

Department of Mathematics and Statistics

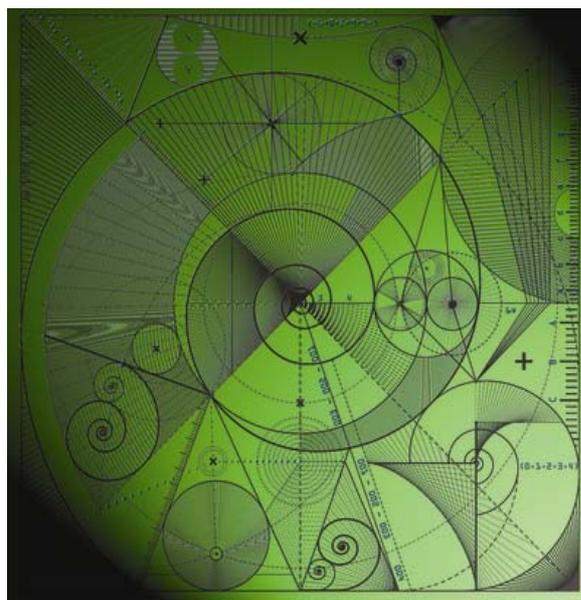
Preprint MPS-2012-26

28 November 2012

Estimating interchannel observation error correlations for IASI radiance data in the Met Office system

by

L.M. Stewart, S.L. Dance, N.K. Nichols, J.R. Eyre
and J. Cameron



Estimating interchannel observation error correlations for IASI radiance data in the Met Office system

L. M. Stewart*, S. L. Dance†, N. K. Nichols
School of Mathematical and Physical Sciences, University of Reading, Reading, RG6 6BB,
J. R. Eyre, J. Cameron
Met Office, FitzRoy Road, Exeter, EX1 3PB

November 2012

Abstract

The optimal utilisation of hyper-spectral satellite observations in numerical weather prediction is often inhibited by incorrectly assuming independent inter-channel observation errors. However, in order to represent these observation error covariance structures, an accurate knowledge of the true variances and correlations is needed. This structure is likely to vary with observation type and assimilation system. The work in this paper presents the initial results for the estimation of IASI inter-channel observation error correlations when the data is processed in the Met Office one-dimensional (1D-Var) and four-dimensional (4D-Var) variational assimilation systems. The method used to calculate the observation errors is the post-analysis diagnostic proposed by Desroziers *et al.* (2005) *Q.J.R.Meteorol.Soc.* **131**:3385–3396.

The results show significant differences in the source and structure of the observation errors when processed in the two different assimilation systems, but also highlight some common features. When the observations are processed in 1D-Var, the diagnosed error variances are approximately half the size of the error variances used in the current operational system and are very close in size to the instrument noise, suggesting that this is the main source of error. The errors are predominantly uncorrelated, with the exception of a handful of neighbouring channels. When the observations are processed in 4D-Var, we again find that the observation errors are being overestimated operationally, but the overestimation is significantly larger for many channels. In contrast to 1D-Var, the diagnosed error variances are often larger than the instrument noise in 4D-Var. This suggests that horizontal errors of representation, not seen in 1D-Var, are a large contributor to the overall error here. Finally, observation errors diagnosed from 4D-Var are found to contain strong, consistent correlation structures for channels sensitive to water vapour and surface properties.

1 Introduction

In numerical weather prediction (NWP) the governing equations used to describe the behaviour of the atmosphere are sampled by millions of observations in a 6 hour period. Satellite observations are the dominant contributor to the data assimilation algorithms used in NWP in both quantity and impact on forecast accuracy. [Bouttier and Kelly \(2001\)](#) and [Joo *et al.* \(2012\)](#) show that a significant contribution to this dominance is given by hyperspectral sounders, such as the Infrared Atmospheric Sounding Interferometer (IASI) and the Atmospheric Infrared Sounder (AIRS), which sample the infrared spectrum using thousands of channels at a close spectral proximity.

The contribution an observation makes to the data assimilation system is governed by the error associated with its measurement and representation; observations with large errors will receive low weighting in the assimilation while those with small errors have a larger impact. Satellite observation errors, including those from hyper-spectral sounders, are usually assumed to be uncorrelated in both the horizontal

*MetOffice Reading, Meteorology Building, University of Reading, Reading, RG6 6BB

†Corresponding author email: s.l.dance@reading.ac.uk

and the vertical. For IASI observations, the assumption of horizontal uncorrelation is supported by intelligent thinning of the data to avoid assimilating observations which are too close spatially.

Ensuring vertically independent observation errors is more difficult. A measurement in any IASI channel will be sensitive to the temperature and humidity profile over several atmospheric levels; the distribution being represented by the broad weighting functions of the instrument. Therefore the errors in channels spectrally close to each other are likely to be correlated if some mis-representation of the observation is common over several levels; for example, the sensitivity to a certain trace gas. Additionally, correlated errors of representation will be present between channels that observe spatial scales or features within the instrument's 12km field-of-view which the model does not represent well. Also, any errors in the forward model, such as those of spectroscopy, may be correlated between channels.

The current IASI channel selection procedure attempts to lessen the impact of vertically correlated errors by avoiding the assimilation of adjacent channels. But it is generally acknowledged that vertical correlation structure is still present among those channels chosen for assimilation. However, because of a lack of knowledge of the true error correlation structure and the perceived expense of including this structure in the assimilation, the observation errors are treated as uncorrelated between channels. To avoid mis-weighting the IASI observations in the data assimilation, the error variances are inflated to account for the lack of correlation. The level of inflation is determined by the wavelength band of the channels. This technique is not only used at the Met Office, but at the time the work was undertaken, was the technique Meteo-France and ECMWF applied to treat IASI observation error correlations.

Mis-treating observation errors as uncorrelated has been shown to be detrimental to the analysis accuracy and information content when using hyper-spectral satellite observations (Rabier *et al.*, 2002; Liu and Rabier, 2003; Collard, 2004; Stewart *et al.*, 2008; Stewart, 2010; Stewart *et al.*, 2012). When used in conjunction with data thinning or channel constraining algorithms, it inhibits the impact of using greater volumes of satellite data, as shown in Dando *et al.* (2007) and Collard (2004). With the increasing resolution of NWP models, such a restriction on data usage is hardly desirable. This motivated studies into the quantification, and future use in assimilation, of these observation error correlations.

A number of methods exist for quantifying error correlations in data assimilation; however the application of such methods is not without difficulty. The most commonly used estimation technique is the observational method, often known as the Hollingsworth-Lönnerberg method. Hollingsworth and Lönnerberg (1986) describe the method which separates background and observation errors using background innovation statistics, under the assumption that the background errors carry spatial correlations while the observation errors do not. With additional inputs, it can be modified to account for correlated errors in the observations. Garand *et al.* (2007) applied the method to AIRS data, showing significant inter-channel error correlations. More recently, Bormann and Bauer (2010a) and Bormann *et al.* (2010b) applied the method to ATOVS, AIRS and IASI data used in the ECMWF analysis, again demonstrating considerable correlation structures in certain wavelength bands.

