimensional GRIB fields for the day and time required (currently available at 09 and 21 UTZ). The resolution is a T159/N80 reduced Gaussian grid with approximately 125 km equal resolution globally.

The variable transform, which in particular removes the variable inter-correlation between water vapour and temperature, requires the full fields of temperature, water vapour, and logarithm of surface pressure, and these are also retrieved from ECMWF for the day and time required.

The interface with the 1D-VAR satellite retrieval application interpolates the background errors bilinearly to the retrieval location, which includes the interpolation of model fields and error standard deviations as well as correlation matrices.

The program package that calculates the background errors is self-contained, including documentation. It is maintained and regularly updated by ECMWF to reflect major model and resolution upgrades. Following major upgrades, new correlation matrices are needed, and these will be available from ECMWF as part of the latest version of the software.

1 Introduction

At ECMWF the background error formulation used in the data assimilation is wavelet based and varies in space and time to take account of geographical differences and the flow of the day (Fisher and Andersson, 2001). The size of a full background error matrix is several orders of magnitude larger than a full model state, which makes it impractical to store the matrix in memory. Because of this the full background error is only available as a sequence of simpler operators, currently consisting of: standard deviations, which change for each assimilation cycle based on the latest available ensemble standard deviations or by other methods; horizontal and vertical correlation matrices, which show geographical variation reflecting climatological averages; and balance operators to account for inter-variable correlations, which are global climatological averages. For satellite retrievals which combine observations and background information for a single atmospheric column or line of sight, horizontal correlations are not needed, and a simpler vertical covariance matrix can be used which still captures most of the flow and geographical dependency of the operationally used background errors. This report will describe this simplified vertical background error covariance framework and outline its implementation as a set of stand-alone FORTRAN programs and scripts.

2 The background error covariance model

2.1 Why not use 3D background errors directly in 1D-VAR?

The background errors used in 3D analyses consist of a sequence of operations, including variable transforms, division by standard deviations and applying horizontal and vertical correlations:

- Transform model variables to a set of uncorrelated variables. Example: $\delta T \rightarrow \delta T_u = \delta T - f(u,v)$, where $f(u,v)$ includes geostrophic balance.
- Divide by standard deviations σ_u of the uncorrelated variables.
- Apply horizontal correlations.
- Apply vertical correlations.

For use in 1D-VAR satellite radiance retrievals, the 3D errors are not the right errors to use because the 1D-VAR problem differs in several respects from a 3D data assimilation,

- When no wind is involved, total temperature δT is used instead of the unbalanced temperature δT_u. The absence of wind correlations simplifies the background errors.
- Different standard deviations are needed because the analysis variables are different, e. g. σ_T, not σ_{Tu}.
- No horizontal correlations are needed.
- Different vertical correlation matrices are needed because the analysis variables are different, e. g. for δT, not δT_u.

It can be seen that the background errors suitable for 1D-VAR are related to the 3D errors used in data assimilation, but they are simpler. In particular horizontal correlations and balances between wind and temperature do not need to be included, which is a major simplification. What remains to be used by the 1D-VAR analysis can be illustrated by writing down the the costfunction for background (J_b) and observations (J_o) that is minimized by the analysis:

$$J(\delta \mathbf{x}) = J_b + J_o = \delta \mathbf{x}^T \mathbf{B}^{-1} \delta \mathbf{x} + J_o(\delta \mathbf{x}) \quad (1)$$

The background error costfunction now contains the following sequence of steps:

1. Transform model variables to a set of uncorrelated variables, with $\mathbf{K}^{-1}$ a variable transform operator which approximately removes error correlations between variables, $\delta \mathbf{x}_u = \mathbf{K}^{-1} \delta \mathbf{x}$

2. Divide by standard deviations of the uncorrelated variables, with Σ a diagonal matrix of background error standard deviations, $\delta\mathbf{x}' = \Sigma^{-1}\delta\mathbf{x}_u$

3. Apply vertical correlations, with $\mathbf{V}$ a vertical correlation matrix, $\delta\chi = \mathbf{V}^{-1/2}\delta\mathbf{x}'$

Inserting this in the background error costfunction gives a few equivalent expressions of the costfunction,

$$\begin{aligned} J_b &= \delta\chi^T\delta\chi \\ &= \delta\mathbf{x}'^T\mathbf{V}^{-1}\delta\mathbf{x}' \\ &= \delta\mathbf{x}_u^T\Sigma^{-T}\mathbf{V}^{-1}\Sigma^{-1}\delta\mathbf{x}_u \\ &= \delta\mathbf{x}^T\mathbf{K}^{-T}\Sigma^{-T}\mathbf{V}^{-1}\Sigma^{-1}\mathbf{K}^{-1}\delta\mathbf{x} \end{aligned} \tag{2}$$

We now describe in more detail how the different operators of the background errors are estimated.

2.2 Transforming from model to analysis variables: $\Sigma\mathbf{K}$

The vertical correlation matrix used by the analysis is expressed in terms of a transformed analysis variable

$$\delta\mathbf{x}' = \Sigma^{-1}\mathbf{K}^{-1}\delta\mathbf{x} \tag{3}$$

which for ozone and temperature is just normalization with the error standard deviation, but for humidity another variable is used, which is close to linearized relative humidity (Hólm et al., 2002).

2.2.1 Humidity transform

The transformed humidity variable is designed to reduce the correlation between temperature and humidity errors and normalizing the humidity to make the errors more Gaussian. We show in some detail how it is derived to make the current description self-contained.

