Theoretical insight into diagnosing observation error correlations using observation-minus-background and observation-minus-analysis statistics

RMetS
Royal Meteorological Society

Theoretical insight into diagnosing observation error correlations using observation-minus-background and observation-minus-analysis statistics

J. A. Waller,[a]* S. L. Dance[a,b] and N. K. Nichols[a,b]

[a]*Department of Meteorology, School of Mathematical and Physical Sciences, University of Reading, UK*
[b]*Department of Mathematics and Statistics, School of Mathematical and Physical Sciences, University of Reading, UK*

*Correspondence to: J. A. Waller, Department of Meteorology, University of Reading, Earley Gate, PO Box 243, Reading, RG6 6BB, UK. E-mail: j.a.waller@reading.ac.uk

To improve the quantity and impact of observations used in data assimilation, it is necessary to take into account the full, potentially correlated, observation error statistics. A number of methods for estimating correlated observation errors exist, but a popular method is a diagnostic that makes use of statistical averages of observation-minus-background and observation-minus-analysis residuals. The accuracy of the results it yields is unknown as the diagnostic is sensitive to the difference between the exact background and exact observation error covariances and those that are chosen for use within the assimilation. It has often been stated in the literature that the results using this diagnostic are only valid when the background and observation error correlation length-scales are well separated. Here we develop new theory relating to the diagnostic. For observations on a 1D periodic domain we are able to the show the effect of changes in the assumed error statistics used in the assimilation on the estimated observation error covariance matrix. We also provide bounds for the estimated observation error variance and eigenvalues of the estimated observation error correlation matrix. We demonstrate that it is still possible to obtain useful results from the diagnostic when the background and observation error length-scales are similar. In general, our results suggest that when correlated observation errors are treated as uncorrelated in the assimilation, the diagnostic will underestimate the correlation length-scale. We support our theoretical results with simple illustrative examples. These results have potential use for interpreting the derived covariances estimated using an operational system.

Key Words: data assimilation diagnostic; correlated observation errors; innovation statistics

Received 2 March 2015; Revised 25 August 2015; Accepted 1 September 2015; Published online in Wiley Online Library

1. Introduction

Data assimilation techniques combine model states, known as forecasts or backgrounds, with observations, weighted by their respective errors, to provide a best estimate of the state, known as the analysis. The accurate representation of these error statistics is essential for obtaining an accurate estimate of the state. Until recently, research has predominately concentrated on how to estimate and represent the background error covariance matrix e.g. Bannister (2008). Observation errors have been assumed uncorrelated and often the data is thinned or 'superobbed' in an attempt to satisfy this assumption (Lorenc, 1981). However, with the desire and need to make better use of the observations, especially for high-resolution forecasting, the understanding and accurate representation of these statistics needs to be addressed.

Observation errors can be attributed to a number of different sources, some of which may be state dependent and dependent on the model resolution (Lorenc, 1986; Janjic and Cohn, 2006; Waller, 2013; Waller *et al.*, 2014a, 2014b). A number of methods exist for estimating the observation error covariances, but none are without fault. Previously, estimates of the observation error covariance matrix have been calculated using methods such as those proposed by Hollingsworth and Lönnberg (1986), Dee and Da Silva (1999) and Desroziers and Ivanov (2001). At present, a popular method is the diagnostic proposed in Desroziers *et al.* (2005) (see section 2 for a detailed discussion of this diagnostic). Initially proposed as a consistency check, the diagnostic uses the statistical average of observation-minus-background and observation-minus-analysis residuals to provide an estimate of the observation error covariance matrix. The diagnostic provides

an exact estimate of the observation error covariance matrix if the assumed background and assumed observation error statistics that are used in the assimilation are correct. In practice the statistics used in the assimilation will not be exact, but Desroziers *et al.* (2005) show that in this case the diagnostic may still be used to gain an estimate of the observation error variances and correlations. It is also shown that the result may be improved if successive iterations of the diagnostic are applied. As well as providing a diagnostic for estimating the observation error covariance matrix, Desroziers *et al.* (2005) provide a diagnostic for estimating the background error covariance matrix mapped into observation space. They also provide one diagnostic that does not rely on the background and observation error statistics that are used in the assimilation; this diagnostic calculates the statistical expectation of the observation-minus-background residuals and provides a result equal to the sum of the observation error statistics and the background error statistics mapped into observation space. This relation, first suggested by Hollingsworth and Lönnberg (1986), has been used recently to diagnose both background and observation errors e.g. Bormann and Bauer (2010) and Bormann *et al.* (2003). However, when estimating correlated errors, determining how to split the estimated quantity into observation and background errors is non-trivial.

Todling (2015) highlighted the limitations of the diagnostics of Desroziers *et al.* (2005) in an observing system simulation experiment framework. Despite the limitations the diagnostic has been successfully used in some studies to estimate observation error variances and correlations. It has been used in simple model experiments in both variational (Stewart, 2010) and ensemble (Li *et al.*, 2009; Miyoshi *et al.*, 2013) data assimilation systems and to estimate time varying observation errors (Waller *et al.*, 2014a). The diagnostic has also been applied to operational numerical weather prediction (NWP) observation types such as ATOVS, AIRS and IASI to calculate inter-channel error covariances (Bormann and Bauer, 2010; Bormann *et al.*, 2010; Stewart *et al.*, 2014; Weston *et al.*, 2014). When the correlated errors calculated using the diagnostic have been accounted for in the assimilation, it has been shown to lead to a more accurate analysis (Healy and White, 2005; Stewart, 2010; Stewart *et al.*, 2013), the inclusion of more observation information content (Stewart *et al.*, 2008) and improvements in the forecast skill score (Weston *et al.*, 2014). Indeed, Stewart *et al.* (2013) and Healy and White (2005) show that even the use of a crude approximation to the observation error covariance matrix may provide significant benefit.

In an operational setting the iteration of the diagnostic for estimating spatially correlated observation errors is not feasible as the use of correlated errors in assimilation systems is in its infancy. In many cases iterating the diagnostic will be costly and time consuming and may produce disappointing results due to the many assumptions that are already required to permit operational assimilation. In some cases the computational framework for including these correlated errors in the assimilation is not yet developed and this in itself is a necessary challenge to overcome if the correlated observation errors are to be used and the diagnostic is to be iterated.

As the popularity of this diagnostic grows, it is important to have a better understanding of the results it produces. Theoretical results relating to the diagnostic under some simplifying assumptions have been previously published, both in the original manuscript of Desroziers *et al.* (2005) and in workshop proceedings (Ménard *et al.*, 2009; Desroziers *et al.*, 2009). These results relate to scalar cases or consider the estimation of variances or the convergence of the method under iteration. We discuss these results in further detail in section 4. In this work we develop new theoretical results relating to the diagnostic for observations on a 1D periodic domain and provide insight into how well the observation error covariances may be estimated using just one application of the diagnostic. We support these theoretical results with illustrative examples using some simple correlation matrices. We provide results that show what effect changes in the error statistics used in the assimilation have on the estimated observation error covariance matrix. From the results discussed in section 5, we show that:

- The estimated observation error variance decreases as assumed background error variance increases.
- The estimated observation error variance increases as assumed observation error variance increases.
- The power in the largest scales of the estimated observation error correlation matrix decreases as assumed background error variance increases.
- The power in the largest scales of the estimated observation error correlation matrix increases as assumed observation error variance increases.

The power in the largest scales can be obtained by considering the eigenstructure of the estimated matrix and provides some insight into the behaviour of the estimated correlation length-scale. These results provide an understanding of the diagnostic that can aid the interpretation of results when the diagnostic is used to estimate spatial correlations in an operational setting e.g. Waller *et al.* (2015). We note that in the results presented here, the statistical nature of the estimation is not considered since results are calculated directly and not from samples of the analysis and background residuals. When estimates are calculated in this way it is inevitable that further noise will be introduced.

We begin in section 2 by describing in detail the diagnostics of Desroziers *et al.* (2005). We present the diagnostic in spectral form in section 3. In section 4 we prove some new theoretical results and we illustrate and expand these results in section 5. We conclude in section 6.

2. The diagnostic of Desroziers *et al.* (2005)

2.1. Notation

Data assimilation techniques combine observations, $\mathbf{y} \in \mathbb{R}^p$, available at time t, with a model prediction of the state, the background, $\mathbf{x}^b \in \mathbb{R}^n$, which is often determined by a previous forecast. Here p and n denote the dimensions of the observation and model state vectors respectively. In the assimilation, the observations and background are weighted by their respective errors, using the background and observation error covariance matrices $\mathbf{B} \in \mathbb{R}^{n\times n}$ and $\mathbf{R} \in \mathbb{R}^{p\times p}$, to provide a best estimate of the state $\mathbf{x}^a \in \mathbb{R}^n$, known as the analysis. The calculation of the analysis requires a comparison of the background with the observations. To achieve this the background is mapped into the observation space using the, possibly non-linear, observation operator $\mathcal{H} : \mathbb{R}^n \to \mathbb{R}^p$. After an assimilation step the analysis is then evolved forward in time using a (possibly nonlinear) model, to provide a background at the next assimilation time.

2.2. The diagnostics

The diagnostics described in Desroziers *et al.* (2005) provide estimates of the observation and background error covariance matrices given that the analysis is determined using,

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + \tilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}} + \tilde{\mathbf{R}})^{-1}\mathbf{d}^{\mathrm{o}}_{\mathrm{b}}, \tag{1}$$

where $\mathbf{H}$ is the observation operator linearised about the current state and $\tilde{\mathbf{R}}$ and $\tilde{\mathbf{B}}$ are the assumed observation and background error statistics used to weight the observations and background in the assimilation. The background residual, also known as the innovation,

$$\mathbf{d}^{\mathrm{o}}_{\mathrm{b}} = \mathbf{y} - \mathcal{H}(\mathbf{x}^{\mathrm{b}}), \tag{2}$$

is the difference between the observation $\mathbf{y}$ and the mapping of the forecast vector, $\mathbf{x}^{\mathrm{b}}$, into observation space by the observation

operator $\mathcal{H}$. The analysis residuals,

$$\mathbf{d}_{\mathrm{a}}^{\mathrm{o}} = \mathbf{y} - \mathcal{H}(\mathbf{x}^{\mathrm{a}}), \tag{3}$$

$$\approx \mathbf{y} - \mathcal{H}(\mathbf{x}^{\mathrm{b}}) - \mathbf{H}\mathbf{K}\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}. \tag{4}$$

are similar to the background residuals, but with the forecast vector replaced by the analysis vector $\mathbf{x}^{\mathrm{a}}$.

We now restrict ourselves to the use of a linear operator $\mathbf{H}$ (this allows Eq. (4) to be written as an equality) and make the assumption that the forecast and observation errors are uncorrelated. Taking the statistical expectation of the outer product of the background residuals results in

$$E[\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}\mathbf{d}_{\mathrm{b}}^{\mathrm{o}\mathrm{T}}] = \mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R} = \mathbf{S}, \tag{5}$$

the sum of the observation and mapped background error statistics. Note that $\mathbf{S}$ is a direct estimate of the sum of the true statistics and does not depend on the background and observation error statistics used in the assimilation. However, in practice the estimate will be subject to sampling error.