In other techniques, Dee and da Silva (1999) used a maximum likelihood method to estimate information on error statistics. This work resulted in the derivation of statistical parameters that varied in time. Desroziers and Ivanov (2001) used statistics from analysis innovations to tune background and observation error parameters, resulting in a successful description of the observation error parameters in a 3D-Var framework.

A method that addressed the separation of correlated observation and background errors was proposed in Desroziers *et al.* (2005). This method uses post-analysis diagnostics from linear estimation theory to statistically approximate the covariances of the observation errors. This approach has been used to estimate error variances and inter-channel correlations in the Meteo-France, (Desroziers *et al.*, 2005), Met Office (Stewart *et al.*, 2009) and ECMWF, (Bormann *et al.*, 2010b), variational assimilation systems. Findings include an over-estimation of observation error variances and strong inter-channel correlations at certain wavelengths.

Once observation error correlations have been quantified, it is useful to be able to identify their

sources. As mentioned above, errors in the meteorology, the forward model, and resolution representation can be correlated in the vertical. Current methods make no attempt to calculate the separate contribution from each source of error.

The work described in this paper considers the initial application of the Desroziers observation error diagnostic to estimate the inter-channel error correlation structure of IASI observations used in both the Met Office 1D-Var and 4D-Var assimilation system. For each assimilation system, we will compare the diagnosed observation error variances with those currently used, and establish the level of correlation between the observation errors. Attention will be paid to any variation in structure between different groups of channels. Unlike previous papers on the subject, we will also compare the error variance and correlation structures diagnosed for the two systems, so as to draw conclusions as to the origin of the errors and their dependence on the assimilation process used. Some results and discussion of this work have been previously presented in [Stewart *et al.* \(2009\)](#) and [Stewart \(2010\)](#).

The paper will be structured as follows. Firstly in section 2, we describe the Desroziers technique used to estimate observation error covariances, and the two assimilation systems that we will be using to process the observations. Section 2 then continues to describe the IASI data used and the experimental set up. The results are contained in section 3. Section 3 is separated into the results for processing the IASI observations in 1DVar, the results for the processing in 4D-Var, and a comparison of the two methods. Finally, a summary and conclusions from the paper are given in section 4.

2 Methodology

Below we describe the methods and data structure used in this paper.

2.1 Desroziers technique of error approximation

The technique proposed by [Desroziers *et al.* \(2005\)](#) is based on linear estimation theory, where the optimal analysis describing the true state of the atmosphere, \mathbf{x}^a , can be expressed in terms of the background state, \mathbf{x}^b , and the background innovation vector, \mathbf{d}_b^o ,

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}\mathbf{d}_b^o. \quad (1)$$

where \mathbf{R} and \mathbf{B} are the observation and background error covariance matrices, respectively, and \mathbf{H} is the linearised version of the nonlinear observation operator H . The background innovation vector is the difference between the observations, \mathbf{y} , and their background counterparts, $H(\mathbf{x}^b)$,

$$\mathbf{d}_b^o = \mathbf{y} - H(\mathbf{x}^b). \quad (2)$$

Similarly, the analysis innovation vector, \mathbf{d}_a^o , is given by the differences between the observations and their analysis counterparts, $H(\mathbf{x}^a)$,

$$\mathbf{d}_a^o = \mathbf{y} - H(\mathbf{x}^a). \quad (3)$$

Desroziers *et al.* showed that by taking the statistical expectation, \mathbb{E} , of the product of (2) and (3) under the assumption of mutually uncorrelated observation and background errors, an approximation of the observation error covariance matrix is obtained:

$$\mathbb{E}[\mathbf{d}_a^o(\mathbf{d}_b^o)^T] \approx \mathbf{R}. \quad (4)$$

This relation is satisfied exactly provided the covariance matrices used in (3) are consistent with the true observation and background error covariances. However, given that we know that we are mis-representing the observation error covariance structure, Desroziers *et al.* suggest the use of the estimation technique as an iterative procedure; using the previously diagnosed \mathbf{R} matrix at each iteration should result in an \mathbf{R} matrix closer and closer to reality.

[Desroziers *et al.* \(2005\)](#) used the estimation technique to successfully diagnose background and observation error variances for radiosonde wind observations in the French ARPEGE system. For a toy

problem they were also able to iteratively tune observation error variances and recover observation error covariances, starting from a mis-specification of both. [Stewart *et al.* \(2009\)](#), [Bormann and Bauer \(2010a\)](#) and [Bormann *et al.* \(2010b\)](#) applied the estimation technique on a larger scale, estimating the error covariance structure for clear-sky sounder radiances used in the Met Office and ECMWF assimilation systems respectively. Results for IASI observations showed noticeable correlation structure in the surface-sensitive and short-wave temperature-sounding channels, and a significant degree of correlation between humidity-sounding channels.

2.2 Assimilation systems

The results presented here will show the inter-channel error correlation structure for IASI observations processed in the Met Office 1D-Var and 4D-Var systems. The two systems are used together for the processing of satellite data in NWP at the Met Office.

The objective of both 1D-Var and 4D-Var retrieval systems is to minimise a cost function that penalises distance from the observations taken at a sequence of times, \mathbf{y}_i , and a previous short-range forecast (or background state), \mathbf{x}^b ,

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2} \sum_{i=1}^n (\mathbf{y}_i - H_i(\mathbf{x}))^T \mathbf{R}_i^{-1}(\mathbf{y}_i - H_i(\mathbf{x})) \quad (5)$$

where n is the total number of observations. The observation operator H_i is comprised of a Radiative Transfer for TIROS Operational Vertical Sounder (RTTOV) radiative transfer model as described in [Saunders *et al.* \(2005\)](#) and [Matricardi *et al.* \(2004\)](#); it accurately predicts brightness temperatures given first-guess model fields of temperature and humidity, as well as surface air temperature, skin temperature and surface humidity. The product of the cost function minimisation is an updated estimate of the analysis state of the atmosphere.

The 1D-Var retrieval system, as described in [Hilton *et al.* \(2009\)](#), is used on individual observations prior to their assimilation into the NWP model. It is called by the Observation Processing System (OPS) which is used to pre-screen and quality control the observations. Since the 1D-Var retrieval is applied to single observations, the cost function (5) is minimised with $n = 1$.

One quality control procedure in the OPS is to identify observations which are too different from the forecast background, and hence may cause problems in the later assimilation. A large value of the cost function (5), or a slow convergence rate, is an indicator of inconsistency between the background forecast and the observations, and hence observations with these attributes are eliminated at this stage. The observations that ‘pass’ the OPS analysis checks are deemed suitable for assimilation in the Met Office 4D-Var assimilation system (VAR).

The 1D-Var retrieval also provides estimates of the atmospheric variables not represented in 4D-Var, but required for radiative transfer calculations. The analysed variables in 4D-Var are a subset of the full state vector variables, and those variables, such as skin temperature, which are not included are unmodifiable. It is therefore crucial to the success of the assimilation that these variables are accurately specified prior to the 4D-Var assimilation. The full state vector, i.e. all variables, is used in the 1D-Var retrieval, and the analysis values of those variables not analysed in 4D-Var are passed there. The first set of statistics in our experiments will be generated using the background, \mathbf{d}_b^o , and analysis, \mathbf{d}_a^o , innovations from these 1D-Var analyses.