One main contributor to the correlation between temperature and water vapour errors is condensation/evaporation following temperature changes in clouds. In a cloud, the water vapour mixing ratio is by definition at its saturation value $q^c = q_s$, where q^c is the in-cloud value of q. For a given background temperature T_b the Clausius-Clapeyron equation (see e. g. Rogers and Yau 1989) gives how the mixing ratio changes in response to a temperature change dT,

$$\frac{de_s}{dT} = \frac{e_s(T_b)L}{R_vT_b^2} \Rightarrow dq^c = dq_s \approx \frac{q_s(T_b)L}{R_vT_b^2}dT \tag{4}$$

where e_s is the saturation vapour pressure, L is the latent heat for mixed phase, R_v is the gas constant for water vapour and q_s is the saturation water vapour missing ratio. Here we have used $q_s = (R_d/R_v)e_s/(p-e_s) \approx (R_d/R_v)e_s/p$ because the pressure $p \gg e_s$ in the troposphere and the stratosphere (R_d is the gas constant for dry air). Only a certain fraction of a model gridbox is saturated however, and the total gridbox increment in water vapour mixing ratio is a sum of the in-cloud changes correlated with temperature errors, and all other changes δq_u that are uncorrelated with temperature errors,

$$\delta q = \delta q_u + Q_{qT}rh_b\delta q^c = \delta q_u + Q_{qT}rh_b\frac{q_s(T_b)L}{R_vT_b^2}\delta T \tag{5}$$

Here $Q_{qT}rh_b$ takes into account that the correlation between humidity and temperature errors is mainly confined to clouds. Q_{qT} is determined as a correlation coefficient between $\delta q/q_s(T_b)$ and $\frac{rh_bL}{R_vT_b^2}\delta T$ and provided as a climatological polynomial function of relative humidity and model level. The value of Q_{qT} ranges from 0 at ca 80% relative humidity to 1 at saturation. Rearranging this expression we get finally

$$\frac{\delta q_u}{q_s(T_b)} = \frac{\delta q}{q_s(T_b)} - Q_{qT}\, rh_b \frac{L}{R_v T_b^2} \delta T \tag{6}$$

It is useful to divide by $q_s(T_b)$ because it makes the background error statistics more Gaussian (Hólm et al., 2002). The above expression is very close to linearized relative humidity, except for the presence of Q_{qT},

$$\delta rh = \delta(\frac{e}{e_s}) \approx \delta(\frac{q}{q_s(T)}) \approx \frac{\delta q}{q_s(T_b)} - rh_b \frac{q_s(T_b)L}{R_v T_b^2} \delta T \tag{7}$$

Because of this relationship, it can be shown that the error variances for $\frac{\delta q_u}{q_s(T_b)}$ are nearly the same as those of δrh (because close to $rh_b = 1$ the relationships are very similar and for $rh_b \ll 1$ the temperature contribution to the variances is order of magnitude smaller (see Hólm et al., 2002).

2.2.2 *Direct and inverse transforms and their adjoints*

There are four variable transform operations that could be required within the analysis. First, direct transform from analysis to model variables and inverse transform from model to analysis variables are required. Second, for some 1D-VAR applications the adjoints of the direct and the inverse transforms are also required.

We start with the inverse transformation from the model variables as above,

$$\delta T' = \delta T / \sigma_T \tag{8}$$

$$\delta q' = \left(\delta q / q_s(T_b) - Q_{qT} \frac{rh_b L}{R_v T_b^2} \delta T \right) / \sigma_{rh} = (\delta q / q_s(T_b) - Q_{qT}\, \alpha \delta T) / \sigma_{rh} \tag{9}$$

$$\delta o3' = \delta o3 / \sigma_{o3} \tag{10}$$

Here the primed variables are those for which the vertical correlation matrices are valid, see second line of Eq. 2, thus for example $\delta q' = (\delta q_u / q_s(T_b))/\sigma_{rh}$. This can be expressed in matrix form as follows:

$$\delta \mathbf{x}' = \begin{pmatrix} \delta T' \\ \delta q' \\ \delta o3' \end{pmatrix} = \begin{pmatrix} 1/\sigma_T & 0 & 0 \\ -Q_{qT}\, \alpha / \sigma_{rh} & 1/(q_s \sigma_{rh}) & 0 \\ 0 & 0 & 1/\sigma_{o3} \end{pmatrix} \begin{pmatrix} \delta T \\ \delta q \\ \delta o3 \end{pmatrix} = \Sigma^{-1} \mathbf{K}^{-1} \delta \mathbf{x} \tag{11}$$

So given profiles of background variables, background error standard deviations and increments, a transformed increment is calculated which can be used together with the correlation matrices within 1D-VAR.

The direct transform to the model variables can be derived from the above inverse transform equations as

$$\delta T = \sigma_T \delta T' \tag{12}$$

$$\delta q = q_s(T_b) \left(\sigma_{rh} \delta q' + Q_{qT}\, \alpha \delta T \right) \tag{13}$$

$$\delta o3 = \sigma_{o3} \delta o3' \tag{14}$$

or in matrix form:

$$\delta\mathbf{x} = \begin{pmatrix} \delta T \\ \delta q \\ \delta o3 \end{pmatrix} = \begin{pmatrix} \sigma_T & 0 & 0 \\ q_s Q_{qT}\alpha\sigma_T & q_s\sigma_{rh} & 0 \\ 0 & 0 & \sigma_{o3} \end{pmatrix} \begin{pmatrix} \delta T' \\ \delta q' \\ \delta o3' \end{pmatrix} = \mathbf{K}\Sigma\delta\mathbf{x}' \tag{15}$$