Desroziers *et al.* (2005) show that an estimate of the observation error correlation matrix can be obtained by taking the statistical expectation of the outer product of the analysis and background residuals,

$$E[\mathbf{d}_{\mathrm{a}}^{\mathrm{o}}\mathbf{d}_{\mathrm{b}}^{\mathrm{o}\mathrm{T}}] = \tilde{\mathbf{R}}(\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}+\tilde{\mathbf{R}})^{-1}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R}) = \mathbf{R}^{\mathrm{e}}, \tag{6}$$

where $\mathbf{R}^{\mathrm{e}}$ is the estimated observation error covariance matrix and $\mathbf{B}$ and $\mathbf{R}$ are the exact background and observation covariance matrices. If the observation and forecast errors used in the assimilation are exact, $\tilde{\mathbf{R}} = \mathbf{R}$ and $\tilde{\mathbf{B}} = \mathbf{B}$, then

$$E[\mathbf{d}_{\mathrm{a}}^{\mathrm{o}}\mathbf{d}_{\mathrm{b}}^{\mathrm{o}\mathrm{T}}] = \mathbf{R}. \tag{7}$$

A further diagnostic makes use of the background residual and the difference between the mapping of the background and analysis into observation space $\mathbf{d}_{\mathrm{b}}^{\mathrm{a}} = \mathbf{H}(\mathbf{x}^{\mathrm{a}}) - \mathbf{H}(\mathbf{x}^{\mathrm{b}})$. This allows the estimation of $\mathbf{H}\mathbf{B}^{\mathrm{e}}\mathbf{H}^{\mathrm{T}}$,

$$E[\mathbf{d}_{\mathrm{b}}^{\mathrm{a}}\mathbf{d}_{\mathrm{b}}^{\mathrm{o}\mathrm{T}}] = \mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}+\tilde{\mathbf{R}})^{-1}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R}) = \mathbf{H}\mathbf{B}^{\mathrm{e}}\mathbf{H}^{\mathrm{T}}. \tag{8}$$

Again $\mathbf{H}\mathbf{B}^{\mathrm{e}}\mathbf{H}^{\mathrm{T}} = \mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}}$ when the matrices $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{R}}$ used in the analysis update Eq. (1) are exact.

It has been shown by Desroziers *et al.* (2005) that using Eq. (6) it is possible to obtain a reasonable estimate of $\mathbf{R}^{e}$ even if the matrices $\tilde{\mathbf{R}}$ and $\tilde{\mathbf{B}}$ used in the assimilation are not correctly specified, and successive applications of the diagnostic and assimilation scheme may be applied to converge to a solution. Ménard *et al.* (2009) prove, in the case of scalar observation and background error variances, that if the observation error variance is unknown and the background error variance is known then successive applications of Eq. (6) will lead to the convergence of the observation error variance to the exact value. Similarly if the observation error variance is known, then an unknown background error variance may be obtained by iterating Eq. (8). However, in the case where both the observation and background error variances are unknown, it is shown that the iteration of either Eq. (6) or Eq. (8) will converge to a solution, but one that does not match the exact statistics. Ménard *et al.* (2009) also show, again in the scalar case, that if both variances are iterated concurrently then the diagnostics converge in one iteration to a solution, possibly incorrect, that depends on the assumed error variances. The results may or may not be close to the true values, but the estimate cannot be improved. Ménard *et al.* (2009) then extend these results to a system with periodic 1D domain; however, we note that the results shown in Ménard *et al.* (2009) do not hold in the multi-dimensional case as an application of the diagnostic changes both the variances and correlations in the observation error covariance matrix, a process which is overlooked in equation (46) of Ménard *et al.* (2009).

Desroziers *et al.* (2005) show for the multi-dimensional case that the method may be used to estimate error variances and correlations, but state that it 'appears that the adjustment of background and observation error variances is only relevant if those errors have different structures'. As a result it is often stated that the method will not yield an accurate result if the scales in the background and observation error statistics are similar (Bormann and Bauer, 2010; Bormann *et al.*, 2010; Stewart *et al.*, 2014; Weston *et al.*, 2014). However, it is actually the convergence of the iterations that may be slow or even fail if the scales in the true observation and background or assumed observation and background error covariance matrices are proportional. Although this scale separation causes problems for the iteration procedure, it may not result in the failure of the diagnostic. Previously in Waller *et al.* (2014a), using twin experiments, the diagnostic produced accurate estimates of the observation error statistics even when the length-scales of the background and observation error statistics were similar. If one is fortunate with the underlying statistics or the choice of assumed statistics, then it is possible that the diagnostic will give a reasonable estimate of the observation error covariance matrix after one iteration and the failure of the convergence will not be an issue (see section 4 for an example).

In operational systems many of the assumptions made for using the diagnostics are violated. The $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{R}}$ used in the assimilation are not exact and the linearisation of the observation operator will introduce further error. To attempt to understand the impact of the incorrect $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{R}}$ used in the assimilation we will consider the performance of the diagnostic in an idealised framework that allows the spectral form of the diagnostics to be used.

3. The diagnostic in Fourier space

Desroziers *et al.* (2005, 2009) and Ménard *et al.* (2009) show that by making some simplifying assumptions the diagnostics may be written in Fourier space. We now assume that observations have uniform density over a 1D periodic domain and that the observation operator is the identity. We also require the true and assumed observation and background errors to be homogeneous with $\mathbf{R} = \rho\mathbf{C}_{\mathrm{r}}$, $\tilde{\mathbf{R}} = \tilde{\rho}\tilde{\mathbf{C}}_{\mathrm{r}}$, $\mathbf{B} = \beta\mathbf{C}_{\mathrm{b}}$ and $\tilde{\mathbf{B}} = \tilde{\beta}\tilde{\mathbf{C}}_{\mathrm{b}}$ where ρ, $\tilde{\rho}$, β and $\tilde{\beta}$ are the exact and assumed observation and background error variances and $\mathbf{C}_{\mathrm{r}}$, $\tilde{\mathbf{C}}_{\mathrm{r}}$, $\mathbf{C}_{\mathrm{b}}$ and $\tilde{\mathbf{C}}_{\mathrm{b}}$ are the circulant exact and assumed observation and background error correlation matrices. (A circulant matrix is a matrix where each row is determined by cyclically shifting the preceding row; Gray, 2006). Our approach also applies if we instead assume a non-identity observation operator so long as $\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}}$ and $\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}$ are circulant (in this case the matrices in following Eqs (9)–(12) would be of size $p \times p$), although here we choose the identity operator as it allows us to consider directly how changes in the assumed background error statistics alter the performance of the diagnostic. Under these assumptions the matrices share common eigenvectors and it is possible to write,

$$\mathbf{B} = \beta\mathbf{F}\boldsymbol{\Gamma}\mathbf{F}^{\mathrm{T}}, \tag{9}$$

$$\mathbf{R} = \rho\mathbf{F}\boldsymbol{\Lambda}\mathbf{F}^{\mathrm{T}}, \tag{10}$$

$$\tilde{\mathbf{B}} = \tilde{\beta}\mathbf{F}\tilde{\boldsymbol{\Gamma}}\mathbf{F}^{\mathrm{T}}, \tag{11}$$

$$\tilde{\mathbf{R}} = \tilde{\rho}\mathbf{F}\tilde{\boldsymbol{\Lambda}}\mathbf{F}^{\mathrm{T}}, \tag{12}$$

where $\mathbf{F}$ is a $n \times n$ orthogonal matrix of common eigenvectors such that $\mathbf{F}^{\mathrm{T}} = \mathbf{F}^{-1}$ and $\boldsymbol{\Gamma}$, $\boldsymbol{\Lambda}$, $\tilde{\boldsymbol{\Gamma}}$ and $\tilde{\boldsymbol{\Lambda}}$ are $n \times n$ diagonal matrices that contain the eigenvalues γ_k, λ_k, $\tilde{\gamma}_k$ and $\tilde{\lambda}_k$, of $\mathbf{C}_{\mathrm{b}}$, $\mathbf{C}_{\mathrm{r}}$, $\tilde{\mathbf{C}}_{\mathrm{b}}$ and $\tilde{\mathbf{C}}_{\mathrm{r}}$ respectively. Since the correlation matrices are positive definite and circulant, the eigenvalues are positive and can be found using a discrete Fourier transform with the eigenvectors being the discrete Fourier basis (Gray, 2006). These eigenvalues are ordered according to wave number. In this case the order of the eigenvalues has a relation to the length-scales in the correlation matrix, with the first eigenvalue relating to the eigenvector with the largest

Table 1. Summary of parameters used in Eqs (16)–(18).

Parameter	Exact quantity	Assumed quantity	Estimated quantity
Observation error variance	ρ	$\tilde{\rho}$	ρ^{e}
Eigenvalues of the observation error correlation matrix	λ_k	$\tilde{\lambda}_k$	λ_k^{e}
Background error variance	β	$\tilde{\beta}$	β^{e}
Eigenvalues of the background error correlation matrix	γ_k	$\tilde{\gamma}_k$	γ_k^{e}
Background residual variance	σ	–	–
Eigenvalues of the background residual correlation matrix	ω_k	–	–

length-scales. This ordering is not linked to the magnitude of the eigenvalues.

Substituting Eqs (9)–(12) into Eq. (5) allows the background residual covariance, given by Eq. (5), to be written as

$$\sigma\mathbf{\Omega} = \beta\mathbf{\Gamma} + \rho\mathbf{\Lambda}, \tag{13}$$

where σ is the background residual variance and $\mathbf{\Omega}$ is a matrix of eigenvalues ω_k of the background residual correlation matrix $\mathbf{S}$. The matrix $\mathbf{S}$ may be written in this form since it is the sum of circulant matrices, and therefore circulant itself.

Using Eqs (9)–(12) and (6) a relationship for $\mathbf{R}^{\mathrm{e}}$, defined as in Eq. (7), may be written as,

$$\rho^{\mathrm{e}}\mathbf{\Lambda}^{\mathrm{e}} = \tilde{\rho}\tilde{\mathbf{\Lambda}}(\tilde{\rho}\tilde{\mathbf{\Lambda}} + \tilde{\beta}\tilde{\mathbf{\Gamma}})^{-1}(\rho\mathbf{\Lambda} + \beta\mathbf{\Gamma}), \tag{14}$$

where ρ^{e} is the estimated variance and Λ^{e} is the matrix containing eigenvalues, λ_k^{e}.

Similarly a relationship for $\mathbf{B}^{\mathrm{e}}$, defined as in Eq. (8), may be written as,

$$\beta^{\mathrm{e}}\mathbf{\Gamma}^{\mathrm{e}} = \tilde{\beta}\tilde{\mathbf{\Gamma}}(\tilde{\rho}\tilde{\mathbf{\Lambda}} + \tilde{\beta}\tilde{\mathbf{\Gamma}})^{-1}(\rho\mathbf{\Lambda} + \beta\mathbf{\Gamma}), \tag{15}$$

where β^{e} is the estimated variance and Γ^{e} is the matrix containing eigenvalues, γ_k^{e}.

