Atmospheric
Chemistry
and Physics

Technical Note: Spectral representation of spatial correlations in variational assimilation with grid point models and application to the Belgian Assimilation System for Chemical Observations (BASCOE)

Q. Errera[1] and R. Ménard[2]

[1]Insitut d'Aéronomie Spatiale de Belgique, BIRA-IASB, Belgium
[2]Air Quality Research Division, Environment Canada, Dorval (Qc), Canada

Correspondence to: Q. Errera (quentin@oma.be)

Abstract. The formulation of the background error covariances represented in the spectral space is discussed in the context of univariate assimilation relying on a grid point model, leaving out all the aspects of balances between the different control variables needed in meteorological assimilation. The spectral transform operations are discussed in the case of a spherical harmonics basis and we stress that there is no need for an inverse spectral transform and of a Gaussian grid. The analysis increments are thus produced directly on the model grid. The practice of producing analysis increments on a horizontal Gaussian grid and then interpolating to an equally spaced grid is also shown to produce a degradation of the analysis. The method discussed in this paper allows the implementation of separable and non-separable spatial correlations. The separable formulation has been implemented in the Belgian Assimilation System for Chemical ObsErvations (BASCOE) and its impact on the assimilation of O_3 observed by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) is shown. To promote the use of this method by other non-meteorological variational systems and in particular chemistry, the Fortran code developed is made available to the community.

1 Introduction

One of the critical aspects of any assimilation system is the formulation of the background spatial error covariances matrix, denoted the **B** matrix. It needs to be sufficiently com-

pact to be numerically implemented and sufficiently complex to represent correctly the real error covariances of the first guess field. There are several approaches to achieve this goal. For variational systems, which are the systems we will focus on, three approaches have been developed. The first one historically (Parrish and Derber, 1992), which we shall discuss at length in this paper, uses a spectral representation of the correlation matrix. It is based on the fact that a homogeneous and isotropic horizontal correlation matrix can be represented by a diagonal matrix in the spectral space. This property was discovered in the late sixties by Russian scientists working in statistical fluid mechanics and turbulence theory (see the works of Monin and Yaglom, 1971, 1975; Panchev, 1971). These concepts were extended to the sphere and applied successfully to large-scale atmospheric dynamics and analyses by Boer (1983) and Boer and Shepherd (1983) using spherical harmonics as the orthogonal basis. It is only in the late nineties that this formulation became the cornerstone of 3-dimensional variational meteorological assimilation systems (Courtier et al., 1998), and was applied later to chemical data assimilation (Dethof and Hólm, 2004). More recently, efforts have been made to develop spectral representation of inhomogeneous and flow-dependent correlations (Fisher, 2003).

The second approach for variational systems is based on a diffusion operator which uses the model's diffusion to generate the effect of a Gaussian correlation function. This method was introduced by Derber and Rosati (1989) and further developed by Weaver and Courtier (2001). This method

is particularly useful when the domain has complex boundaries (e.g. ocean data assimilation) where it is difficult to define positive definite correlations. This method is also well suited for models that have non uniform grid on a sphere (e.g. icosahedral grids) because it avoids the interpolation of analysis increments onto that grid (Elbern et al., 2010; Schwinger and Elbern, 2010). The diffusion approach primarily gives rise to a Gaussian covariance model but can be generalized to more complex (i.e. inhomogeneous and anisotropic) correlations. However, the diffusion approach is highly demanding in computer resources (both CPU and memory), especially if one needs to simulate non-Gaussian correlations.

A third approach in variational data assimilation which does not compute explicitly the error covariance is the recursive filter (Purser et al., 2003). Like the diffusion operator approach, the recursive filter approach attemps to compute Gaussian correlations. The method consists in evaluating the effect of a Gaussian correlation model on a state vector, by applying a sequence of 1-D finite difference operators in different directions on the state vector. Repeated applications of these finite difference operators in carefully chosen directions can lead to approximate the smoothing effect of Gaussian homogeneous and isotropic correlations. Although positive-definitness can be obtained, the scheme is approximate and computationally complex.