When the 1D-Var retrieval is performed in the OPS, the observation operator (or forward model) is fitted separately to each individual column of observations, so the position of the observations, and hence any resolution conflict, is already determined. Therefore, it can be argued that the horizontal errors of representation, created by a contrast in model and observation resolution, will appear in the background matrix B . Hence, the error of representation component of the observation error will be only

Description	OPS channel no.	VAR channel no.	IASI channel no. (wavelength cm^{-1})
Temperature sounding	0-123	0-86	16-457 (648.75-759.00)
Window (surface sensitive)	124-152	87-108	515-2245 (773.50-1206.00)
Water vapour sensitive	148,150,153-182	109-138	2019,2119,2741-5405 (1149.50,1174.50,1330.00-1996.00)

Table 1: Groups of channels by similar spectral properties

in the vertical. Therefore, from the 1D-Var diagnostics we expect observation errors to be predominantly attributed to instrument noise, forward model error and vertical errors of representation.

The 1D-Var retrieval produces a quality controlled subset of brightness temperature measurements suitable for assimilation in the Met Office incremental 4D-Var assimilation system. The 4D-Var retrieval system minimises the cost function (5) over the full sequence of observations, \mathbf{y}_i , and assimilates the observations at their measurement times. In 1D-Var, each observation is assimilated at its own horizontal location, while in 4D-Var all the observations are assimilated together at model grid points. A detailed description of the 4D-Var procedure at the Met Office is given in [Rawlins *et al.* \(2007\)](#).

The 4D-Var algorithm generates an optimal analysis increment which is used to update the solution state at the start of the assimilation time window. From this starting state, the nonlinear model is run over the time window to generate the forecast. The forecast model (or analysis) fields are output at model grid points at pre-determined times, and can be interpolated to the observation locations. In the 4D-Var assimilation, all observation information is fitted to the resolution provided by the model, and so correlated errors of representation (both horizontal and vertical) are expected to be contained wholly in \mathbf{R} . The second set of statistics in our experiments will be generated using the analysis innovations from the interpolated analyses, and the background innovations.

2.3 Data

The Desroziers *et al.* estimation technique described above is applied to data from the IASI instrument onboard MetOp-A. IASI is an infrared Fourier transform spectrometer measuring in the spectral interval of $645\text{-}2760\text{cm}^{-1}$ at a resolution of 0.5cm^{-1} . IASI observations are an important component of the global observing system, and their positive impact on forecasting at the Met Office, ECMWF and Meteo France has been demonstrated in [Hilton *et al.* \(2009\)](#), [Collard and McNally \(2009\)](#) and [Rabier *et al.* \(2009\)](#).

IASI has the potential to provide observations in 8461 channels, but at the time of the research only observations from a subset of 314 derived in [Collard \(2007\)](#) were used. This subset was chosen under the assumption that error correlations between channels would not be represented. The channels are predominantly in the CO_2 temperature sounding band, with additional channels chosen to provide information on water vapour, trace gases, surface properties, etc. Because of the need to avoid highly-correlated channels, the water-vapour band is largely under-sampled despite the need for an accurate representation of humidity structures; the increased use of water vapour channels was shown to result in a degradation in analysis accuracy in the current framework. Of this 314, 183 channels are used in the 1D-Var assimilation and 139 are used in the 4D-Var assimilation; fewer channels are used in the 4D-Var assimilation since the inclusion of certain channels was found to have a negative impact on analysis accuracy under current conditions. For the 4D-Var statistics, only observations that pass the 1D-Var quality control are used.

Table 1 groups the channels used in 1D-Var and 4D-Var by similar spectral properties. The channels are shown on a typical IASI spectrum in Figure 1 (1D-Var channels) and Figure 2 (4D-Var channels); the different colours represent the different groups of channels.

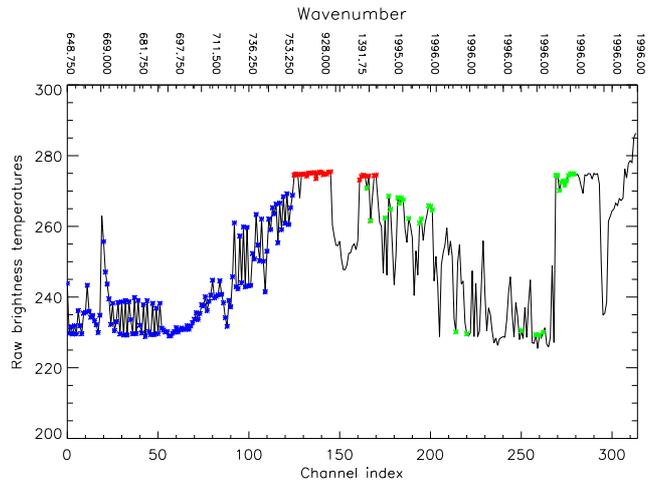


Figure 1: Channels used in 1D-Var on a typical IASI spectrum: Temperature sounding channels indexed 0-123 (blue), window channels indexed 124-152 (red), and water vapour channels indexed 148, 150, 153-182 (green).

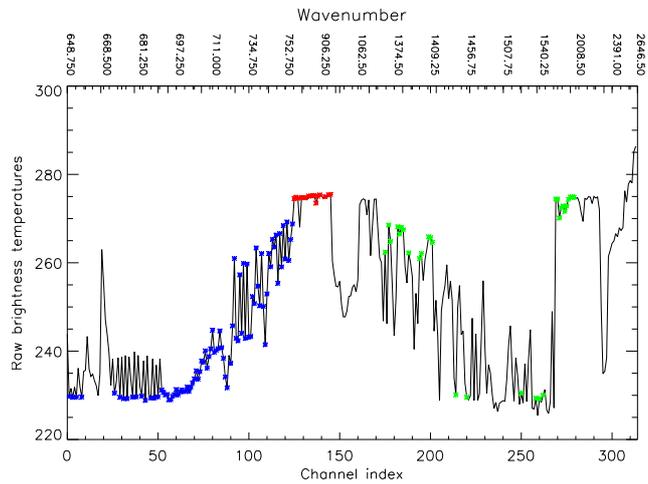


Figure 2: Channels used in 4D-Var on a typical IASI spectrum: Temperature sounding channels indexed 0-86 (blue), window channels indexed 87-108 (red), and water vapour channels indexed 109-138 (green).

2.4 Experimental set-up

The Desroziers diagnostic is calculated for two situations: firstly using the analysis output from the 1D-Var retrieval and secondly using the analysis output from the 4D-Var assimilation. The background and analysis increment statistics are generated from the assimilation of only clear-sky, sea surface IASI observations. Observations will be from both day and night time, with the exception of daytime observations from shortwave channels which will be eliminated. Using only IASI observations in the assimilation avoids the difficulties of attributing the diagnosed error structures to different observation types.