Hence for each wave number, $k = 0, \ldots, n-1$, we have the following relations,

$$\sigma\omega_k = \rho\lambda_k + \beta\gamma_k, \tag{16}$$

$$\rho^{\mathrm{e}}\lambda_k^{\mathrm{e}} = \tilde{\rho}\tilde{\lambda}_k\frac{\rho\lambda_k + \beta\gamma_k}{\tilde{\rho}\tilde{\lambda}_k + \tilde{\beta}\tilde{\gamma}_k}, \tag{17}$$

$$\beta^{\mathrm{e}}\gamma_k^{\mathrm{e}} = \tilde{\beta}\tilde{\gamma}_k\frac{\rho\lambda_k + \beta\gamma_k}{\tilde{\rho}\tilde{\lambda}_k + \tilde{\beta}\tilde{\gamma}_k}. \tag{18}$$

The majority of this manuscript focuses on describing the behaviour of Eq. (17) and deriving results based on this equation. However similar results may be derived for Eq. (18). For convenience we summarise the parameters used in these equations in Table 1.

When considering the result obtained from Eq. (14) or Eq. (15) the eigenvalues will be positive as the products, sums and inverses of positive definite matrices are also positive definite. The estimated covariance matrix will also be symmetric as circulant matrices are commutative and the product of commutative symmetric matrices is symmetric (Gray, 2006).

When the incorrect matrices, $\tilde{\mathbf{R}}$ and $\tilde{\mathbf{B}}$, are used in the assimilation, the diagnostics may contain errors in both the estimation of the variances and in the correlations. Considering the diagnostics in spectral space allows the misspecification of the power in the wave numbers of the correlation function to be assessed.

To prove some further results using Eqs (16)–(18) we make use of the eigenvalue relationship that states that the sum of the eigenvalues α_k of a matrix $\mathbf{A}$ is equal to the trace of that matrix (Golub and Van Loan, 2013), that is,

$$\sum_{k=0}^{n-1}\alpha_k = \mathrm{Tr}(\mathbf{A}). \tag{19}$$

In our case we are considering the eigenvalues of $n \times n$ correlation matrices with ones on the diagonal, and therefore the trace of such a matrix and hence the sum of the eigenvalues will be n.

By applying Eq. (19) to Eq. (16) we are able to show that the background residual variance,

$$\sigma = \rho + \beta \tag{20}$$

is the sum of the observation and background error variances. Furthermore the eigenvalues of the background residual correlation matrix may be written as,

$$\omega_k = \frac{\rho\lambda_k + \beta\gamma_k}{\rho + \beta}. \tag{21}$$

These are quantities that we are able to calculate using the diagnostic that provide valuable information on the unknown true error statistics.

As discussed in section 2, using Eqs (17) and (18) it is possible to show that even when the structure of the background and observation errors are similar the diagnostic may still produce a useful result. Let us assume that the underlying exact background and observation error statistics are identical and the assumed background and observation error statistics are identical, that is $\rho\lambda_k = \beta\gamma_k$ and $\tilde{\rho}\tilde{\lambda}_k = \tilde{\beta}\tilde{\gamma}_k$; it is not necessary that $\rho\lambda_k = \tilde{\rho}\tilde{\lambda}_k$. In this case just one application of Eqs (17) and (18) results in

$$\rho^{\mathrm{e}}\lambda_k^{\mathrm{e}} = \rho\lambda_k, \qquad \beta^{\mathrm{e}}\gamma_k^{\mathrm{e}} = \beta\gamma_k, \tag{22}$$

with the observation and background errors being estimated exactly. This will occur for any choice of exact and assumed observation and background error statistics provided that $\mathbf{R} = \mathbf{B}$ and $\tilde{\mathbf{R}} = \tilde{\mathbf{B}}$. This application of the diagnostic highlights two things: firstly that the diagnostic can produce good results when the observation and background error length-scales are equal, and secondly, if there is any knowledge that the underlying statistics are equal, then to obtain the best performance from the diagnostic, the assumed observation and assumed background error statistics should be chosen to be identical.

The fact that the diagnostic can yield a good result if the scales in the background and observation error statistics are similar can also be shown under other conditions. It is possible to show that $\mathbf{R} = \mathbf{R}^{\mathrm{e}}$ as soon as the background and observation error covariance matrices are proportional, that is $\mathbf{R} = g\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}$ and $\tilde{\mathbf{R}} = g\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}$, where the constants of proportionality g are equal. This result (highlighted by an anonymous referee) provides another set of circumstances under which the diagnostic produces an exact result even when the background and observation error length-scales are similar.

4. The assumption of uncorrelated observation errors

In most operational cases it is assumed that the observation error correlation matrix is diagonal, $\tilde{\mathbf{R}} = \tilde{\rho}\mathbf{I}$ where $\mathbf{I}$ is the identity matrix. As we are attempting to provide information on how the diagnostic may perform in operational cases we now assume that the assimilated observation error variance matrix is diagonal. With this additional assumption, Eq. (17) simplifies to

$$\rho^{\mathrm{e}}\lambda_k^{\mathrm{e}} = \frac{\rho\lambda_k + \beta\gamma_k}{1 + (\tilde{\beta}/\tilde{\rho})\tilde{\gamma}_k}. \tag{23}$$

Using Eq. (19) allows the estimated error variance to be written as,

$$\rho^{\mathrm{e}} = \frac{1}{n}\sum_{k=0}^{n-1}\frac{\rho\lambda_k + \beta\gamma_k}{1+(\tilde{\beta}/\tilde{\rho})\tilde{\gamma}_k}. \tag{24}$$

We see from Eq. (24) that as the assumed observation error variance, $\tilde{\rho}$ increases, the estimated observation error variance, ρ^{e}, increases. We also see that as the assumed background error variance, $\tilde{\beta}$ increases, the estimated observation error decreases. It is also possible to obtain a similar result in the case where the assumed observation errors are correlated by applying the trace rule to Eq. (17).

We now show how to obtain bounds for the estimated observation error variance, which provide further information on how the diagnostic behaves. To calculate the upper bound, we note that since $\tilde{\beta}, \tilde{\rho}, \tilde{\gamma}_k > 0$, the denominator in Eq. (24) is bounded from below by 1. To calculate the lower bound we consider the maximum value of the denominator. This occurs when $\tilde{\gamma}_k = \tilde{\gamma}_{\max}$, the maximum eigenvalue. A further application of Eq. (19) gives an inequality for the estimated error variance,

$$\frac{\sigma}{1+(\tilde{\beta}/\tilde{\rho})\tilde{\gamma}_{\max}} \leq \rho^{\mathrm{e}} \leq \sigma. \tag{25}$$

All quantities in this inequality are known: those denoted with a tilde are assumed in the assimilation and σ may be calculated using Eq. (20). We see that the upper bound is constant and equal to the sum of the true background and observation error variances. From Eq. (25) we see that the lower bound will behave differently depending on the behaviour of the assumed background and observation error variances; the lower bound increases as $\tilde{\rho}$ increases and decreases as $\tilde{\beta}$ increases. We illustrate this with examples in sections 5.3.2 and 5.3.3. In the limit of very large $\tilde{\rho}$ the estimated error variance will tend to the sum of the true background and observation error variances.

The eigenvalues of the estimated observation correlation matrix are given by dividing Eq. (23) by the estimated observation error variance, Eq. (24). The behaviour of the eigenvalues of the estimated correlation matrix will depend on the spectra for the assumed and exact matrices. If there is some knowledge of these it may be possible to say something further about the behaviour of the eigenvalues (see sections 5.3.2 and 5.3.3 for some examples). It is possible to put bounds on the eigenvalues using the bounds on the estimated observation error variance, Eq. (25),

$$\frac{\omega_k}{1+(\tilde{\beta}/\tilde{\rho})\tilde{\gamma}_k} \leq \lambda_k^{\mathrm{e}} \leq \omega_k\frac{1+(\tilde{\beta}/\tilde{\rho})\tilde{\gamma}_{\max}}{1+(\tilde{\beta}/\tilde{\rho})\tilde{\gamma}_k}. \tag{26}$$

As with Eq. (25) all quantities in this inequality are known as they are either assumed or may be calculated. By assuming that **B** and **R** are fixed, and considering the upper and lower bounds we gain some understanding of how the estimated eigenvalues change when the assumed background and observation error statistics are altered. From Eq. (26) we see that the lower bound increases as $\tilde{\rho}$ increases and decreases as $\tilde{\beta}$ increases. By considering the gradient we find that the upper bound is a decreasing function of $\tilde{\rho}$ and an increasing function of $\tilde{\beta}$. From this we see that the upper bound decreases as $\tilde{\rho}$ increases and increases as $\tilde{\beta}$ increases.

These theoretical results allow us to understand some behaviour of the diagnostic under the given assumptions. We now consider the performance of the diagnostic using simple examples.

5. Results in an idealised framework

5.1. The observation and background error covariance matrices

Using Eq. (6), under the assumptions used to write the diagnostics in spectral space, it is possible to test the diagnostic in an idealised framework. The assumptions restrict us to working in a 1D periodic domain. Here we choose a periodic domain of length $l = 32\pi$ and assume that we are calculating statistics for 16 equally spaced points on this domain. We note that the inverse matrices in Eq. (6) are calculated using a standard direct (Gaussian elimination) method since the matrices are small.

To define the correlated error matrices we use the second-order autoregressive function (SOAR), also known as the Balgovind (Balgovind *et al.*, 1983) correlation function, on the finite domain,

$$c(i,j) = \left(1 + \frac{\left|2a\sin\left(\frac{\theta_{i,j}}{2}\right)\right|}{L}\right)\exp\left(\frac{-\left|2a\sin\left(\frac{\theta_{i,j}}{2}\right)\right|}{L}\right), \tag{27}$$

where c is the correlation between two points i and j (Yaglom, 1986). The SOAR function is defined in terms of chordal distance on the periodic domain, a circle of radius $a = l/2\pi$, where $\theta_{i,j}$ is the central angle between points i and j. The length-scale of the correlation function is defined by L. We choose the SOAR function because at large correlation length-scales it resembles the observation error covariance structure found in Bormann *et al.* (2003). The SOAR correlation function has also been used to model the background error correlations in operational systems e.g. Ingleby (2001) and Simonin *et al.* (2014). When defining the observation and background correlation matrices we use either an identity matrix or the function in Eq. (27). In Figure 1 we plot the a row of the correlation matrix generated by the SOAR function for different values of L along with the corresponding eigenspectrum (note only half the correlation and eigenspectrum

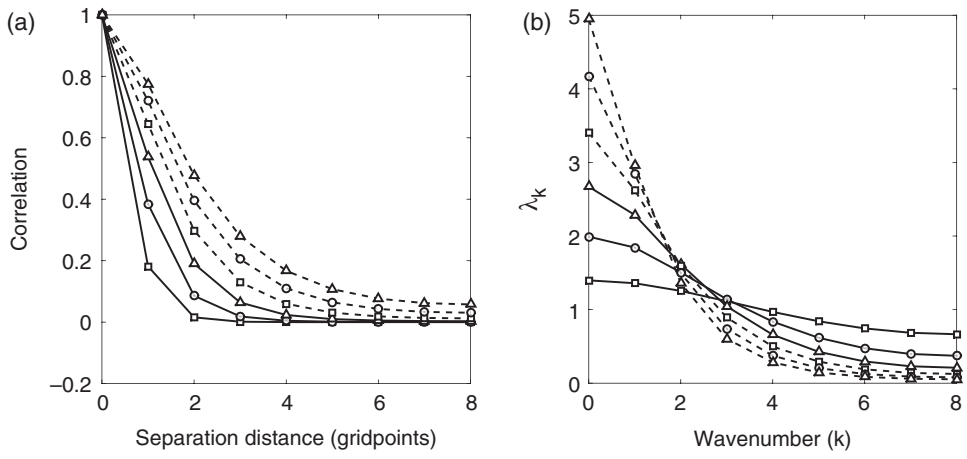

Figure 1. The SOAR function (a) and its eigenspectrum (b). $L = 2$ (solid line squares), $L = 3$ (solid line circles), $L = 4$ (solid line triangles), $L = 5$ (dashed line squares), $L = 6$ (dashed line circles), $L = 7$ (dashed line triangles).