The adjoint of the inverse transform is given by (with superscript * for adjoint operators and variables, and noting that the adjoint of a linear operator is just its transpose)

$$\delta\mathbf{x}^{*'} = \begin{pmatrix} \delta T^{*'} \\ \delta q^{*'} \\ \delta o3^{*'} \end{pmatrix} = \begin{pmatrix} 1/\sigma_T & -Q_{qT}\alpha/\sigma_{rh} & 0 \\ 0 & 1/(q_s\sigma_{rh}) & 0 \\ 0 & 0 & 1/\sigma_{o3} \end{pmatrix} \begin{pmatrix} \delta T^* \\ \delta q^* \\ \delta o3^* \end{pmatrix} = \Sigma^{-1*}\mathbf{K}^{-1*}\delta\mathbf{x}^* \tag{16}$$

Finally the adjoint of the direct transform is

$$\delta\mathbf{x}^* = \begin{pmatrix} \delta T^* \\ \delta q* \\ \delta o3^* \end{pmatrix} = \begin{pmatrix} \sigma_T & q_s Q_{qT}\alpha\sigma_T & 0 \\ 0 & q_s\sigma_{rh} & 0 \\ 0 & 0 & \sigma_{o3} \end{pmatrix} \begin{pmatrix} \delta T^{*'} \\ \delta q^{*'} \\ \delta o3^{*'} \end{pmatrix} = \mathbf{K}^*\Sigma^*\delta\mathbf{x}^{*'} \tag{17}$$

2.3 What about $\log q$?

If the 1D-VAR variables are $\log q$ or $\log o3$, then an additional step is needed to transform from q to $\log q$ etc. Increments of $\log q$ can be approximated by linearizing around the background value q_b, so

$$\delta \log q \approx \frac{\delta q}{q_b} \tag{18}$$

The way to go from q increments to increments of $\log q$ is thus to divide δq by q_b. This is easiest achieved by adding another step to the transform, after δq has been calculated. This would be done outside the current program package, on the 1D-VAR code side.

Alternatively, the control variable transform could be modified to give $\delta q/q_b$ directly, which would require a code change inside the program package. In this case the equations for humidity and ozone direct and inverse transforms above need to be modified as follows:

$$\frac{\delta q}{q_b} = q_s(T_b)\left(\sigma_{rh}\frac{\delta q'}{rh_b} + Q_{qT}\frac{L}{R_v T_b^2}\delta T\right) \tag{19}$$

$$\frac{\delta o3}{o3_b} = \sigma_{o3}\frac{\delta o3'}{o3_b} \tag{20}$$

$$\delta q' = \left(rh_b\frac{\delta q}{q_b} - Q_{qT}\frac{rh_b L}{R_v T_b^2}\delta T\right)/\sigma_{rh} \tag{21}$$

$$\delta o3' = \left(o3_b\frac{\delta o3}{o3_b}\right)/\sigma_{o3} \tag{22}$$

2.4 Background error standard deviations Σ

The background error standard deviations Σ are obtained from the ECMWF ensemble data assimilation system, and are operational products that can be retrieved twice daily (valid at 09UTZ and 21UTZ) from the ECMWF MARS archive (Bonavita et al., 2011). For humidity, the standard deviation of the relative humidity errors is used, because the analysis variable used for humidity (see above) is close to linearized relative humidity. Figure 1 shows examples of the standard deviations for a given day.

2.5 Vertical correlation matrices V

The vertical correlation matrices for temperature, humidity, and ozone are provided as global files with a correlation matrix every 625 km (Fisher and Andersson, 2001). The correlation matrices are climatological averages over a month to a season, and mainly reflect geographical variations, such as land, sea, and orography. It is also of interest to understand how the correlation matrices vary by season and during the diurnal cycle. Currently ECMWF uses a single set of correlation matrices for all seasons, and only updates them following upgrades to the forecasting system, such as changes in vertical resolution or major model and observation system changes. In future it is foreseen that the correlation matrices will be updated more frequently to eventually capture the correlations of the day.

The vertical correlation matrices in Figs 2–10 are shown for the 91 model levels, where level 1 is at the top of the model atmosphere while model level 91 is at surface.

2.5.1 Geographical variation of the vertical correlations

The importance of the geographical variation of the correlation matrices can be seen in the change of the boundary layer correlations in the three examples of 50N (The English Channel), 20N (Sahara), and 20S (South Atlantic subsidence area), all at 0E. Even adjacent correlation matrices can vary significantly, if for example one is over sea and the other over land. The values of each correlation matrix are averages over a 625 km box.

2.5.2 Seasonal variation of the vertical correlations

To investigate the seasonal variation, available samples from the current ECMWF operational system were split into three periods: JFM 2011, MAM 2011, and SO 2011. Each sample contains 19 days by 9 ensemble analysis forecast differences, spaced every 3.5 days in the JFM and MAM samples and every 1.5 days in the SO sample. This particular sampling was chosen because these fields were available as part of calculating new operational ECMWF background errors correlations. It can be seen in Figs. 5–7 that the seasonal variation is small compared with the geographical variation seen in Figs. 2–4. However, there may be other locations where the seasonal variation is larger.