The 1D-Var results presented here are based on the IASI observations taken on 17 July 2008 at 00:00, 06:00, 12:00 and 18:00. The total number of observations used to produce the statistics is 27,854: 9,131 of which were suitable for use in the 4D-Var assimilation and the rest were thinned out. The 4D-Var results are produced from a smaller subset of these IASI observations taken at 18z. A total of 2,073 observations are used to produce the statistics; this is from a subset of the 6,539 observations that are used in the OPS at this time.

In addition to the use of the Desroziers diagnostic, we also use a Hollingsworth-Lönnberg method (Hollingsworth and Lönnberg, 1986) to calculate the error variances from the 4D-Var analyses. A comparison of the two methods for diagnosing observation error covariances has previously been performed for the ECMWF assimilation system in Bormann *et al.* (2010b), and we use it here to verify the robustness of the Desroziers method.

The Hollingsworth-Lönnberg method typically uses background innovations from a dense observing network to estimate background and observation errors, under the assumption that the background errors carry spatial correlations while the observation errors do not. To calculate the correlated and uncorrelated component of inter-channel observation errors, the method is modified slightly and the background innovation covariances for pairs of observations close in time are stratified against horizontal observation separations. By taking the difference between the covariance at zero separation and at a separation representative of model (or analysis) resolution, we can obtain an estimate for the spatially correlated component of the observation error, i.e, forward model error and errors of representation. Using this method, stratifying innovation covariances calculated for the same channel, will provide an estimate of the observation error variances; stratifying innovation covariances calculated for different channels will tell us about the observation error covariances.

Calculating the observation separation representative of the VAR analysis resolution is not a trivial task. For this work a set of observation pairs separated by a range of distances thought to sample the analysis resolution were selected. The average distance between observations was 84.8km. The observation error estimate was found by differencing the covariance of the innovations for all observations in the data set separately (zero separation) and the covariance of the innovations across observation pairs (typically 84.8km). A similar method was used by Bormann and Bauer (2010a) and was shown to give more robust results compared to the use of a correlation function.

3 Results

In this section we present the observation error structures derived when IASI observations are assimilated in the Met Office 1D-Var and 4D-Var systems. The observation errors are derived using the Desroziers technique described in Section 2.1. Comparisons are made with the operational observation error variances, and the observation error correlation matrices are calculated. Diagnosed inter-channel correlation structures will be compared for the two retrieval systems, and conclusions will be drawn as to the origin of the observation error correlations.

3.1 1D-Var assimilation

First we consider the application of the Desroziers diagnostic to the 1D-Var analyses. Figure 3 shows the global location of all the observations used in the 1D-Var retrieval, and the size of their background

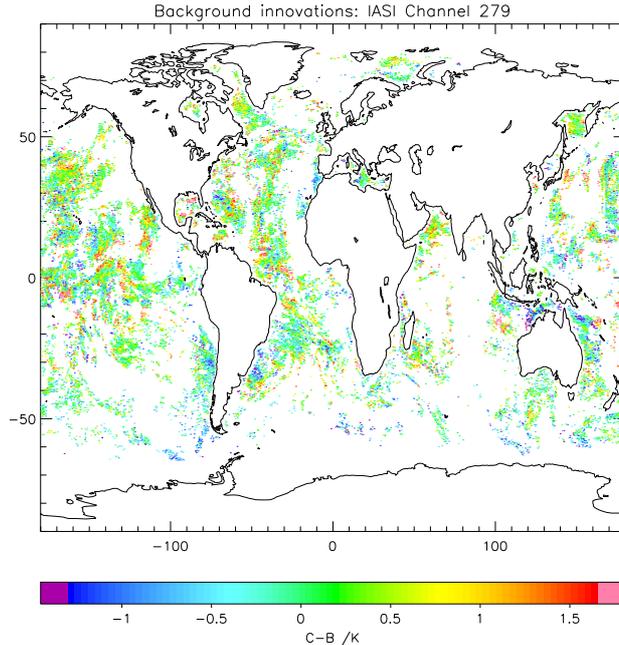


Figure 3: Global location and background innovation value $O - B$ (degrees Kelvin) for observations in IASI channel 5403 (MetDB channel 279).

innovations for IASI channel 5403 which is highly sensitive to water vapour.

The observation error variances calculated using the Desroziers diagnostic, the operational error variances used in 1D-Var at the Met Office and an approximation of the IASI instrument noise are shown in Figure 4. The operational error variances are comprised of the instrument noise plus a forward model error of $0.2^2 K$. Results are shown for the 183 channels used in the 1D-Var retrieval.

The structure of the operational and diagnostic error variances in Figure 4 is very similar, but the diagnosed error variance is noticeably lower than the operational error variance for all channels. The largest difference is in OPS channels 148 - 180 which are highly sensitive to water vapour. Also the diagnosed error variances are very close in size to the instrument noise. This suggests that the forward model component of the error variances is being overestimated in 1D-Var, especially in channels sensitive to water vapour, and that a large part of the true observation error is from the instrument noise.

Figure 5 shows the observation error covariances for the 183 channels used in the 1DVar retrieval. The error covariance plot is diagonally dominant; the diagonal values in Figure 5 correspond to the values plotted on the red line in Figure 4. The darker colours towards the top of the diagonal in Figure 5 correspond to the larger variance values in the water-vapour sensitive OPS channels 165-172 in Figure 4. The covariances appear relatively weak between temperature sounding and surface sensitive OPS channels indexed under 120, with the exception of channel 20. However, channel 20 is a high-peaking channel in the temperature sounding band, which is not used in the 4D-Var assimilation because of the stratospheric ringing of its innovations.

The error correlation matrix, \mathbf{C} , can be determined easily from the error covariance matrix using the identity $\mathbf{R} = \mathbf{D}^{1/2} \mathbf{C} \mathbf{D}^{1/2}$ where \mathbf{D} is the diagonal matrix of error variances. The diagnosed error correlation matrix is shown in Figure 6. Noticeably, the correlation structure is not uniformly symmetric, which is a necessity for an error correlation matrix. However, this is expected since we are violating one of the diagnostic assumptions by using an error covariance matrix to compute the cost function that is not entirely representative of the true error structure, i.e. $\mathbf{R}_{\text{operational}} \neq \mathbf{R}_{\text{truth}}$. Potentially an

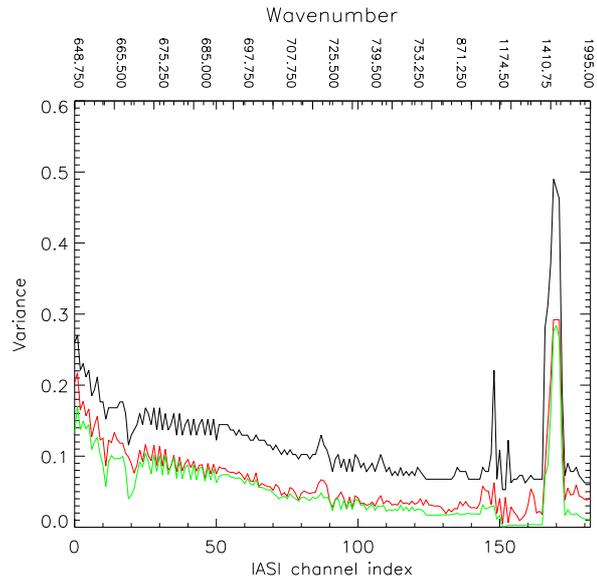


Figure 4: Operational error variances (black line), diagnosed error variances (red line) and instrument noise (green line).