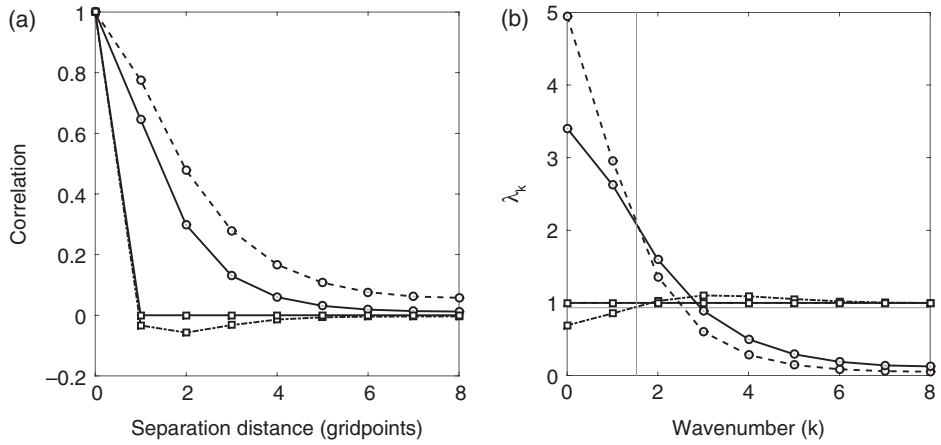

Figure 2. Correlation functions (a) and corresponding eigenvalues (b). Estimated observation error correlations (dot-dashed line squares) when $\mathbf{R} = \tilde{\mathbf{R}} = I$ (solid line squares) and the length-scale in the assumed background error covariance (black dashed line circles) is $L = 7$ compared to the actual background covariance function (solid line circles) where $L = 5$. On (b) the thin horizontal line shows the $1/\rho^e = 0.93$ value and the vertical line divides the plot into regions where $\lambda_k^e < 1/\rho^e$ (left of line) and $\lambda_k^e > 1/\rho^e$ (right of line).

are plotted due to their symmetric nature). In this and subsequent figures we show piecewise linear plots of discrete valued functions.

From Figure 1 we see that all eigenvalues are positive and the eigenvalues for a particular correlation function decrease as the wave number increases; this is a property of circulant matrices that have been constructed using positive coefficients (Gray, 2006). As we noted in section 3, the eigenvalues are ordered such that the first eigenvalue is associated with the longest length-scale. We see from the figure that as the correlation length-scale increases, the magnitude of the first eigenvalue increases. Given the result in Eq. (19) the sum of the eigenvalues must be conserved so that an increase in one, or some eigenvalues, must result in a decrease in other eigenvalues.

When considering the result of Eq. (6), there is no guarantee that the coefficients of the estimated circulant matrix will be positive and the eigenvalues may not be decreasing as a function of wave number.

5.2. *Exact uncorrelated observation errors*

We begin by considering the case where both $\mathbf{R}$ and $\tilde{\mathbf{R}}$ are diagonal. In the fortunate situation where the true variances are equal $\rho = \beta$ (note that the value of this variance may be found using Eq. (20)) and the assumed variances are equal $\tilde{\rho} = \tilde{\beta}$ Eq. (17) reduces to,

$$\rho^e \lambda_k^e = \rho \frac{1 + \gamma_k}{1 + \tilde{\gamma}_k}, \tag{28}$$

and the eigenvalues may be estimated using,

$$\lambda_k^e = \frac{\rho}{\rho^e} \left(\frac{1 + \gamma_k}{1 + \tilde{\gamma}_k} \right). \tag{29}$$

We see that,

$$\begin{aligned} &\text{if } \gamma_k < \tilde{\gamma}_k \text{ then } \lambda_k^e < \frac{\rho}{\rho^e}, \\ &\text{if } \gamma_k > \tilde{\gamma}_k \text{ then } \lambda_k^e > \frac{\rho}{\rho^e}, \\ &\text{and if } \gamma_k = \tilde{\gamma}_k \text{ then } \lambda_k^e = \frac{\rho}{\rho^e}. \end{aligned} \tag{30}$$

In a practical situation we would not know the eigenvalues corresponding to the true background error correlation matrix; however we are able to calculate the estimated observation error variance, ρ^e, and the eigenvalues, λ_k^e, of the estimated observation error correlation matrix. Therefore, if we believe we are in the situation where the observation errors are uncorrelated but we are unsure of the background error structure, then the above equations may give some insight on where the eigenvalues corresponding to the assumed background error correlation matrix are too small or too large.

We now consider a simple example to show if misspecifying a length-scale in $\tilde{\mathbf{B}}$ can introduce correlations in the estimated observation error matrix even where both $\mathbf{R}$ and $\tilde{\mathbf{R}}$ are diagonal. We set the exact background error length-scale to be $L = 5$ and the assumed background length-scale to be $L = 7$. All assumed and exact variances are set to be $\rho = \beta = \tilde{\rho} = \tilde{\beta} = 1$. We plot the exact and estimated correlation functions and corresponding eigenvalues for our SOAR example in Figure 2. We find that:

- Despite the correct matrix being used in the assimilation, the estimated observation error variance in this case is slightly larger than the actual variance ($\rho^e = 1.07$).
- The misspecified length-scale in $\tilde{\mathbf{B}}$ results in correlations in the estimated observation error covariance matrix, even though the matrix $\tilde{\mathbf{R}}$ used in the assimilation is correct.

From Eq. (30) and Figure 2 we see that the value of $1/\rho^e$ and the eigenvalues of the estimated observation error correlation matrix provide information on where the assumed background error correlation eigenvalues are larger or smaller than the true background error eigenvalues. In particular at low wave numbers, $\lambda_k^e < 1/\rho^e = 0.93$, indicating that the assumed background eigenvalues are too large. Hence in the estimated eigenvalues, we have less power in the larger scales. For the high wave numbers, $\lambda_k^e > 1/\rho^e = 0.93$, indicating that the assumed background eigenvalues are too small so in the estimated eigenvalues we expect to find increased power in the smaller scales.

5.3. *Exact observation errors correlated but assumed uncorrelated in the assimilation*

In all the following cases we set the assumed observation error covariance matrix to be diagonal, $\tilde{\mathbf{R}} = \tilde{\rho}\mathbf{I}$, with error variance $\tilde{\rho}$. Correlations for the exact observation error and exact and assumed background errors are defined using the SOAR function, Eq. (27); for the exact observation error covariance we set $L = 2$ and for the exact background error matrix we set $L = 5$. Both the error variances are chosen to be $\beta = \rho = 1$. The length-scales are chosen with the background error correlation length-scale longer than that of the observation error correlation length-scale as in an operational system it is expected that the background correlation length-scale would be larger than those of the observations. We consider how well the diagnostic estimates $\mathbf{R}^e$, both in terms of variance and length-scale. Details of the assumed background error variance, $\tilde{\beta}$, assumed observation error variance, $\tilde{\rho}$ and length-scale for $\tilde{\mathbf{B}}$ chosen for use in the assimilation for each

Table 2. Estimated observation error variances when length-scales (defined using the SOAR function in Eq. (27)) and variances in $\tilde{\mathbf{R}}$ and $\tilde{\mathbf{B}}$ used in the assimilation are incorrect. The exact observation and background error variances are set to $\rho = \beta = 1$ and length-scales to $L = 2$ and $L = 5$ respectively. The matrix $\tilde{\mathbf{R}}$ used in the assimilation is always diagonal.

Exp. label	$\tilde{\rho}$	$\tilde{\beta}$	$\tilde{\mathbf{B}}$ Length-scale (L)	ρ^c
Control	1	1	5	0.94
ρ0.5	0.5	1	5	0.68
ρ1.1	1.1	1	5	0.98
ρ2	2	1	5	1.22
ρ10	10	1	5	1.73
β0.5	1	0.5	5	1.22
β0.75	1	0.75	5	1.06
β0.99	1	0.99	5	0.94
β1.5	1	1.5	5	0.78
β2	1	2	5	0.68
L3	1	1	3	0.91
L4	1	1	4	0.92
L6	1	1	6	0.97
L7	1	1	7	1.00
$\rho 2\beta 1.5$L6	2	1.5	6	1.08
$\rho 2\beta 2$L6	2	2	6	0.97
$\rho 2\beta 1.5$L7	2	1.5	7	1.10
$\rho 2\beta 2$L7	2	2	7	1.00

experiment are detailed in Table 2 along with the estimated observation error variance.

5.3.1. *Control experiment*

We first consider the estimation of the observation error covariance matrix when the exact background error covariance matrix is used in the assimilation, $\tilde{\mathbf{B}} = \mathbf{B}$, but the observation error covariance matrix is assumed diagonal. In this case the only misspecified quantity is the length-scale in the observation error covariance matrix. This provides a reference solution that helps us understand how assuming uncorrelated observation errors affects the diagnostic. We plot a row of the exact, assumed and estimated correlation matrices and corresponding eigenvalues in Figure 3, the estimated observation error variance is given in Table 2, for the experiment labelled Control.

We see that although the correct background error is used, the assumption that the observation error covariance matrix is diagonal results in:

- An underestimated variance.
- An underestimated correlation length-scale.

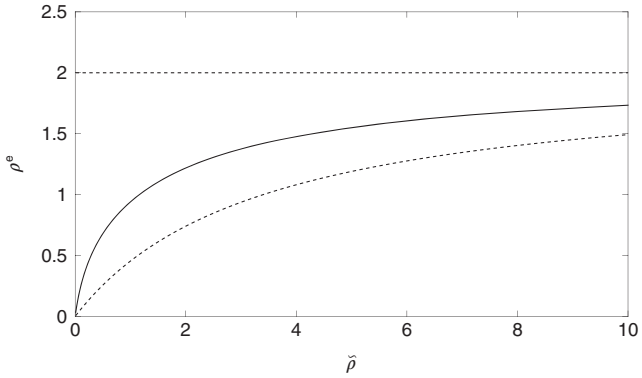

Figure 4. Estimated observation error variance and bounds when $\tilde{\mathbf{B}} = \mathbf{B}$ and $\tilde{\mathbf{R}} = \tilde{\rho}\mathbf{I}$. Change in estimated observation error variance with increasing assumed observation error variance (solid line). Upper and lower bounds given in Eq. (25) are also shown (dashed lines).