2.5.3 Diurnal variation of the vertical correlations

The diurnal variation can only be captured twice daily, at 09UTZ and 21UTZ, from the current ECMWF analysis system. Two samples of 19 days by 9 differences were created from the JFMAM 2011 sample above, one sample every seven days for 09UTZ and 21UTZ. As for the seasonal variation, the diurnal

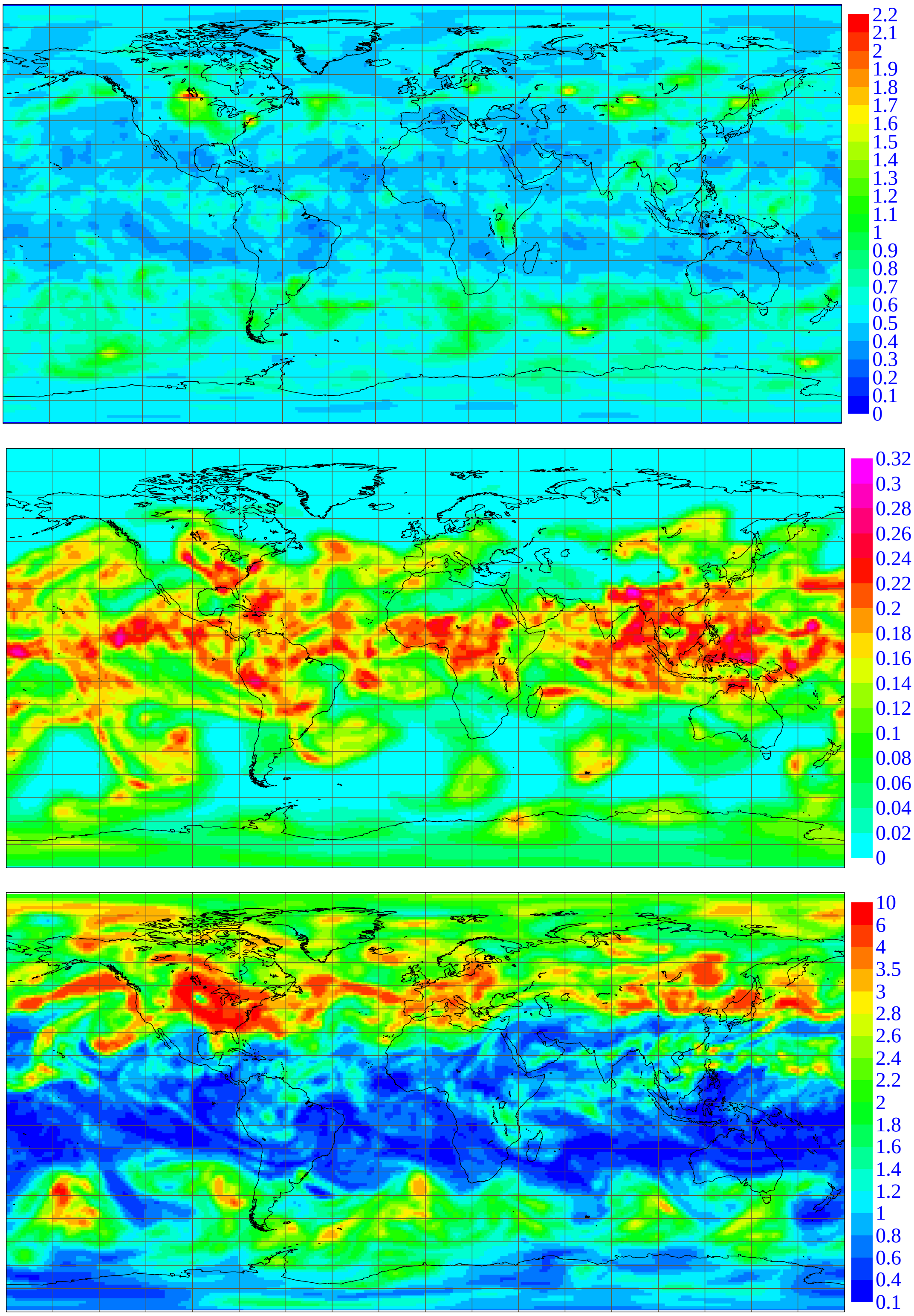

Figure 1: Background error standard deviations at 200hPa for temperature [K] (top), relative humidity [0-1] (middle) and ozone mass mixing ratio [kg/kg $\times$ 1e8] (bottom).

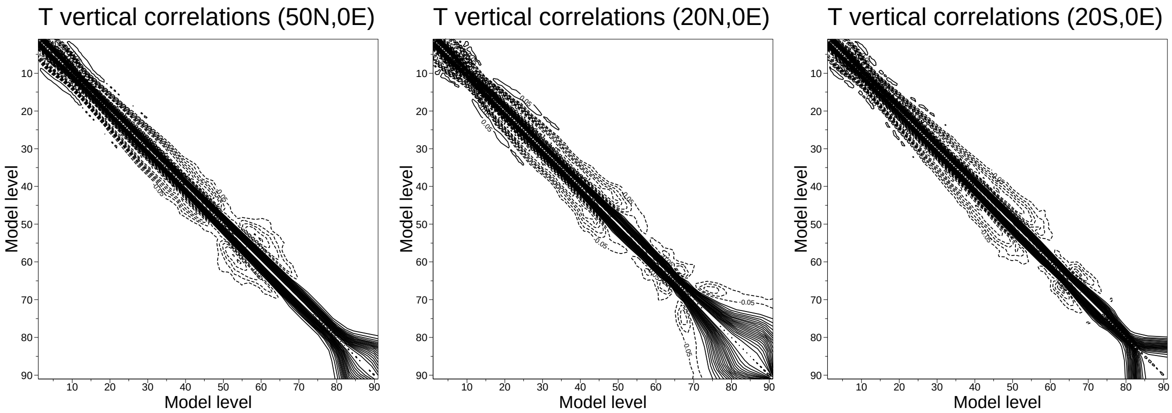

Figure 2: Temperature vertical correlation matrices at 0E and 50N (left), 20N (middle), and 20S (right).