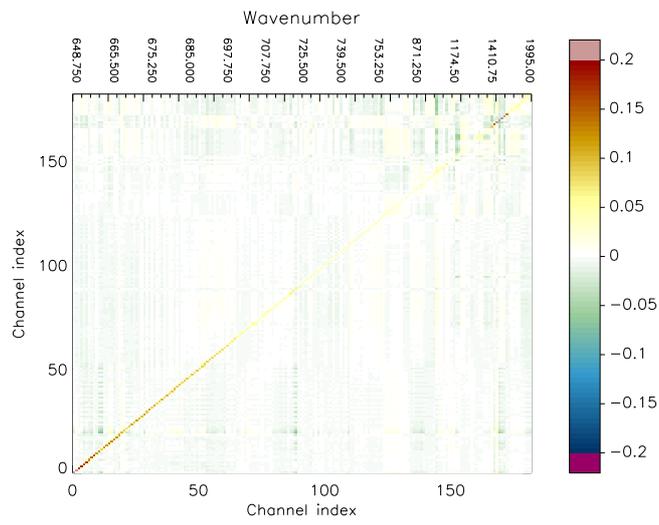


Figure 5: Diagnosed observation error covariance matrix for the OPS.

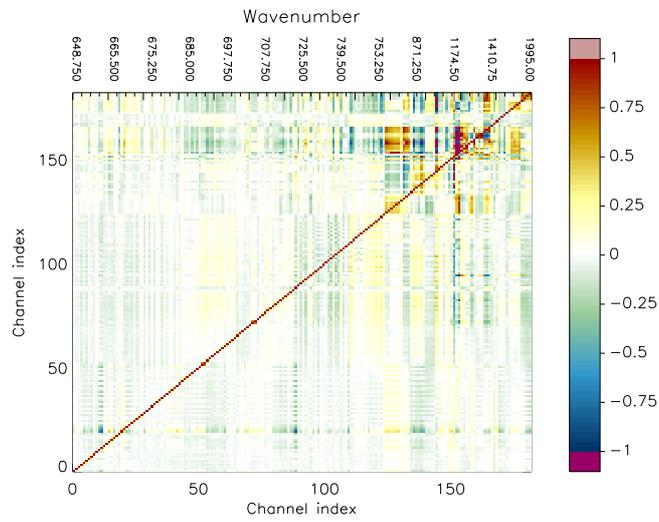


Figure 6: Diagnosed observation error correlation matrix for the OPS.

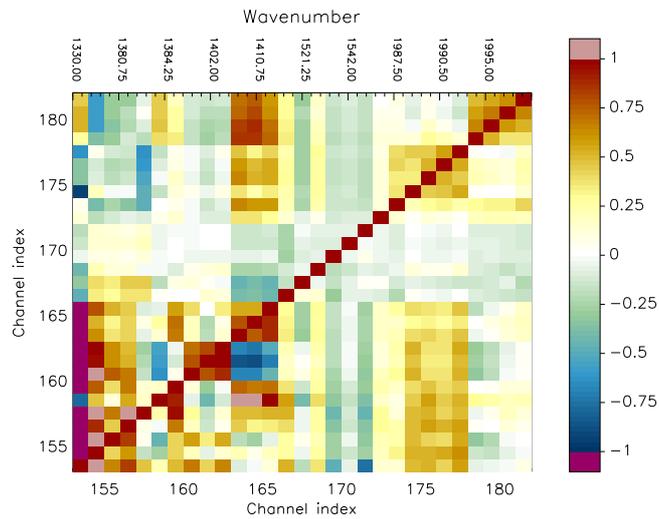


Figure 7: Diagnosed observation error correlation matrix for water vapour channels used in the OPS.

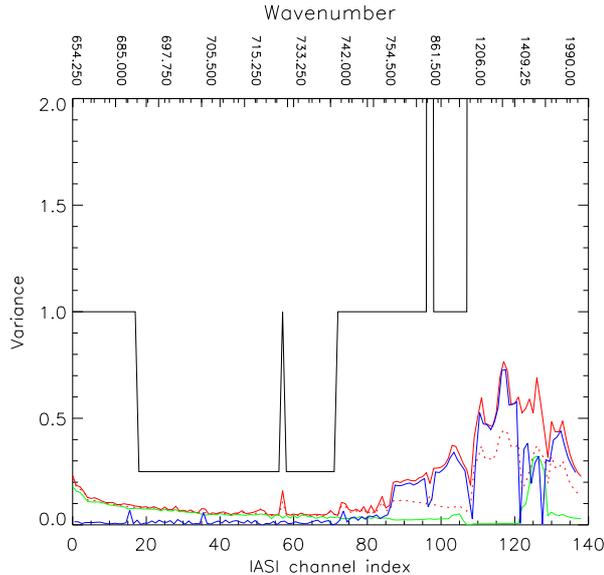


Figure 8: Operational error variances (black line), diagnosed error variances (red line) and the first off-diagonal covariance (blue line), Hollingsworth-Lönnerberg diagnosed error variances (red dashed line) and instrument noise (green line).

iterative procedure for updating the error variances, as proposed in Desroziers *et al.* (2005), would make the error covariance matrix \mathbf{R} closer to symmetric as we moved closer to using the true error correlation matrix; this will however be influenced by the specification of the background error covariance matrix.

Looking explicitly at the error correlations in Figure 6 makes it easier to identify any strong structures. In the temperature sounding OPS channels indexed below 120, there is little identifiable correlation structure with the exception of two channels which are strongly correlated with their neighbours. These two channels were chosen for monitoring purposes and are spectrally closer their neighbours than the other temperature sounding channels (0.25cm^{-1} compared with 0.5cm^{-1}). In the water-vapour sensitive OPS channels there is some significant correlation structure. However, there does not appear to be any consistent correlation between channels with similar properties. This can be seen more clearly in Figure 7 which provides a close up of the error correlation structure for the water vapour sensitive channels. There are a few small blocks of correlation structure corresponding to channels very close together on the spectrum (Figure 1), but the majority of correlation structure appears incoherent.

We can conclude that when the IASI observations are analysed using 1D-Var statistics, the errors are largely uncorrelated with the exception of a small number of neighbouring channels sensitive to water vapour. Figure 4 showed the closeness of the derived error variances to the instrument noise, suggesting that this plus some forward model error was the main contributor to the overall observation error in 1D-Var. We can therefore surmise that these error sources are largely independent with IASI channel choice.

3.2 4D-Var assimilation

We now calculate the observation error covariances using the analysis innovations derived from the 4D-Var assimilation of IASI data.