We note that the estimated observation error correlation structure shows some signature of the assumed observation error correlation structure. This result is similar to those shown in Todling (2015). Nevertheless, the estimated observation error covariance matrix is a better approximation of the exact observation error correlation compared to the diagonal assumed observation error covariance matrix. From the eigenvalues in Figure 3 it is clear that the estimated eigenvalues of the observation error correlation matrix are under (over) estimated when the assumed observation error eigenvalues are too small (large). The first eigenvalue is related to the eigenvector with the longest length-scales and hence an under (over) estimation of this eigenvalue will result in an under (over) estimation of the power in the largest scales.

We now examine in more detail the effect on the estimated observation error matrix of misrepresenting the observation and background error statistics in the assimilation.

5.3.2. *Impact of misspecifying the observation error variance*

We begin by considering how changing the variance of the assumed diagonal matrix $\tilde{\mathbf{R}}$ affects the estimates of the observation error covariance matrix. Table 2 Experiments ρ0.5, ρ2 and ρ10 show the estimated error variance when the assumed background error matrix is correct, $\tilde{\mathbf{B}} = \mathbf{B}$ and the assumed observation error covariance matrix is diagonal with an incorrect variance. We also plot in Figure 4 the change in estimated observation error variance with increasing assumed observation error variance and the values of the upper and lower bounds of the variance calculated using Eq. (25) when $\tilde{\mathbf{B}} = \mathbf{B}$ and $\tilde{\mathbf{R}} = \tilde{\rho}\mathbf{I}$.

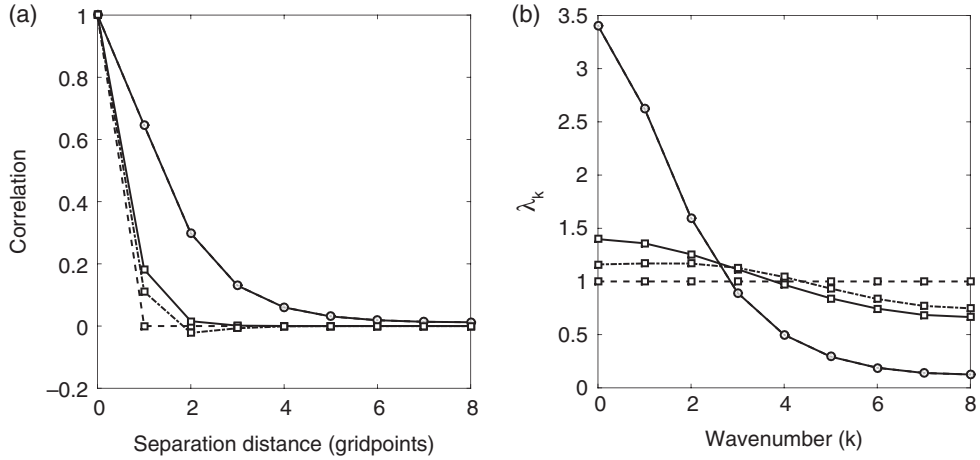

Figure 3. Correlation functions (a) and corresponding eigenvalues (b) for the Control Experiment. Estimated observation error (dot-dashed line squares), when $\tilde{\mathbf{B}} = \mathbf{B}$ (solid line circles), with $L = 2$ for $\mathbf{R}$ (solid line squares) and $\tilde{\mathbf{R}} = I$ (black dashed line squares).

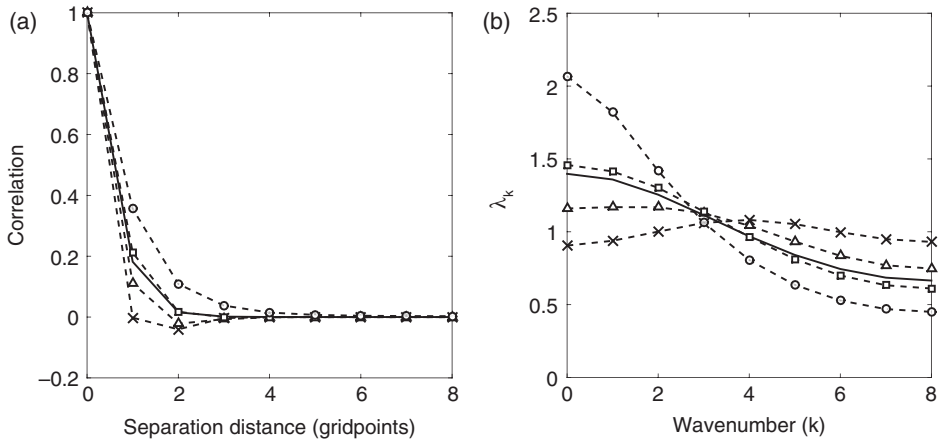

Figure 5. Estimated observation error correlations (a) and corresponding eigenvalues (b) for Experiments $\rho 0.5$ ($\tilde{\rho} = 0.5$, dashed line crosses), Control ($\tilde{\rho} = 1$, dashed line triangles), $\rho 2$ ($\tilde{\rho} = 2$, dashed line squares) and $\rho 10$ ($\tilde{\rho} = 10$, dashed line circles) when observation errors are assumed uncorrelated and the variance misspecified in the assimilation. The exact observation error correlation function (solid line) is plotted for comparison.

The results in the figure and table verify Eq. (24), that is, as the assumed observation error variance increases the estimated observation error variance increases. We also note that the theoretical bounds are respected in the experimental case.

Ménard *et al.* (2009) show for the case of scalar background and observation error variances that when the assumed observation error variance is too small (large) then the estimated observation error variance is underestimated (overestimated). Experiments $\rho 0.5$, $\rho 2$ and $\rho 10$ support this conclusion for the multi-dimensional case. However, experiment $\rho 1.1$ shows that this conclusion is not general as, for example, a choice of $\tilde{\rho} = 1.1$ results in $\rho^{\mathrm{e}} = 0.98$. The underestimation of the error variance is a result of the observation error correlations being neglected.

Next we consider the effect on the estimated length-scales. By using Eqs (23) and (24) we are able to consider the first eigenvalue,

$$\lambda_0^{\mathrm{e}} = \frac{1}{\frac{1}{n}\sum_{k=0}^{n-1} \frac{\rho\lambda_k + \beta\gamma_k}{1 + (\tilde{\beta}/\tilde{\rho})\tilde{\gamma}_k}} \left(\frac{\rho\lambda_0 + \beta\gamma_0}{1 + (\tilde{\beta}/\tilde{\rho})\tilde{\gamma}_0} \right). \tag{31}$$

This provides useful information as the first eigenvalue is related to the power in the largest scales and therefore can tell us something about the estimated correlation length-scale.

In section 4 we provided bounds for the estimated eigenvalues but not any results relating to the specific behaviour of the eigenvalues. Since here we are considering observation and background error covariance matrices that are diagonal or defined by the SOAR function, the assumed and exact correlation matrices have only non-negative coefficients. For any correlation matrix with non-negative coefficients the eigenvalues decrease as the wave number increases. Using this additional information we are able to determine what happens to the first eigenvalue. As $\tilde{\gamma}_0 \geq \tilde{\gamma}_k$, the derivative of Eq. (31) with respect to $\tilde{\rho}$ is positive and hence Eq. (31) is an increasing function with $\tilde{\rho}$. Therefore, we expect the estimated first eigenvalue, and hence the power in the lowest wave numbers, to increase as a function of $\tilde{\rho}$.

We plot the estimated correlation function and corresponding eigenvalues in Figure 5. From this we verify our theoretical result, that the power in the lowest wave numbers (largest scales) also increases as the assumed observation error variance increases. We find that error covariance length-scales are underestimated when the assumed variance is too small or correct and overestimated when the assumed variance is too large. From the eigenvalues we see that an assumed observation error variance that is too small results in the power in the large scales being underestimated and small scales overestimated.

In summary, in the case of misspecified observation error variances we are able to show that:

- As the assumed observation error variance increases the estimated observation error variance increases.
- As the assumed observation error variance increases the estimated power in the largest scales increases.
- In the case where observation errors are assumed uncorrelated, it does not hold that an assumed observation error variance that is too small (large) will result in an estimated observation error variance that is underestimated (overestimated).

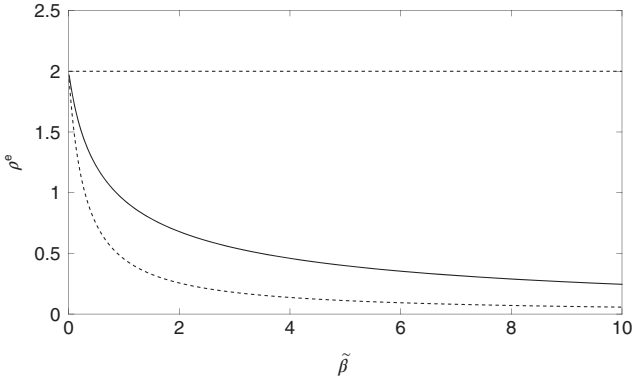

Figure 6. Change in estimated observation error variance with increasing assumed background error variance (solid line). Upper and lower bounds are also shown (dashed lines). Here $\tilde{\mathbf{R}} = \rho\mathbf{I}$ and $\tilde{\mathbf{B}} = \tilde{\beta}\mathbf{C}_b$.

5.3.3. *Impact of misspecifying the background error variance*

In Experiments $\beta 0.5$, $\beta 0.75$, $\beta 1.5$ and $\beta 2$ we consider how the estimate of $\mathbf{R}^e$ alters when the assumed background error variance $\tilde{\beta}$ is misspecified, but the length-scale in the assumed background correlation matrix is correct. We provide the assumed background variances and estimated observation error variances in Table 2. We also plot in Figure 6 the change in estimated observation error variance with increasing assumed background error variance and the values of the upper and lower bounds of the variance from Eq. (25). Again the results in the figure and table verify what is shown in Eq. (24), that as the assumed background error variance is increased the estimated observation error variance decreases. We also note that the theoretical bounds are respected in the experimental case. Experiments $\beta 0.5$, $\beta 0.75$, $\beta 1.5$ and $\beta 2$ support the conclusion from Ménard *et al.* (2009) that for the case of scalar background and observation error variances that when the assumed background error variance is too small (large) then the estimated observation error variance is overestimated (underestimated). However, the assumption of uncorrelated errors means that this conclusion is not valid in the multidimensional case: for example Experiment $\tilde{\beta} 0.99$ shows that using an assumed background error variance that is too small,

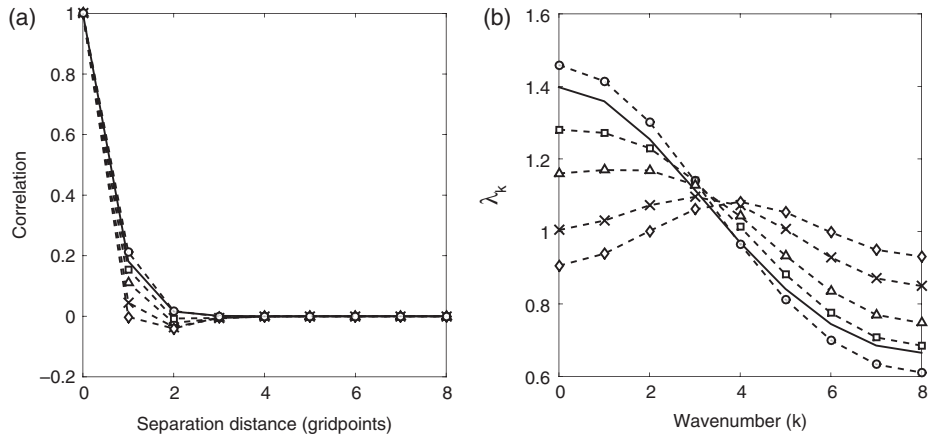

Figure 7. Estimated observation error correlations (a) and corresponding eigenvalues (b) for Experiments $\beta 0.5$ ($\tilde{\beta} = 0.5$, dashed line circles), $\beta 0.75$ ($\tilde{\beta} = 0.75$, dashed line squares), Control ($\tilde{\beta} = 1.0$, dashed line triangles), $\beta 1.5$ ($\tilde{\beta} = 1.5$, dashed line crosses) and $\beta 2$ ($\tilde{\beta} = 2.0$, dashed line diamonds) when the variance in the background error covariance matrix is misspecified in the assimilation. The exact observation correlation function (solid line) is plotted for comparison.