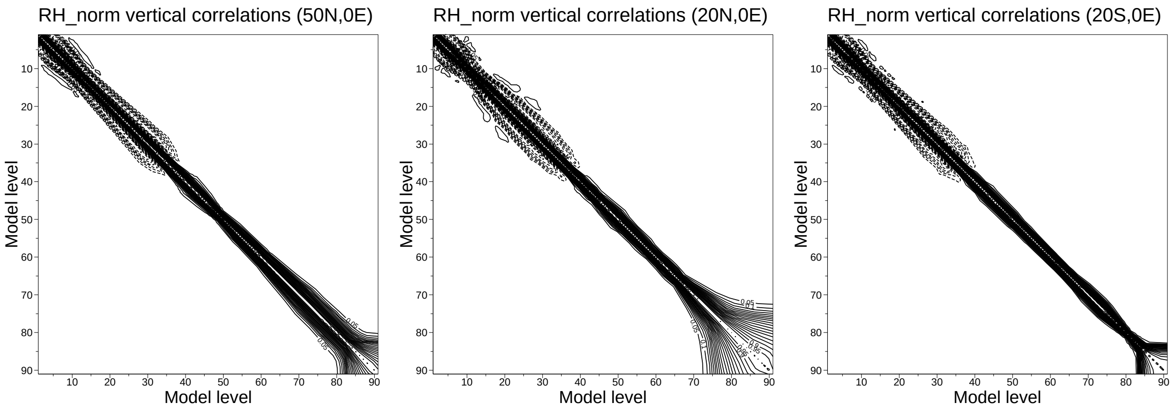

Figure 3: Humidity vertical correlation matrices at 0E and 50N (left), 20N (middle), and 20S (right).

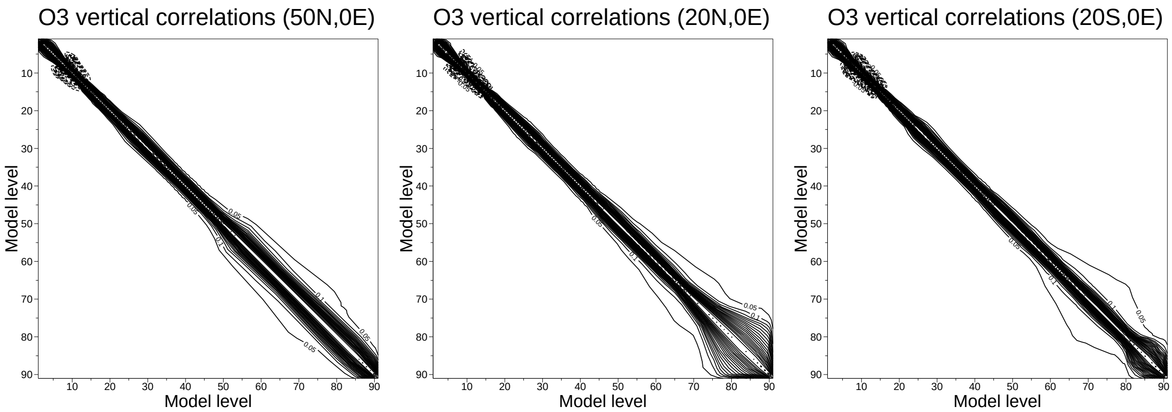

Figure 4: Ozone vertical correlation matrices at 0E and 50N (left), 20N (middle), and 20S (right).

variation in Figs. 8–10 is small compared with the geographical variation seen in Figs. 2–4. This does not exclude that the diurnal variation is important at other locations though.

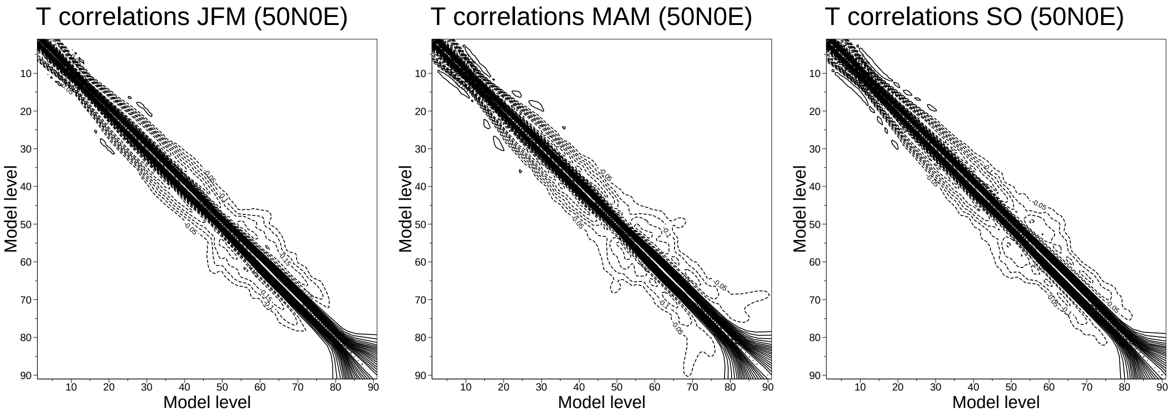

Figure 5: Temperature vertical correlation matrices at (0E,50N) for JFM (left), MAM (middle), SO (right).

