] and it has been suggested that part of the correlation comes from representativity error rather than the instrument error or errors in the observation operator [Stewart, , Weston,]. Until recently it has been assumed that it is too expensive to include correlated observation error matrices in assimilation schemes and that it is only feasible to use a diagonal observation error covariance matrix. The effect of correlated error is reduced by using techniques such as observation thinning [Lahoz et al.,] or superobbing [Daley, 99], and variance inflation [Hilton et al., 9, Whitaker et al., 8]. Calculations are also simplified by assuming that the observations errors are the same at each model level [Dee and Da Silva, 999]. Efforts are being made to find methods of reducing the cost of using correlated observation error matrices [Stewart et al., a, Stewart, , Healy and White, 5, Fisher, 5]. Once these methods are in place it will be important to have accurate estimates of the covariance matrices, as these are required to obtain the optimal estimate from any data assimilation system [Houtekamer and Mitchell, 5, Stewart et al., 8]. It is therefore important to understand how to estimate representativity error.

Despite the di culties in calculating correlated error, there have been some attempts. The Hollingsworth and Lönnberg method [Hollingsworth and Lönnberg, 986] has been used to calculate the statistics of the innovations. A method proposed by Desroziers et al. [5] makes use of information from the first guess and analysis departures and yields an approximation to the observation error covariance matrix. Once the innovation statistics or the observation error covariances have been calculated, the background and/or instrument error terms can be subtracted to leave an approximation of forward model error for specific observing instruments. Other methods [Daley, 99 , Liu and Rabier,] assume that observations can be written as a projection of a high resolution model state on to observation space with the representativity error being the difference between this high resolution projection and the model representation of the observation. Many of these approaches yield a static approximation of representativity error, but Janjic and Cohn [6] show that it is state dependent and correlated in time.

Work has been carried out by Stewart [ ], Stewart et al. [9, b] and Bormann and Bauer [ ] to calculate estimates of the full observation error covariance matrix. They show that the observation error covariance matrices for observing instruments such as IASI, AMSU-A, HIRS and MHS contain significant correlations. In particular the correlations for the humidity channels are more significant than those for temperature. The calculated matrices contain contributions from both the representativity error and the instrument error. Due to the complex nature of observation error statistics it is not known what portion of the error is representativity error. As humidity fields contain smaller scale features than temperature fields, it is possible that it is the representativity error that contributes to the more significant error correlations.

In this paper we investigate whether the significant correlations are representativity error. We calculate the representativity error for temperature and humidity data over the UK using a method described by Daley [99] and Liu and Rabier [ ]. We investigate whether representativity error is more significant for humidity than temperature, and whether one approximation of representativity error is suitable for all pressure levels.

In section we describe the method used for calculating representativity error. We then describe the model and available data in section . Our experimental design is given in section 4 and we

2 Representativity Error

2.1 Representativity Error

Forward model error,

$$\epsilon^{\mathrm{R}} = \mathbf{y} - \mathbf{H}(\mathbf{x}\ ), \qquad (\ )$$

is the difference between the noise free observation vector, $\mathbf{y}$, of length $\mathbf{p}$ and the mapping of the exact model state vector, $\mathbf{x}$, of length $\mathbf{N}$ into observation space using the possibly non-linear observation operator $\mathbf{H}$. The noise free observation vector is a theoretical construct that represents an observation measured by a perfect observing instrument, i.e. with no instrument error. It is related to the actual measurement via the equation

$$\mathbf{y}^{\mathrm{o}} = \mathbf{y} + \epsilon^{\mathrm{I}}, \qquad (\ )$$

where $\mathbf{y}^{\mathrm{o}}$ is the observation vector and ϵ^{I} is the instrument error.

The covariance of the forward model error $\mathbf{E}[\epsilon^{\mathbf{R}}\epsilon^{\mathbf{R}^{\mathbf{T}}}] = \mathbf{R}^{\mathrm{H}}$ is included in the observation error covariance matrix $\mathbf{R} = \mathbf{R}^{\mathrm{H}} + \mathbf{R}^{\mathrm{I}}$, where $\mathbf{R}^{\mathrm{I}} = \mathbf{E}[\epsilon^{\mathbf{I}}\epsilon^{\mathbf{I}^{\mathbf{T}}}]$ is the instrument error covariance matrix. To calculate the representativity errors in this paper we use a method defined by Daley [99] and Liu

3 The Model and Data

In this study we calculate representativity error for both temperature and specific humidity over the UK. The calculation of representativity error by the method of Liu and Rabier [ ] assumes that the actual state can be taken from a high resolution model. As our actual state we take data from the Met O ce UKV model. The UKV model is a variable resolution model that covers the UK; at its highest resolution the domain has .5km grid boxes. The model data used is at this resolution. The boundary data for the model is interpolated to the .5km grid from the 4km resolution regional model.

In this work we calculate representativity error using the assumption that the model state is a truncation of high resolution data. For the majority of our experiments we chose a truncation factor that gives a model grid spacing equivalent to the grid spacing that is used in the Met O ce NAE model. The Met O ce NAE model has a grid spacing of km (in mid-latitudes) and covers Europe and the North Atlantic.