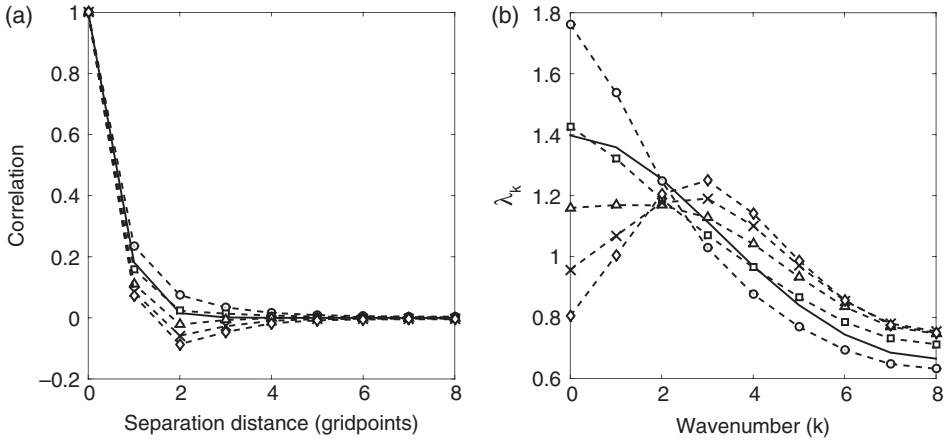

Figure 8. Estimated observation error correlations (a) and corresponding eigenvalues (b) for Experiments L3 (assumed background correlation length-scale $L = 3$, dashed line circles), L4 ($L = 4$, dashed line squares), Control ($L = 5$, dashed line triangles), L6 ($L = 6$, dashed line crosses) and L7 ($L = 7$, dashed line diamonds) when the length-scale in the background error covariance matrix is misspecified in the assimilation. The exact observation error correlation function (solid line) is plotted for comparison.

$\tilde{\beta} = 0.99$, can result in an underestimate of the observation error variance $\rho^{e} = 0.94$.

We next consider what happens to the estimated length-scale as the assumed background error variance increases. We again consider the first eigenvalue in the theoretical case. In this instance we find that, as $\tilde{\gamma}_0 \geq \tilde{\gamma}_k$, the derivative of Eq. (31) with respect to $\tilde{\beta}$ is negative. Therefore, we expect the estimated first eigenvalue and the estimated correlation length-scale to decrease as a function of $\tilde{\beta}$.

We plot the estimated correlation length-scales for Experiments $\beta 0.5$, $\beta 7.5$, $\beta 1.5$, $\beta 2$ and Control, along with the corresponding eigenvalues in Figure 7.

From the figure we see that the observation error correlation function is shortest, and most underestimated when the background error variance is largest. The observation error correlation length-scale remains underestimated as the background error variance decreases. The observation error correlation length-scale is only overestimated when the assumed background error variance is half the value of the actual background error variance or less. Considering the eigenvalues of $\mathbf{R}^{e}$ we see that unless the assumed background error variance is much smaller than the true background error variance, the power in the low wave numbers (large scales) will be underestimated and the power in the high wave numbers (small scales) will be overestimated. This is consistent with the theoretical result that the first eigenvalue decreases as $\tilde{\beta}$ increases.

In summary, in the case of misspecified background error variances we are able to show that:

- As the assumed background error variance increases the estimated observation error variance decreases.
- As the assumed background error variance increases the estimated power in the largest scales decreases.
- In the case where observation errors are assumed uncorrelated, it does not hold that an assumed observation error variance that is too small (large) will result in an estimated observation error variance that is overestimated (underestimated).

5.3.4. Impact of misspecifying the background error correlation length-scale

We now consider what happens when the background error variance is correctly specified but the correlation length-scale is misspecified. We give the assumed background correlation function length-scales and estimated observation error variances in Table 2, Experiments L3, L4, L6 and L7 and we plot the estimated observation error correlation functions and corresponding eigenvalues in Figure 8. Again we plot the result from the control experiment for comparison.

From the table and figure we see that:

- As the assumed background error length-scale increases, the estimated observation error variance increases.
- As the assumed background error length-scale increases, the estimated correlation length-scale and leading eigenvalues decrease.

However, in all but the case of the largest length-scale, the observation error variances are underestimated. We see that, when the assumed background correlation length-scale is too large, the estimated observation error length-scale is underestimated.

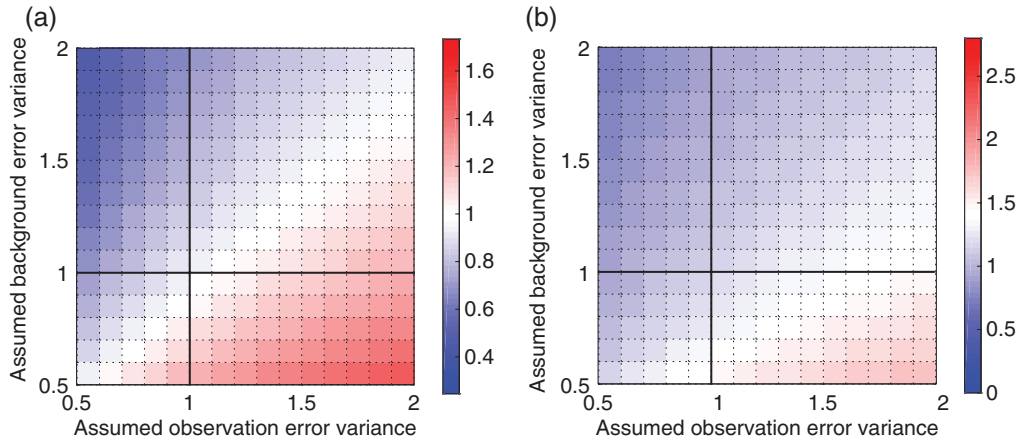

Figure 9. Change in estimated variance and first eigenvalue as a function of assumed background and observation error variance. The thick black lines show the true values of the background and observation error variance. (a) Estimated observation error variance, ρ^{e}, as a function of assumed background and observation error variance. Red shows an overestimate, white an accurate estimate and blue an underestimate. (b) First eigenvalue, γ_0^{e}, of the estimated observation error correlation as a function of assumed background and observation error variance. Red shows an overestimate, white an accurate estimate and blue an underestimate.

5.3.5. Impact of misspecifying the observation and background error variance

We now consider what happens when both the background and observation error variance are misspecified. The observation errors are still assumed uncorrelated, but the assumed background error correlation is chosen to be exact. Theoretically it is complex to prove results when multiple variables are changing. Here we consider the general trends for the estimated observation error variance and length-scales for a variety of values for $\tilde{\beta}$ and $\tilde{\rho}$.

In sections 4 and 5.3.2 we showed that as the assumed observation error variance increases, the estimated observation error variance increases. In sections 4 and 5.3.3 we showed that as the assumed background error variance increases, the estimated observation error variance decreases. So for a combination of too small (large) background variances and too large (small) observation error variances we expect an overestimated (underestimated) ρ^{e}. This is shown in Figure 9(a) where we plot the change in estimated observation error variance for different values of assumed background and observation error variance. However, it is also clear from the figure that when the effects of changes in $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{R}}$ conflict with each other, the over- or underestimation of ρ^{e} will be dominated by whichever assumed variance has the larger error. From Figure 9(b) we see that in this experimental situation the first eigenvalue is almost always underestimated and it is only in the case of low assumed background error variances or large assumed observation error variance that the leading eigenvalue is overestimated. This suggests that in most cases the length-scale of the estimated correlation matrix will be too short. This is likely to be a result of assuming uncorrelated observation errors (diagonal $\tilde{\mathbf{R}}$).

In summary when both the assumed observation and background variance are misspecified we find that:

- Knowledge of the impact of individual errors on the estimated quantities may be combined to provide information about the impact on the estimated quantities when multiple errors are present.
- A combination of too small (large) background error variances and too large (small) observation error variances will result in an overestimated (underestimated) observation error variance.
- When observation errors are assumed uncorrelated, it is likely that the diagnostic will produce an estimated observation error correlation where the length-scale is underestimated.

5.3.6. Impact of misspecifying the background error variance and correlation length-scale

We now consider what happens when both the length-scale and variance of the assumed background error matrix are misspecified. The observation errors are still assumed uncorrelated, but the assumed observation error variance is chosen to be exact, $\tilde{\rho} = \rho$. Again we consider the general trends for the estimated observation error variance and length-scales for a variety of values for $\tilde{\beta}$ and the length-scale in $\tilde{\mathbf{B}}$, before considering some specific cases.

In Figure 10(a,b) we show how the estimated variance and leading eigenvalue vary when different values are used for the assumed background error statistics.

From Eq. (24) we proved that as the assumed background error variance was increased the estimated observation error variance decreased and this can be seen by comparing any given row of Figure 10(a). However, it is more complex to say whether the variance will be over- or underestimated in any given circumstance. It is clear in this case that it is the assumed variance that has the largest impact on the estimated error variance as the horizontal gradient is much larger than the vertical gradient.

It is clear from Figure 10(b) that the estimated leading eigenvalue decreases as assumed background error variance and length-scale increases. Hence it is likely that the correlation length-scale will decrease as assumed background error variances and length-scales are increased. When estimating the leading eigenvalue it appears that the largest change is caused by the change in assumed background error length-scale rather than background error variance.

Expert opinion can often provide information on the anticipated relationship between the true and assumed statistics. In the Met Office operational convection-permitting data assimilation scheme it is suggested that both the background error variance and correlation length-scale are overestimated. Results in Waller *et al.* (2015) used the diagnostic to estimate spatial observation error statistics for Doppler radar radial winds assimilated using two different sets of background error statistics. The qualitative results agree with the findings given here, namely that when both the background error variance and correlation length-scale are overestimated the diagnostic will provide an underestimate of both the observation error variance and correlation length-scale. This suggests that although the results here are derived under a simplified framework, they still may be able to provide useful information in the operational case.