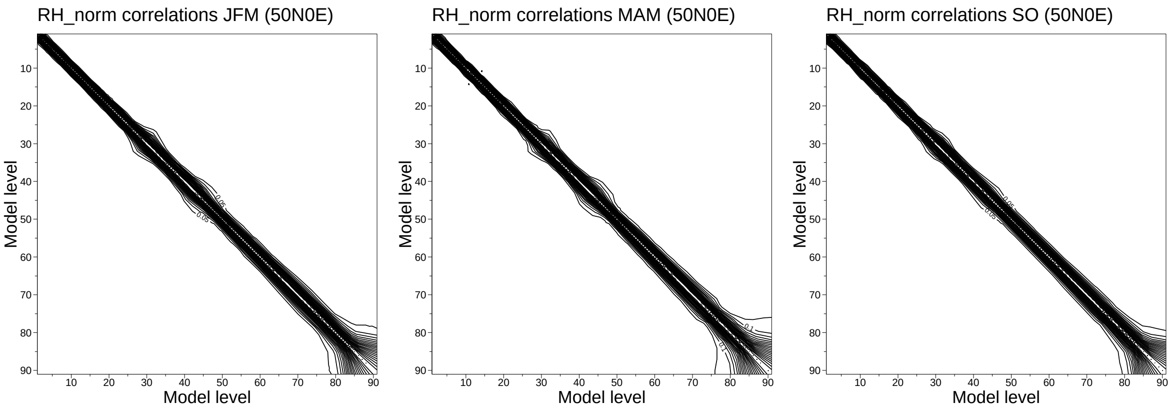

Figure 6: Relative humidity vertical correlation matrices at (0E,50N) for JFM (left), MAM (middle), SO (right).

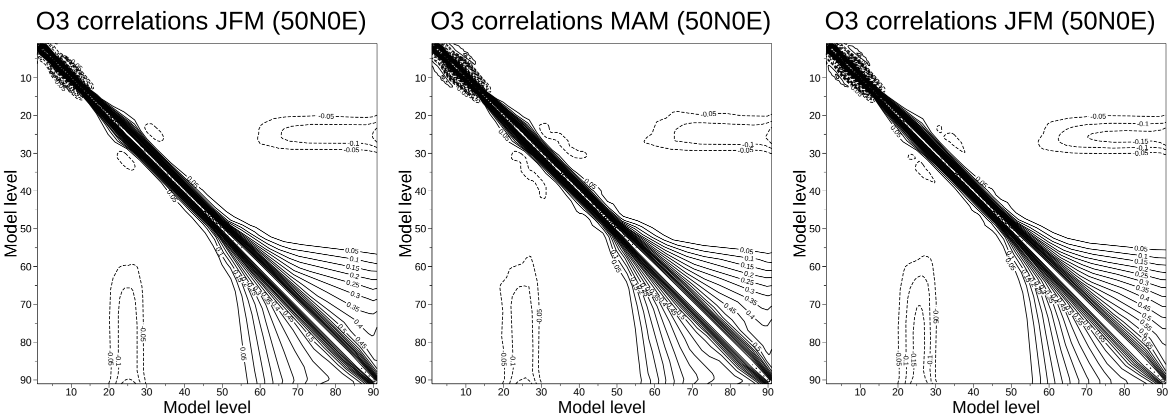

Figure 7: Ozone vertical correlation matrices at (0E,50N) for JFM (left), MAM (middle), SO (right).

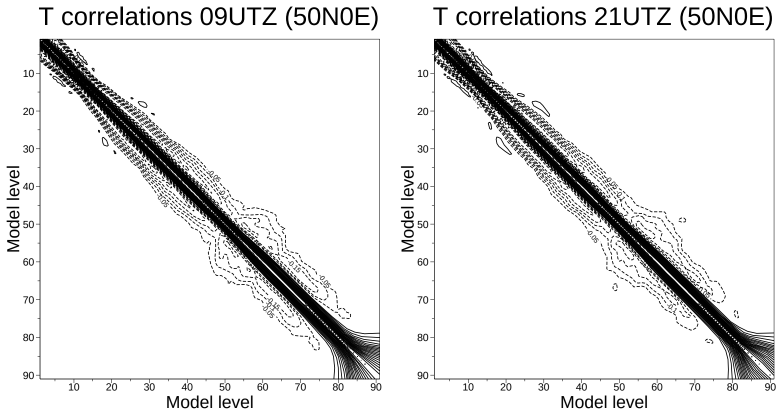

Figure 8: Temperature vertical correlation matrices at (0E,50N) for 09UTZ (left) and 21UTZ (right).

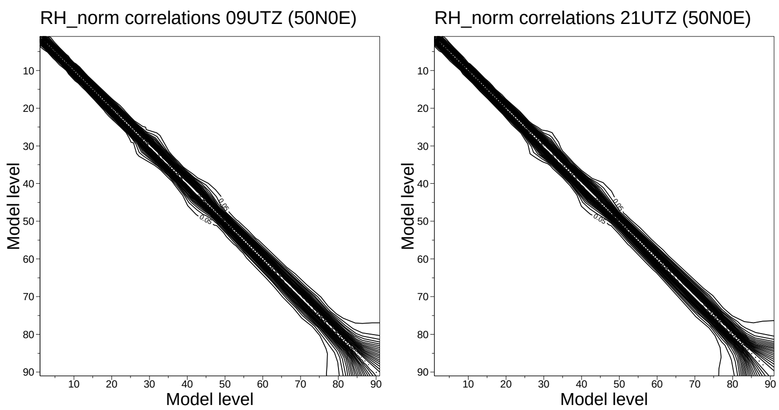

Figure 9: Relative humidity vertical correlation matrices at (0E,50N) for 09UTZ (left) and 21UTZ (right).