3.1 The data available

We use temperature and humidity data over the UK available for two cases. The first case, Case , consists of data from 7 August 7 at times 8 UTC, 9 UTC and 9 UTC on an area over the southern UK that covers $-$. 4°W to .7 °E and 49. 8°N to 5 . 6°N. In this case there are partly clear skies with convection occurring over the south east [Eden, 7]. The second set of data, Case , is from 6 September 8 at 4 UTC, 4 UTC and 5 UTC and covers $-$5. °W to . °E and 5 .5°N to 56. °N. In this case a deep depression is tracking slowly east-northeast across England [Eden, 8]. The data is available on a × grid of latitude and longitude lines at each of 4 pressure levels.

3.2 Creating samples from the data

There are some limitations to the data. Data near the boundar

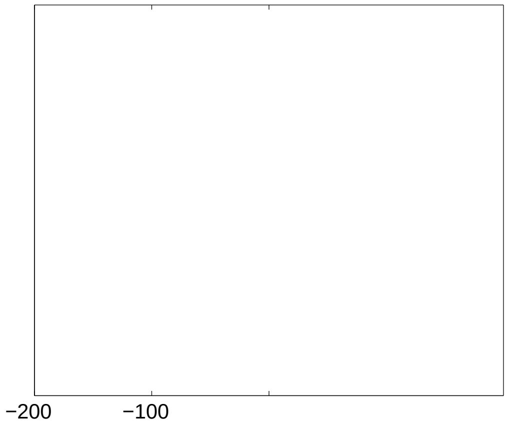
Table 1 – Variances for the true state

	Temperature (K^2)	log(Specific Humidity) ($(kg/kg)^2$)
Case	.66 8	. 8
Case	. 9 4	. 78

3.3 Data processing

To create surrogate samples from each available sample the data must be detrended. Detrending gives data on a homogeneous field; this is required by our chosen method for calculating representativity error. Data is detrended by removida be46(g)-r.99849(t)-5.86(fd)-6(t)- -6(l) (i) (d)- (e)- 54(u)-6(s)-

Table 2 – Representativity error (RE) variances for Case . The values given in brackets are a comparison


−200
−100

Acknowledgments

The authors thank Laura Stewart for providing the data. This work was funded by NERC as part of the National Center for Earth Observation and the Met O ce through a CASE studentship.

A Representativity error variance

Here we show why the representativity error variance does not alter when calculated with different numbers of observations. We do this by considering the calculation of the elements of the representativity error matrix.

Representativity error is calculated using Eq. ( ). The matrices are:

- $\mathbf{F}$ with $\mathbf{j} = \ \ldots \mathbf{p}$ and $\mathbf{k} = \ \ldots \mathbf{M}$. Elements defined as in Eq. (4).
- $\mathbf{F}$ with $\mathbf{j} = \ \ldots \mathbf{p}$ and $\mathbf{k} = \ \ldots \mathbf{M_m}$. Elements defined as in Eq. (4).
- $\mathbf{W}$ with $\mathbf{j} = \ \ldots \mathbf{M}$ and $\mathbf{k} = \ \ldots \mathbf{M}$. [illegible]

$$\mathbf{R}^{H}_{j,j} = \sum_{l=1}^{M} \exp\left(\frac{2ijl\pi}{p}\right)\hat{w}_l \hat{s}_l \hat{w}_l \exp\left(\frac{-2ijl\pi}{p}\right), \quad (8)$$

$$= \sum_{l=1}^{M} \hat{w}_l \hat{s}_l \hat{w}_l. \quad (9)$$

This does not depend on p and hence we do not expect the variance to change when we use different numbers of observations to calculate representativity error.

We now show that the correlation structure depends only on the distance between observations and not the number of observations.

Our model has N_m grid points separated by a spacing Δx and we have p observations. The distance between consecutive observations is $\frac{N_m \Delta x}{p}$. Suppose we have two observations separated by a distance d and assume that these are observation j and observation k. Then we have

$$d = \frac{(jk)(N_m \Delta x)}{}$$

N. Bormann, A. Collard, and P. Bauer. Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. II :application to AIRS and IASI data. *Quarterly Journal of the Royal Meteorological Society*, 6: 5 – 6 , .

R. Daley. *Atmospheric Data Analysis*. Cambridge University Press, 99 .

R. Daley. Estimating observation error statistics for atmospheric data assimilation. *Ann. Geophysicae*, :6 4–647, 99 .

D. P. Dee and A. M. Da Silva. Maximum-likelihood estimation of forecast and observation error covariance parameters. part I: Methodology. *Monthly Weather Review*, 7: 8 – 84 , 999.

G. Desroziers, L. Berre, B. Chapnik, and P. Poli. Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, : 85 – 96, 5.

P. Eden. Weather log: August 7. *Weather*, 6 :i – iv, 7.

P. Eden. Weather log: September 8. *Weather*, 6 :i – iv, 8.

M. Fisher. Accounting for correlated observation error in the ECMWF analysis. Technical report, ECMWF, 5. MF/ 5 6.

S. B. Healy and A. A. White. Use of discrete Fourier transforms in the d-var retrieval problem. *Quarterly Journalof the Royal Meteorological Society*, :6 – 7 , 5.

F. Hilton, A. Collard, V. Guidard, R. Randriamampianina, and M. Schwaerz. Assimilation of IASI radiances at european NWP centres. *In Proceedings of Workshop on the assimilation of IASI data in NWP, ECMWF, Reading, UK, 6-8 May 2009*, pages 9–48, 9.

A. Hollingsworth and P. Lönnberg. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I:(y)]TJ/R .9sh

L. Stewart, S. Dance, and N. Nichols. Data assimilation with correlated observation errors: analysis accuracy with approximate error covariance matrices. Technical report, Univer-