Figure 8 shows the observation error variances used in 4D-Var (black line), the instrument noise (green line), the error variances (red line) and first off-diagonal (blue line) from the symmetrised matrix diagnosed using the Desroziers *et al.* method, $\mathbf{R}_{sym} = \frac{1}{2}(\mathbf{R} + \mathbf{R}^T)$, and the error variances diagnosed using the modified Hollingsworth-Lönnerberg method as described in Section 2.4 (red dashed line). The

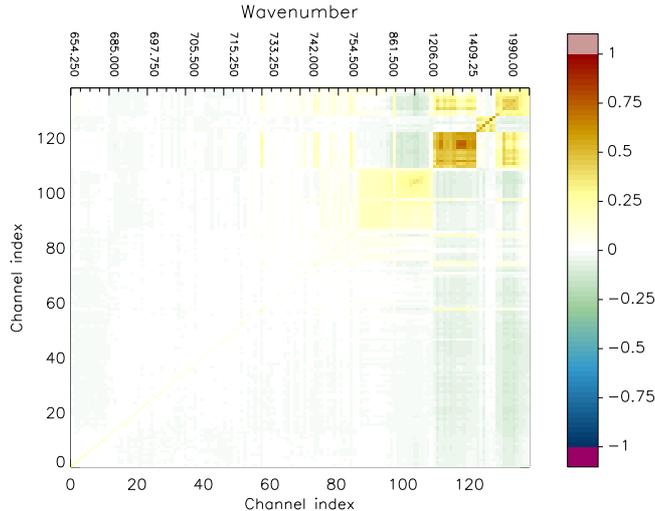


Figure 9: Diagnosed observation error covariance matrix for VAR.

values out of range are for channels where the operational error variance is 4.0. For all channels the diagnosed error variances from the Desroziers *et al.* method are significantly smaller than those being used operationally, implying that the error variances are heavily overestimated. However, the size of the first off-diagonal covariance value indicates why this overestimation is necessary. For the majority of the window and water vapour channels used in the 4D-Var assimilation, the first off-diagonal covariance value is very close in size to the diagonal variance value; therefore ignoring this value and other off-diagonal elements will result in the incorrect over-weighting of the observations in the analysis. We conclude that it is necessary to inflate the error variances if the large off-diagonal covariances are ignored.

Interestingly, the diagnosed error variances are significantly larger than the instrument noise for surface and water vapour sensitive channels; for some water vapour sensitive channels, the instrument noise is less than a fifth of the total error variance. We can conclude that instrument noise is not the main source of error in these channels when IASI observations are assimilated in 4D-Var. For much of the temperature sounding channels, the instrument noise appears to make up the main contribution to the diagnosed error, as with the 1D-Var results.

Comparing the error variances diagnosed with the Desroziers *et al.* technique with those derived using the Hollingsworth-Lönnerberg method, we observe that the Hollingsworth-Lönnerberg method produces very similar error variances for the temperature sounding channels, and slightly smaller error variances for the window and water vapour channels. Although the Hollingsworth-Lönnerberg variances are smaller in these latter channels, the magnitudes are comparable with those derived using the Desroziers technique and are much smaller than those currently being used. These results give us confidence in the robustness of the Desroziers results.

Figure 9 shows the diagnosed observation error covariances for the 139 channels used in 4D-Var. There are four significant block structures of covariance centred around the diagonal: the first is for the window channels, and the latter three for channels sensitive to water vapour in different parts of the spectrum. These four groups of channels are easily identifiable on the typical IASI spectrum shown in Figure 2. The block error structure suggests strongly correlated errors within the window and water vapour channels in similar parts of the spectrum.

The strong block structure is even more clearly visible in the diagnosed observation error correlation matrix shown in Figure 10. As suggested by the previous figure, very strong error correlations are present within the window and water vapour channels in similar parts of the spectrum. Figure 11 shows a closer look at the three blocks of water vapour sensitive channels. Many of the most strongly correlated channels

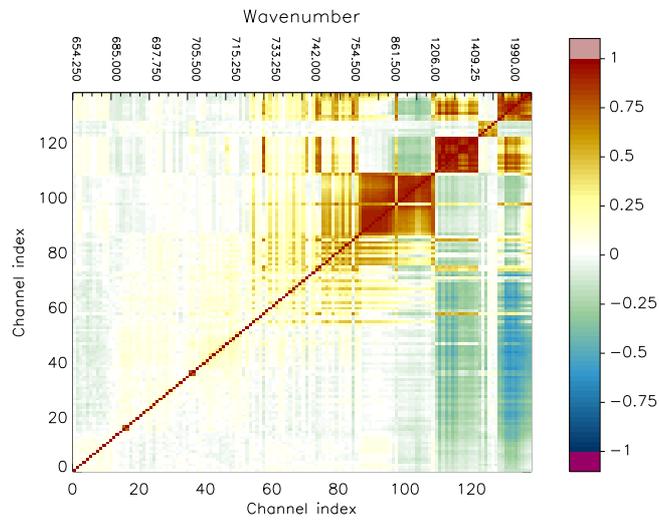


Figure 10: Diagnosed observation error correlation matrix for VAR.

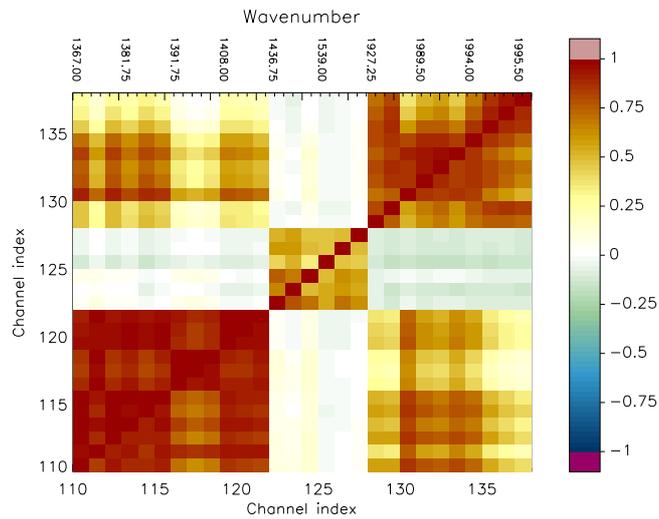


Figure 11: Diagnosed observation error correlation matrix for water vapour channels used in VAR.

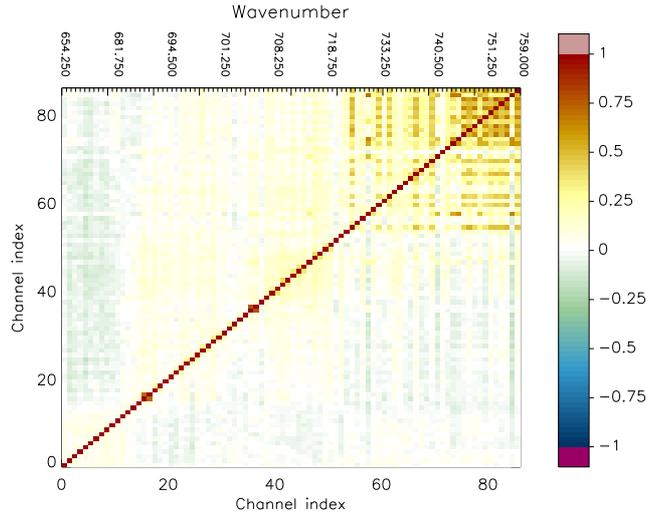


Figure 12: Diagnosed observation error correlation matrix for the temperature sounding channels used in VAR.

have error correlations larger than 0.75. We also observe that not all water vapour channels are highly correlated; it is channels in similar parts of the spectrum (i.e, longwave and shortwave) and with similar sensitivity to water vapour that are the most strongly correlated. Finally, bands of strong correlation are also visible surrounding the first, and largest, block structure in Figure 10. These bands are for channels in the temperature sounding part of the spectrum with a higher than average sensitivity to water vapour.