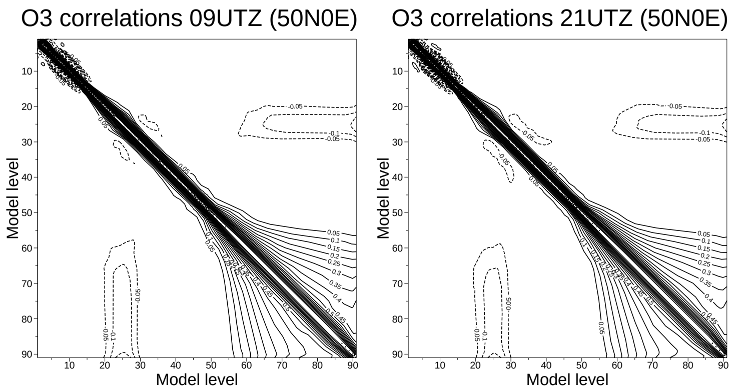

Figure 10: Ozone vertical correlation matrices at (0E,50N) for 09UTZ (left) and 21UTZ (right).

3 Programs

3.1 Requirements on retrieval algorithms using the program package

The program package assumes that the retrieval application provides profiles of temperature, humidity, and ozone increments together with their location, i. e. their latitude and longitude in radiances, and the number of observation locations. These profiles must have been interpolated to the pressure of the ECMWF model levels at each observation location, using the ECMWF surface pressure for that location and time and the ECMWF vertical coordinate parameters. The interpolation of the input profiles is not part of the program package supplied by ECMWF, because it requires knowledge of the retrieval algorithm's vertical grid, which is internal to each retrieval algorithm. Nevertheless, retrieval of surface pressure and its interpolation is part of the program package provided by ECMWF, and this can optionally be used by a retrieval algorithm to obtain the surface pressure and vertical coordinate parameters needed for the vertical interpolation. Another point to note is that only the correlation matrix is provided, and any operations on the matrix, like taking its square root or inverse, will be handled by the retrieval algorithms as needed.

3.2 What the ECMWF background error program package provides

The program and script package supplied by ECMWF (contact the authors for the latest version of the software) provides what is needed to use ECMWF derived background errors for 1D-VAR and similar retrieval applications, given a set of increment profiles and locations (see above). A full documentation is included in the program package itself, in particular start from reading

- *README.txt*: Instructions on how to install, compile, test, and use the package on ibm and linux platforms.

The available operations can be summarized as follows:

- A *vertical correlation matrix* file is provided with the program package, which has a global set of matrices for temperature, relative humidity and ozone mass mixing ratio. The matrices are approximately equally distributed on the globe, as an average over ca 625 km by 625 km area (32 in latitude by $NINT[64cos(latitude)]$ in longitude). A global mean vertical correlation matrix is also provided. For any observation location an index is created pointing to the nearest neighbour correlation matrix for use in retrieval applications. The correlation matrix file can be updated periodically by ECMWF reflecting changes in vertical resolution, horizontal density of the matrices or other major upgrades. There is also an option for seasonal or more frequent updates to the correlation matrices, in line with ECMWF developments in this area.

- A retrieval script *scripts/retrieve.ksh* (see Appendix) contains all that is needed to retrieve the background model fields and the errors of the day from the ECMWF MARS archive (see available dates in next section). The fields retrieved are all global model level fields of temperature, specific humidity, ozone mass mixing ratio, and logarithm of surface pressure. These are the background fields. In addition, all global model levels of background error standard deviations estimated by the ECMWF ensemble data assimilation system are retrieved for the same time, including temperature, relative humidity, and ozone mass mixing ratio. The errors are currently available at T159/N80 (ca. 125 km) reduced Gaussian grids, whereas the model fields are available at up to T1279/N640 (16 km) reduced Gaussian grid. The error fields should be retrieved in their native grid, but the model fields can be retrieved at a higher resolution.

- Interpolation subroutines interpolate all correlation matrices (*retrieve_correlations*), model and error fields (*retrieve_profiles*) bi-linearly to a set of observation locations, providing profiles of background fields and errors matching the input increment profiles (see code *sources/bgerr.F90* for these subroutines with full documentation of input and output variables) .

- A variable transformation subroutine *mod2anvar* transforms the input increment profiles to the variables used in the analysis (see code *sources/vartransform.f90* for *mod2anvar* with full documentation of input and output variables). For temperature and ozone this is simply a normalization by the background error standard deviations, but for humidity the transform takes into account the correlation between humidity and temperature errors by forming what is close to linearized relative humidity. This is the reason why relative humidity background error standard deviations are used in the normalization of the humidity control variable. The variable transform program does an 'inverse' transform, 'direct' transform, and the adjoint of both. Each of these operations may be required by different retrieval algorithm versions.

- A simple test program *tests/bgerr_test.f90* and test files containing complete model states and errors are provided for testing the installation of the program package in different environments. A README file gives instructions on compilation and use of the program package.

3.3 Available dates for *scripts/retrieve.ksh*

Because the background errors here rely on the ensemble data assimilation system (EDA) to provide background error standard deviations, the retrieval script *scripts/retrieve.ksh* (see Appendix) for the errors is not compatible with dates before the current form of the EDA was implemented. Here is a summary of the relevant dates, and options for previous dates.