In summary when both the assumed background variance and length-scale are misspecified we find that:

- Errors in the assumed background error variance have the largest impact on the estimated observation error variance.

- Errors in the assumed background error length-scale have the largest impact on the estimated observation error eigenvalues.
- A combination of too small (large) assumed background error variances and too small (large) assumed background length-scale will result in an overestimation (underestimation) of the power in the largest scales of the estimated observation error covariance matrix.

5.3.7. *Impact of misspecifying the observation error variance and background correlation length-scale*

We next consider how the diagnostic behaves when the assumed observation error variance and assumed background error length-scales are misspecified. In Figure 11(a,b) we show how the estimated variance and leading eigenvalue vary when different values are used for the assumed background error length-scale and assumed observation error variance.

We see that the observation error variance is always underestimated when the assumed observation error variance is too small; it is overestimated when the assumed observation error variance is much too large. The first eigenvalue is over (under) estimated when both the assumed observation error variance is too large (small) and assumed background error length-scale too small (large). When both the assumed observation error variance and background error length-scale are too large (small) the eigenvalue may be over or under estimated. Again we see that the change in assumed observation error variance appears to be the dominating factor in the change in estimated observation error variance, with the variance underestimated when the assumed variance is underestimated. The dominant factor in the change of the first eigenvalue appears to be the change in length-scale of the assumed background error correlation.

In summary when the assumed observation error variance and assumed background error length-scale are misspecified we find that:

- Errors in the assumed background error variance have the largest impact on the estimated observation error variance.
- Errors in the assumed background error length-scale have the largest impact on the estimated observation error eigenvalues.
- A combination of too small (large) assumed observation error variances and too large (small) assumed background length-scale will result in an underestimation (overestimation) of the power in the largest scales of the estimated observation error covariance matrix.

5.3.8. *Impact of misspecifying all assumed error variances and length-scales*

In an operational setting it is likely that none of the assumed error statistics are exact. In the previous three sections we have seen that it may be possible to combine our knowledge of the results from sections 5.3.2 to 5.3.4 to give an understanding of what may happen to the estimated quantities when two of the assumed quantities are in error. We now consider if it is possible to make any similar conclusions when all assumed error variances and length-scales are misspecified.

Here again we restrict ourselves to consider only cases that are relevant to operational assimilation. Experiments $\rho 2\beta 1.5$L6, $\rho 2\beta 2$L6, $\rho 2\beta 1.5$L7 and $\rho 2\beta 2$L7 detail how the diagnostic performs when the observation error variance is too large and the length-scales and error variances in $\tilde{\mathbf{B}}$ are too large.

From Table 2 and Figure 12 we see that for each of the Experiments $\rho 2\beta 1.5$L6, $\rho 2\beta 2$L6, $\rho 2\beta 1.5$L7 and $\rho 2\beta 2$L7 the observation error correlation length-scale is underestimated and for experiments $\rho 2\beta 1.5$L6 and $\rho 2\beta 1.5$L7 we find that the observation error variance is overestimated.

This case, where all assumed quantities are misspecified, is a nonlinear problem, so it is not clear that combining individual results will give a clear prediction of what happens when all parameters are varied together. The exact performance of the diagnostic will always be complex to predict without detailed knowledge of the assumed and exact correlation structures. However in some circumstances, where there is some expert opinion on the assumed structures and their relation to the true statistics, it may be possible to make a valid prediction on the behaviour of the diagnostic.

6. Conclusions

To make better use of observations in data assimilation it is necessary to understand and correctly represent their associated error statistics in the assimilation method. One popular method for estimating observation error statistics, which makes use of information in the background and analysis residuals, is the method of Desroziers *et al.* (2005). Although this method has been used both in simple experiments and operational systems to provide estimates of the observation error statistics, the behaviour of the diagnostic is not well understood. In this work we have developed a theoretical understanding of the non-iterative application of this diagnostic and illustrated this with simple examples. We note that in these cases the statistical nature of the diagnostic is not considered, as the values are calculated directly and not from samples of the analysis and background residuals. When estimates are calculated in this way it is inevitable that further noise will be introduced.

To prove theoretical results relating to the diagnostic it is necessary to introduce some simplifying assumptions. We assume that the observation and background errors are homogeneous and that the observations have uniform density over a periodic domain and that the observation operator is the identity. The approach used here would also apply if we instead assume a non-identity observation operator so long as $\mathbf{HBH}^{\mathrm{T}}$ is circulant.

We begin by showing that the diagnostic can provide a satisfactory solution when the structures of the background and observation errors are similar. We highlight that the documented failure of the method in this case is a failure on the iteration of the method. Therefore if the first estimates produced from the diagnostic are close to the truth then, although no closer approximation can be obtained by iteration, the estimate may well be good enough to provide information on the error statistics. We are also able to show that results relating to under or overestimation of variance in the scalar case do not always hold in the multidimensional case.

We then restricted the theoretical work to consider only cases where the observations were assumed uncorrelated and only a single application of the diagnostic is performed. This case is of particular interest as in many operational systems the observations are assumed uncorrelated and it is not feasible to iterate the diagnostic. Under the simplifying assumptions we find that an error in just one of the assumed variances or length-scales will have an impact on both the estimated observation error variance and length-scales. In section 4 we provided bounds for the estimated observation error variance and eigenvalues of the estimated correlation matrix. From these we are able to show that the estimated observation error variance can never be larger than the sum of the true background and observation error variances. We are also able to prove that:

- The estimated observation error variance increases as assumed observation error variance increases.
- The estimated observation error variance decreases as assumed background error variance increases.

We are able to verify this through our experiments and show that the bounds on the variance are respected in the experimental cases. Under the additional assumption of the exact and assumed correlation matrices having non-negative coefficients, we are also

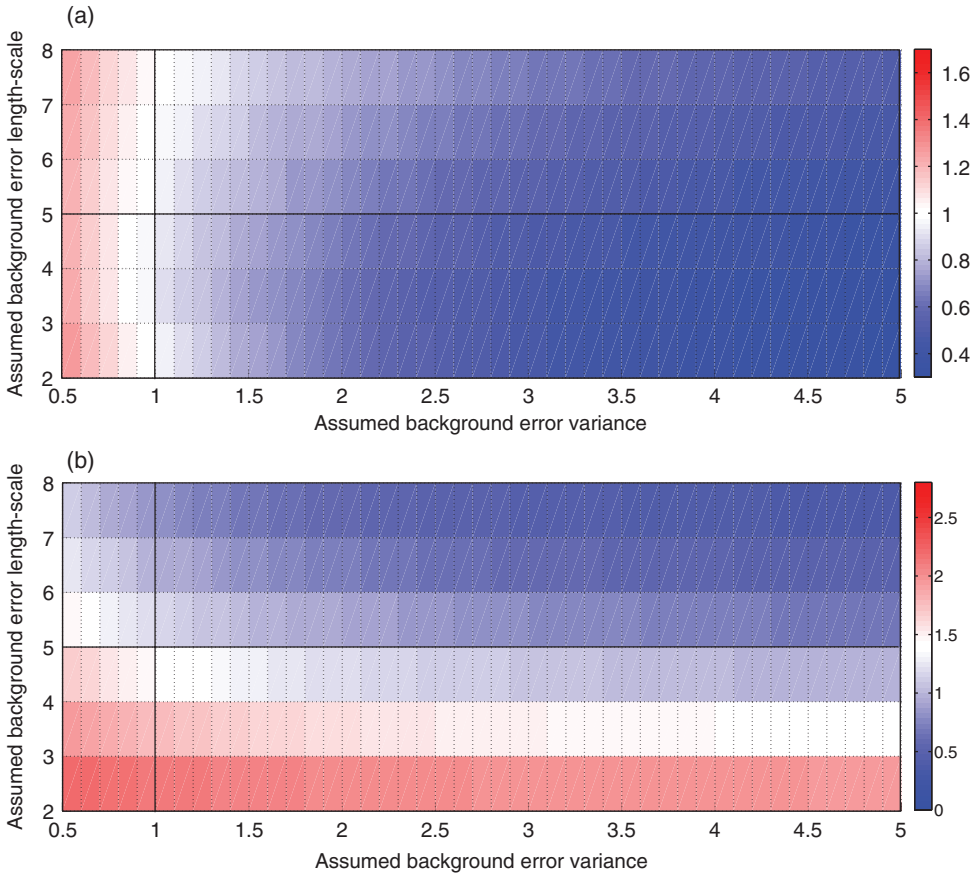

Figure 10. Change in estimated variance and first eigenvalue as a function of assumed background error variance and correlation length-scale. The thick black lines show the true values of the background error variance and length-scale. (a) Estimated observation error variance, ρ^{ϵ}, as a function of assumed background error variance and correlation length-scale. Red shows an overestimate, white an accurate estimate and blue an underestimate. (b) First eigenvalue, γ_0^{ϵ}, of the estimated observation error correlation as a function of assumed background error variance and correlation length-scale. Red shows an overestimate, white an accurate estimate and blue an underestimate.

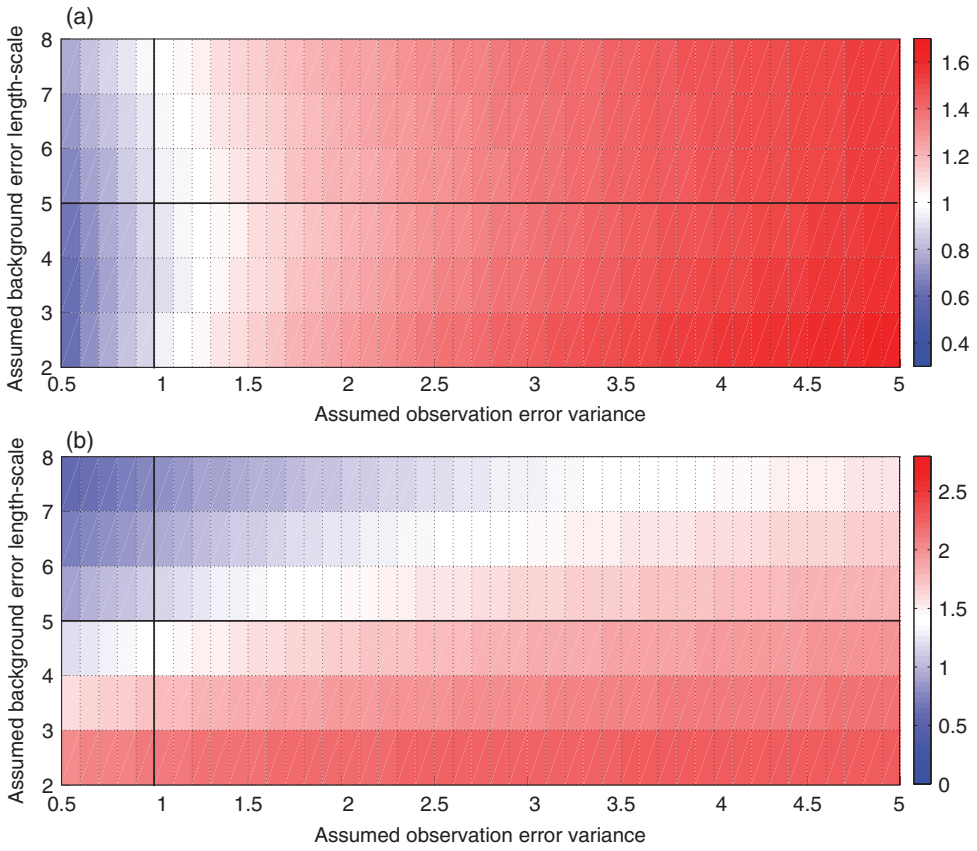

Figure 11. Change in estimated variance and first eigenvalue as a function of assumed observation error variance and background correlation length-scale. The thick black lines show the true values of the observation error variance and background correlation length-scale. (a) Estimated observation error variance, ρ^{ϵ}, as a function of assumed observation error variance and background correlation length-scale. Red shows an overestimate, white an accurate estimate and blue an underestimate. (b) First eigenvalue, γ_0^{ϵ}, of the estimated observation error correlation matrix as a function of assumed observation error variance and background correlation length-scale. Red shows an overestimate, white an accurate estimate and blue an underestimate.