Although correlations are largest in those channels highly sensitive to water vapour or surface properties, a weaker level of correlation is also present in the channels used in temperature sounding. Figure 12 shows two fainter blocks of correlation centred on the diagonal for the upper temperature sounding channels (Var channels 0-50), in addition to bands of correlations for the water vapour sensitive lower temperature sounding channels (Var channels 51-86). Many channels within these two blocks are spectrally close to each other, and therefore we would expect some level of error correlation structure. The differences in measurements between these channels can be used to capture fine scale information on humidity and temperature profiles; it is therefore desirable to include even a weak level of correlation structure in an attempt to lower the operational error variances and hence retain more information.

The results from 4D-Var are similar to those found in [Bormann *et al.* \(2010b\)](#). For channels strongly sensitive to the surface or water vapour, observation errors larger than the instrument noise, but considerably smaller than the operational error variances, were diagnosed using the Desroziers method amongst others. Strong inter-channel error correlations were also present between these channels, which were not seen to the same extent in the temperature sounding band.

3.3 Comparison of 1D-Var and 4D-Var results

As described in Section 2.2, IASI observations in the Met Office system are initially processed in a 1D-Var retrieval before their use in the forecast model using a 4D-Var assimilation method. We are interested in how the observation error covariances vary with these two procedures.

When IASI observations are processed in 1D-Var, we expect the observation errors to be largely made up of forward model error and instrument noise, with some vertical error of representation. The diagnosed observation error variances were shown to be very close to the instrument noise for most of the channels used, implying that most of the error comes from this source, with a small contribution from the forward model error and vertical errors of representation (Figure 4).

In a 4D-Var assimilation, errors of representation in the horizontal also contribute to the observation error covariance matrix. When the IASI observations are processed in the 4D-Var retrieval, the diagnosed observation error variances were much larger than the instrument noise in channels sensitive to surface properties and water vapour (Figure 8). This suggests that instrument noise is no longer the main contributing factor to the errors in these channels, as in 1D-Var. We attribute most of the remaining error to horizontal errors of representation which contribute in a 4D-Var analysis but not in a 1D-Var retrieval.

Comparing the derived error variances with those used operationally, we observe that the overestimation of the operational error variances in 4D-Var was much larger than that seen in the 1D-Var results. In 1D-Var the overestimation was approximately double the diagnosed error variances (Figure 4), while in 4D-Var the overestimation is up to 8 times the diagnosed error variance for the water vapour sensitive channels (Figure 8). Also, comparing Figure 9 to Figure 5, we see that both the variances and covariances are notably larger in the 4D-Var error covariance matrix; the largest covariances are greater than 0.75 compared to less than 0.20.

The observation errors calculated from the 1D-Var statistics were shown to be largely uncorrelated (Figure 6), with the exception of a handful of spectrally close water vapour sensitive channels. However, when statistics from 4D-Var were used in error diagnostics, strong correlations were diagnosed for channels sensitive to water vapour and surface properties (Figure 10). Many of the strongly correlated channels had correlations larger than 0.75; the error correlations in 1D-Var were rarely at this level.

However, not all the water vapour sensitive channels used in the 4D-Var assimilation had highly correlated errors; it was channels with similar spectral properties and sensitivity to water vapour that were the most strongly correlated. This again is different to the 1DVar results, where there was little consistency of correlation within the longwave and shortwave water vapour channels. For the temperature sounding channels, as with the water vapour sensitive channels, the consistency and level of error correlation structure is far greater in the 4D-Var results than in those derived from 1D-Var statistics.

4 Summary and conclusions

In order to model successfully observation error correlations, an accurate knowledge of the true correlation structure is needed. This structure varies with observation type, and as shown here, assimilation system. In this paper we present the initial results from using a post-analysis diagnostic derived from variational data assimilation theory to calculate IASI observation error correlations when the data is used in the 1D-Var and 4D-Var assimilation systems at the Met Office.

The need for an accurate specification of error correlation structure when assimilating high-resolution satellite data has become increasingly important in recent years. A number of methods exist for quantifying error correlations, and have been applied for different data types and operational frameworks. In this paper we use the Desroziers technique of error approximation, and compare the diagnosed error variances and correlations with those used in the current operational system.

When IASI observations are analysed in the Met Office 1D-Var retrieval system, the diagnosed error variances are approximately half the size of the current operational errors and are very close in size to the instrument noise. The errors are also largely uncorrelated with the exception of a handful of neighbouring water vapour sensitive channels. We conclude that current operational errors are being overestimated, and that uncorrelated instrument noise is the main contributor to the observation error.

When the IASI observations are processed in the Met Office 4D-Var retrieval system we again found that the observation errors were being overestimated operationally. However, using this assimilation system resulted in an overestimation of up to eight times in channels highly sensitive to water vapour. The diagnosed errors were noticeably larger than the instrument noise for surface and water vapour sensitive channels, suggesting that another error source had a significant contribution. Because horizontal errors of representation are expected in 4D-Var processing but not in 1D-Var processing, we attribute the majority of this extra error to this source.

Also, the diagnosed errors from 4D-Var were found to contain significant correlation structure. The findings can be broken down into three main features: (i) a strong consistent block correlation structure in the water vapour and surface sensitive channels, (ii) bands of correlation in the temperature sounding channels sensitive to water vapour, and (iii) a weaker but still significant level of block correlation structure in the upper temperature sounding channels. We can conclude that in the 4D-Var assimilation system observation errors are again being overestimated, and strong correlations within the water vapour and surface sensitive channels are being ignored.

Finally, there were several differences in the diagnosed errors from the 1D-Var and 4D-Var statistics. These can be summarised as:

- the source of the observation error was predominantly instrument noise in 1D-Var, while in 4D-Var additional horizontal errors of representation made a significant contribution
- the overestimation of the error variances was much larger in 4D-Var for many of the channels used
- the error correlation structure is much stronger and more consistent between channels with similar spectral properties in 4D-Var.

We can conclude that the assimilation system used to process the observations has a significant impact on the size and structure of the observation error correlations.

We have been able to draw some strong conclusions from the findings above, but it is important to note that the Desroziers *et al.* diagnostic is not unflawed. The lack of symmetry in the diagnosed error covariance and correlation matrices is a visual indicator that some of the assumptions used in the derivation of the diagnostic have been violated. We are knowingly using an incorrect error covariance matrix in both assimilation systems, and we suspect that the operational background error covariance matrices used are also different to the truth. Nevertheless, simple model experiments have shown that the method yields good results even when violating these assumptions (Stewart, 2010, chapter 7), and results using other diagnosis methods are similar e.g., the Hollingsworth-Lönnberg method considered in this paper; Bormann and Bauer (2010a); Bormann *et al.* (2010b).