- After 2010110909 the ECMWF archive contains all that is needed in the script. We only have the background error fields available at 09UTZ and 21UTZ as 3 hour forecasts, and the time and step in the retrieval of the errors and the forecast used in the errors should always be:

  – step=3
  – time=06 or 18

  The background errors will be most accurate for times close to 09 or 21, but they can also be assumed to be the best available estimate for a time window of 6 hours on either side of these times, or a time window of 12 hours after the valid time of the errors. The errors will of course be less accurate the further away one is from their valid time. Although the degree of accuracy has not been quantified as a function of time differences, the errors should still be an improvement on climatological errors for up to 12 hour time differences. These errors can then be combined with retrievals using forecasts at different ranges in the 12-hour window. Note the errors are more accurate after 2011051800 when they began to be used more extensively in the operational system.

- Between 2006020109-2010110821 one can replace the error standard deviations (type=ses) by a less accurate estimate (type=ef). The retrieval script *scripts/retrieve.ksh* will need modification in this case and we recommend this only for expert users.

- Between 2004062909- 2006013121 there are 60-level ”ef” standard deviations available. However, another correlation matrix is needed (for 60 levels, which ECMWF would need to produce), and we recommend staying within the 91 level period.

- Before 2004062909 there are no ”ef” fields for the standard deviations.

- For intercomparison, we do also still produce the ”ef” fields today (end 2012, but may not in future) - so a date post 2010110909 can be used to understand the difference between ”ef” and ”ses”. The ”ses” have more geographic variability and gives a better estimate (calibrated against analysis) of the background error standard deviation, whereas ”ef” give flatter fields, less reflecting day-to-day variations.

4 Conclusions

The program package provided gives flow and geographically varying background error covariance matrices for temperature, humidity and ozone, together with the required variable transforms needed to be close to the ECMWF implemented humidity analysis formulation. The use of flow dependent background error standard deviation of the day from the operational ECMWF ensemble data assimilation system and geographically varying correlation matrices is a considerable improvement compared with earlier global average version of background errors used in 1D-VAR applications. A first look at the vertical correlation matrices seasonal and diurnal variations showed these to be small compared with the geographical variation, but further studies are needed to quantify the variation globally. This initial look does however give confidence that correlation matrices averaged over one season are valid for other seasons and at all times of day. Further developments of this program package are foreseen after experience has been gained in using the package in 1D-VAR applications, and it is possible to extend the framework with further variables, in particular cloud related variables and trace gasses. For cloud and trace gas variables, the background error estimation is not as mature as for humidity, temperature, and ozone, which all are part of the ECMWF operational framework, and attention needs to be given to the scientific validity of these additional background errors in addition to their technical implementation.

Acknowledgements

This work was funded in part by EUMETSAT contract 11/204090: MTG-IRS L2 Processing - Potential of ensemble forecast to provide background information. The covariance processing program package combines selected parts of subroutines previously developed by ECMWF for diagnosing background errors.

References

Bechtold, P., and J. W. M. Cuijpers (1995) Cloud perturbations of temperature and humidity: a LES study. *Boundary-Layer Meteorology*, **76**, 377-386.

Bonavita, M., L. Isaksen and E. Hólm (2011) On the use of EDA background error variances in the ECMWF 4D-Var. *ECMWF Technical Memorandum 664*.

Fisher, M., and E. Andersson (2001) Developments in 4D-Var and Kalman Filtering. *ECMWF Technical Memorandum 247*.

Hólm, E., E. Andersson, A. Beljaars, P. Lopez, J.-F. Mahfouf, A. J. Simmons and J.-N. Thepaut (2002) Assimilation and Modelling of the Hydrological Cycle: ECMWF’s Status and Plans. *ECMWF Technical Memorandum 383*.

Rogers, R. R, and M. K. Yau (1989) A Short Course in Cloud Physics. Third Edition. Oxford: Pergamon Press.

Appendix: Script retrieval.ksh

```
#!/bin/ksh

nlev=91
truncation=159
target_errors="errors.grib"
target_fields="fields.grib"

usage="Usage: ${0/*\//} [options] -d <date> -t <time_in_hours> -s <step>\n
\nRetrieve errors and first guess fields from MARS.\n
\nOptions:\n
\t-l NLEV\t\t number of model levels (default $nlev)\n
\t-T TRUNC\t spectral truncation (default $truncation)\n
\t-e TARGET\t target file for errors (default \"$target_errors\")\n
\t-f TARGET\t target file for model fields (default \"$target_fields\")
"

[[ $# == 0 ]] && { echo $usage; exit 1; }

while getopts "hd:t:s:l:T:e:f:" option
do
    case $option in
        h) echo $usage; exit 0;;
        d) date=$OPTARG;;
        t) time=$OPTARG;;
        s) step=$OPTARG;;
        l) nlev=$OPTARG;;
        T) truncation=$OPTARG;;
        e) target_errors=$OPTARG;;
        f) target_fields=$OPTARG;;
        *) echo "Error: Unknown option: -$option"; echo $usage; exit 1;;
    esac
done
shift $((OPTIND - 1))

date=${date:?"Error: missing date"}
time=${time:?"Error: missing time"}
step=${step:?"Error: missing step"}
grid=$((($truncation + 1) / 2))

mars <<MARS_REQUEST
retrieve,
    expver=0001,
    class=od,
    date=$date,
    time=$time,
    step=$step,
```

```
    level=1/to/$nlev,
    levtype=ml,
    type=ses,
    stream=enda,
    repres=gg,
    param=130/157/203,
    target="$target_errors"
retrieve,
    type=fc,
    stream=dcda,
    repres=sh,
    param=130/152,
    resol=av,
    grid=$grid,
    gaussian=reduced,
    target="$target_fields"
retrieve,
    repres=gg,
    param=133/203.128,
    target="$target_fields"
MARS_REQUEST
```