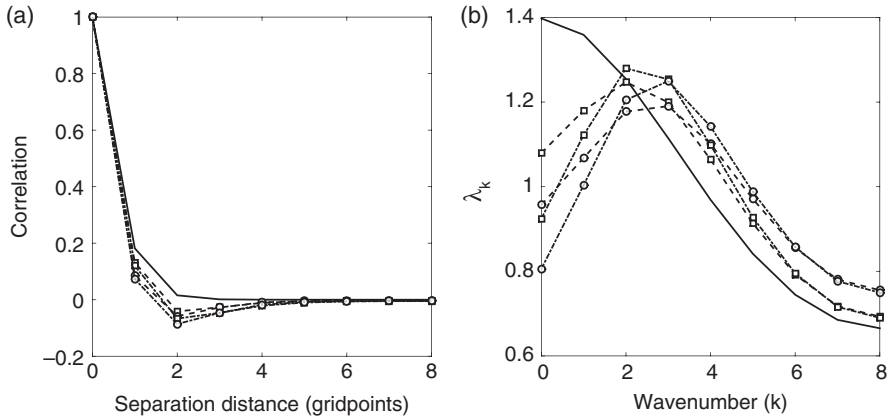

Figure 12. Rows of the exact (solid) and estimated observation error covariance matrices (a) for Experiments $\rho 2\beta 1.5$L6 (dashed black line squares), $\rho 2\beta 2$L6 (dashed black line circles), $\rho 2\beta 1.5$L7 (dash dot-line squares) and $\rho 2\beta 2$L7 (dash dot-line circles). The corresponding eigenvalues are also plotted (b).

able to prove results relating to the first eigenvalue, which provide some information about the estimated correlation length-scale. We prove that:

- The power in the large scales of the estimated observation error correlation matrix increases as assumed observation error variance increases.
- The power in the large scales of the estimated observation error correlation matrix decreases as assumed background error variance increases.

This provides some insight into the behaviour of the estimated correlation length-scale. In general we are able to show that if observation error correlations are neglected in the assumed observation error covariance matrix, then it is likely that the diagnostic will underestimate the strength of the correlations, though the result from the diagnostic will be a better estimate of **R** than one that is assumed diagonal.

The theoretical results are more complex when the background error length-scales are misspecified, so to aid our understanding we considered the results of some simple experiments. A more detailed knowledge of the exact and assumed spectra are required to predict whether the variance will increase or decrease as the assumed background error length-scales are increased. It does appear, however, in the case of the SOAR function that an increase in the assumed background error length-scales causes a reduction in the estimated observation error length-scales.

Misspecification of more than one of the background or observation error variances and length-scales is likely. Using illustrative examples we are able to show that in the case of multiple misspecification in the assimilation:

- Errors in the assumed variances will have larger impacts on the estimated observation error variances.
- Errors in the assumed length-scales will have a larger impact on the estimated observation error length-scales.

So with this knowledge and knowledge of the impact of the individual variables it is possible to hypothesize what may happen to the estimated quantities.

Another important conclusion drawn from the examples is that if the observation error covariance matrix is assumed diagonal in the assimilation, then the observation error correlation matrix calculated by the diagnostic is likely to have underestimated correlation length-scales unless the observation error variance is greatly overestimated. This is an important conclusion to bear in mind when considering operational results as in many cases the observations are assumed uncorrelated, and although the observation error variance may have been inflated it is unlikely that it is large enough for the correlations to be estimated accurately using the method of Desroziers *et al.* (2005).

The results detailed here are dependent on specific underlying assumptions, and no theoretical proof is given that these results would hold in a more general framework. However, the diagnostic has been used to estimate observation error statistics for operationally assimilated observations (Waller *et al.*, 2015). Where operational scenarios are similar to those presented here, the results are consistent. We note that understanding the estimated error statistics may not be possible if nothing is known about the relationship of the true to the assumed statistics. However, expert opinion can often provide information on the nature of this relationship. In this case, one can use the results presented here to make an informed decision on how to interpret the results obtained from the diagnostics.

Acknowledgements

This work is funded in part by the NERC Flooding from Intense Rainfall programme, and the NERC National Centre for Earth Observation and European Space Agency. We thank Daniel Hodyss of the US Navy Research Laboratory at Monterey, and Anna Shlyaeva of Environment Canada for their valuable discussions. We also thank Ricardo Todling and two anonymous reviewers whose comments were greatly appreciated.

References

Balgovind R, Dalcher A, Ghil M, Kalnay E. 1983. A stochastic-dynamic model for the spatial structure of forecast error statistics. *Mon. Weather Rev.* **111**: 701–722.

Bannister RN. 2008. A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances. *Q. J. R. Meteorol. Soc.* **134**: 1951–1970.

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, Saarinen S, Kelly G, Thepaut J. 2003. The spatial structure of observation errors in atmospheric motion vectors from geostationary satellite data. *Mon. Weather Rev.* **131**: 706–718.

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.

Dee DP, Da Silva AM. 1999. Maximum-likelihood estimation of forecast and observation error covariance parameters. Part I: Methodology. *Mon. Weather Rev.* **127**: 1822–1843.

Desroziers G, Ivanov S. 2001. Diagnosis and adaptive tuning of observation-error parameters in a variational assimilation. *Q. J. R. Meteorol. Soc.* **127**: 1433–1452, doi: 10.1002/qj.49712757417.

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, Berre L, Chapnik. 2009. 'Objective validation of data assimilation systems: Diagnosing sub-optimality'. In *Proceedings of ECMWF Workshop on Diagnostics of Data Aassimilation System Performance, 15–17 June 2009*, Reading, UK.

Golub GH, Van Loan CF. 2013. *Matrix Computations*, Chapter 7 (4th edn). The Johns Hopkins University Press: Baltimore, MD.

Gray RM. 2006. *Toeplitz and Circulant Matrices: A Review*. Foundations and Trends in Communications and Information Theory: **2**: 155–239, doi: 10.1561/0100000006.

Healy SB, White AA. 2005. Use of discrete Fourier transforms in the 1D-Var retrieval problem. *Q. J. R. Meteorol. Soc.* **131**: 63–72.

Hollingsworth A, Lönnberg P. 1986. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field *Tellus* **38A**: 111–136.

Ingleby NB. 2001. The statistical structure of forecast errors and its representation in the Met. Office global 3-D variational data assimilation scheme. *Q. J. R. Meteorol. Soc.* **127**: 209, doi: 10.1002/qj.49712757112.

Janjic T, Cohn SE. 2006. Treatment of observation error due to unresolved scales in atmospheric data assimilation. *Mon. Weather Rev.* **134**: 2900–2915.

Li H, Kalnay E, Miyoshi T. 2009. Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter. *Q. J. R. Meteorol. Soc.* **128**: 1367–1386.

Lorenc AC. 1981. A global three-dimensional multivariate statistical interpolation scheme. *Mon. Weather Rev.* **109**: 701–721.

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

Ménard R, Yang Y, Rochon Y. 2009. 'Convergence and stability of estimated error variances derived from assimilation residuals in observation space'. In Proceedings of ECMWF Workshop on Diagnostics of Data Assimilation System Performance, 15-17 June 2009, Reading, UK.

Miyoshi T, Kalnay E, Li H. 2013. Estimating and including observation-error correlations in data assimilation. *Inverse Prob. Sci. Eng.* **21**: 387–398.

Simonin D, Ballard SP, Li Z. 2014. Doppler radar radial wind assimilation using an hourly cycling 3D-Var with a 1.5 km resolution version of the Met Office Unified Model for nowcastings. *Q. J. R. Meteorol. Soc.* **140**: 2298–2314, doi: 10.1002/qj.2298.

Stewart LM. 2010. 'Correlated observation errors in data assimilation', PhD thesis. University of Reading: http://www.reading.ac.uk/maths-and-stats/research/theses/maths-phdtheses.aspx (accessed 1 October 2015).

Stewart LM, Dance SL, Nichols NK. 2008. Correlated observation errors in data assimilation. *Int. J. Numer. Methods Fluids* **56**: 1521–1527.

Stewart LM, Dance SL, Nichols NK. 2013. Data assimilation with correlated observation errors: experiments with a 1-D shallow water model. *Tellus A* **65**, doi: 10.3402/tellusa.v65i0.19546.

Stewart LM, Dance SL, Nichols NK, Eyre JR, Cameron J. 2014. Estimating interchannel observation-error correlations for IASI radiance data in the Met Office system. *Q. J. R. Meteorol. Soc.* **140**: 1236–1244, doi: 10.1002/qj.2211.

Todling R. 2015. A complementary note to 'A lag-1 smoother approach to system-error estimation': The intrinsic limitations of residual diagnostics. *Q. J. R. Meteorol. Soc.*, doi: 10.1002/qj.2546.

Waller JA. 2013. 'Using observations at different spatial scales in data assimilation for environmental prediction', PhD thesis. University of Reading, Department of Mathematics and Statistics, http://www.reading.ac.uk/maths-and-stats/research/theses/maths-phdtheses.aspx (accessed 1 October 2015).

Waller JA, Dance SL, Lawless AS, Nichols NK. 2014a. Estimating correlated observation error statistics using an ensemble transform Kalman filter. *Tellus A* **66**: doi: 10.3402/tellusa.v66.23294.

Waller JA, Dance SL, Lawless AS, Nichols NK, Eyre JR. 2014b. Representativity error for temperature and humidity using the Met Office high-resolution model. *Q. J. R. Meteorol. Soc.* **140**: 1189–1197, doi: 10.1002/qj.2207.

Waller JA, Simonin D, Dance SL, Nichols NK, Ballard SP. 2015. Diagnosing observation error correlations for Doppler radar radial winds in the Met Office UKV model using observation-minus-background and observation-minus-analysis statistics. University of Reading: http://www.reading.ac.uk/web/FILES/maths/Preprint_MPS_15-18_Waller.pdf (accessed 1 October 2015).

Weston PP, Bell W, Eyre JR. 2014. Accounting for correlated error in the assimilation of high-resolution sounder data. *Q. J. R. Meteorol. Soc.* **140**: 2420–2429, doi: 10.1002/qj.2306.

Yaglom A. 1986. *Correlation Theory of Stationary and Related Random Functions I. Basic Results*. Springer-Verlag: New York, NY.