Other potential issues are channel selection choice and observation sampling. The work above was undertaken when there was a fixed IASI channel selection for 1D-Var and 4D-Var; currently the channel selection choice is adaptive and dependent on additional quality control related to cloudy field-of-views. This does not affect the use of the diagnostic nor the substance of the results presented here; it only requires a modification of the application. Also, the statistics used in the calculation of the covariance matrices were taken from a global sample, and hence local effects could be masked. But it is not proposed that we use these exact diagnosed error structures in the Met Office assimilation system; instead that they provide the motivation and framework for future investigations.

When the work in this paper was undertaken, the operational treatment of inter-channel IASI observation errors at the Met Office was to assume independent inter-channel errors in both 1D-Var and 4D-Var. The findings from this paper challenges the validity of this assumption. Motivated by these results, Weston (2011) describes the work at the Met Office currently underway on using error covariances diagnosed using the Desroziers method in the 4D-Var assimilation system. One important question being addressed is whether it is possible to include all diagnosed correlation structure, or if an approximation to the derived structure is necessary for computational purposes. It is expected that a better understanding, and hence representation, of the observation error structures in 1D-Var and 4D-Var will improve the use of IASI data in the Met Office processing systems.

References

Bormann N, Bauer P. 2010. ‘Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. I: Methods and application to ATOVS data’. *Q.J.R.Meteorol.Soc.***136**:1036–1050.

- Bormann N, Collard A, Bauer P. 2010. ‘Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. II: Application to AIRS and IASI data’. *Q.J.R.Meteorol.Soc.* **136**:1051–1063.
- Bouttier F, Kelly G. 2001. ‘Observation-system experiments in the ECMWF 4D-Var data assimilation system’. *Q.J.R.Meteorol.Soc.* **127**:1469–1488.
- Collard A.D. 2004. ‘On the choice of observation errors for the assimilation of AIRS brightness temperatures: A theoretical study’. *ECMWF Technical Memoranda*. **AC/90**.
- Collard A.D. 2007. ‘Selection of IASI channels for use in Numerical Weather Prediction’. *Q.J.R.Meteorol.Soc.* **133**:1977–1991.
- Collard A.D, McNally A.P. 2009. ‘The assimilation of Infrared Atmospheric Sounding Interferometer radiances at ECMWF’. *Q.J.R.Meteorol.Soc.* **135**:1044–1058.
- Dando M.L, Thorpe A.J, Eyre J.R. 2007. ‘The optimal density of atmospheric sounder observations in the Met Office NWP system’. *Q.J.R.Meteorol.Soc.* **133**:1933–1943.
- Dee D.P, da Silva A.M. 1999. ‘Maximum-likelihood estimation of forecast and observation error covariance parameters. Part 1: Methodology’. *Mon. Wea. Rev.* **127**:1822–1834.
- Desroziers G, Berre L, Chapnik B, Poli P. 2005. ‘Diagnosis of observation, background and analysis-error statistics in observation space’. *Q.J.R.Meteorol.Soc.* **131**:3385–3396.
- Desroziers G, Ivanov S. 2001. ‘Diagnosis and adaptive tuning of observation error parameters in variational assimilation’. *Q.J.R.Meteorol.Soc.* **127**:1433–1452.
- Garand L, Heilliette S, Buehner M. 2007. ‘Interchannel error correlation associated with AIRS radiance observations: Inference and impact in data assimilation’. *J.Appl.Meteorol.* **46**:714–725.
- Hilton F, Atkinson N.C, English S.J, Eyre J.R. 2009. ‘Assimilation of IASI at the Met Office and assessment of its impact through observing system experiments’. *Q.J.R.Meteorol.Soc.* **135**:495–505.
- Hollingsworth A, Lönnberg P. 1986. ‘The statistical structure of short-range forecast errors as determined from radiosonde data. Part 1: The wind field’. *Tellus*. **38A**:111–136.
- Joo S, Eyre J.R, Marriott R. 2012. ‘The impact of Metop and other satellite data within the Met Office global NWP system using an adjoint-based sensitivity method’. Met Office Forecasting Research Technical Report **562**.
- Liu Z.-Q, Rabier F. 2003. ‘The potential of high-density observations for numerical weather prediction: A study with simulated observations’. *Q.J.R.Meteorol.Soc.* **129**:3013–3035.
- Matricardi M, Chevallier F, Kelly G, Thepaut J.-N. 2004. ‘An improved general fast radiative transfer model for the assimilation of radiance observations’. *Q.J.R.Meteorol.Soc.* **130**:153–173.
- Rabier F, Fourrié N, Chafai D, Prunet P. 2002. ‘Channel selection methods for Infrared Atmospheric Sounding Interferometer radiances’. *Q.J.R.Meteorol.Soc.* **128**:1011–1027.
- Rabier F, Bouchard A, Faccani C, Fourrié N, Gerard E, Guidard V, Guillaume F, Karbou F, Moll P, Payan C, Poli P, Puech D. 2009. ‘Global impact studies at Meteo-France’. Available from www.crn.meteo.fr/gmap/sat/Glocal-Studies-MF.pdf.
- Rawlins F, Ballard S.P, Bovis K.J, Clayton A.M, Li D, Inverity G.W, Lorenc A.C, Payne T.J. 2007. ‘The Met Office global four-dimensional variational data assimilation scheme’. *Q.J.R.Meteorol.Soc.* **133**:347–362.
- Saunders R, English S, Francis P, Rayer P, Brunel P, Kelly G, Bauer P, Salmond D, Dent D. 2005. ‘RTTOV-8 the latest update to the RTTOV model’. Proc. International TOVS Study Conference XIV, Beijing, China, 25–31 May 2005. Available from <http://cim.ss.ssec.wisc.edu/itwg/itsc/itsc14/proceedings>.

- Stewart L.M. 2010. ‘Correlated observation errors in data assimilation’. PhD thesis, University of Reading. Available from <http://www.reading.ac.uk/maths-and-stats/research/maths-phdtheses.aspx>.
- Stewart L.M, Dance S.L, English S, Eyre J, Nichols N.K. 2009. ‘Observation error correlations in IASI radiance data’. Mathematical Report Series 1/2009. Univ.Reading:Reading, UK. Available from <http://www.rdg.ac.uk/maths/research/maths-report-series.aspx>.
- Stewart L.M, Dance S.L, Nichols N.K. 2008. ‘Correlated observation errors in data assimilation’. *Int.J.Numer.Meth.Fluids*. **56**:1521-1527.
- Stewart L.M, Dance S.L, Nichols N.K. 2012. ‘Data assimilation with correlated observation errors: analysis accuracy with approximate error covariance matrices’. *submitted*. Preprint available from <http://www.reading.ac.uk/maths-and-stats/research/maths-preprints.aspx>.
- Weston P. 2011. ‘Progress towards the implementation of correlated observation errors in 4D-Var’. Met Office Forecasting Research Technical Report **560**.