This paper discusses the spectral representation of background error correlations onto a spherical harmonic basis. Most of the mathematical background presented in this paper has already been published (see the references cited in Sects. 2–4). However, none of those publications provide a complete picture of the problem. For instance, these publications rarely give an introduction on spectral transformations. As numerical weather prediction (NWP) systems usually rely on spectral models, this aspect is supposed to be known. Moreover, those publications describe the method in the meteorological context, that is the use of balance operators between meteorological variables. This makes the background theory of this method difficult to be acquired for people not involved in meteorology. As a consequence, one of the aims of this paper is to describe the spectral representation of the background error correlations for univariate assimilation with systems relying on a grid point model.

A code for this purpose has been developed and implemented in the Belgian Assimilation System for Chemical ObsErvation (BASCOE). In doing so we have identified a rather important property that seemed to have been overlooked in meteorological data assimilation: in variational assimilation only the adjoint spectral transform is necessary, not the inverse transform. As a result the analysis spectral increments are produced directly on the model grid. In other words, there is no need to extract the increments on the (non-equally spaced) Gaussian grid and then to perform a transformation from that grid to the model grid – a transformation which necessarily degrades the analysis. While this method is limited to homogeneous and isotropic correlations, it has

the advantage of being very low demanding in computer resources. Moreover, while not studied in this paper, non-separable three-dimensional error correlations could be implemented easily, meaning that vertical correlations can be different for large and small horizontal spectral features.

This paper is organized as follows. Section 2 and Appendix A introduce representations in spherical harmonics and the spectral representation of homogeneous and isotropic correlations. Section 3 discusses the variational formulation and the control variable change from the model basis to the spherical harmonic basis. The implementation of the spectral transform operators and of the spatial correlations is described in Sect. 4. Section 5 discusses the numerical implementation of the method and illustrates it by the results of an assimilation of a single pseudo-observation. In Sect. 6, the method is applied to real ozone observations using the BASCOE system. The results of this paper are summarized in the conclusions. A Fortran code of this method is provided in the Supplement of this paper. This code is introduced in the Appendix B.

2 Introduction on spectral transform on the sphere

This section provides some useful information about two-dimensional spectral representation on the sphere. For an exhaustive introduction, we refer to the books of Satoh (2004) and Krishnamurti et al. (2006), and the unpublished lectures of P. Swarztrauber*.

In this section, we denote a 2-D field on the sphere by $\Psi(\lambda, \mu)$ where λ is the longitude, $\mu = \sin\phi$ and ϕ is the latitude. Alternatively, the notation Ω will be used in place of (λ, μ). The field $\Psi(\lambda, \mu)$ can be represented by a series of spherical harmonics where the coefficients are noted Ψ_n^m. The values of n and m are specified below. In this section, the term *grid* will be associated with the *horizontal grid* and not the usual three dimensional grid of the model.

2.1 Spherical harmonics

Spherical harmonics, denoted by Y_n^m, are defined as solutions of the Laplace equation. They take the following form:

$$Y_n^m(\lambda, \mu) = N_n^m P_n^m(\mu) e^{im\lambda} \tag{1}$$

where N_n^m are the normalization constants, $P_n^m(\mu)$ are the *associated Legendre functions* and $e^{im\lambda}$ are the Fourier exponentials where $i^2 = -1$. Spherical harmonics form an orthogonal basis and their normalization is defined by the normalization constants. This is discussed in detail in Appendix A1. Here the geodesy 2π normalization is adopted

---

*available at:
www.cisl.ucar.edu/css/staff/pauls/tutorials/index.html

(while Courtier et al., 1998, uses the geodesy 4π) such that:

$$\int_0^{2\pi}\int_{-1}^{+1} Y_n^m(\lambda,\mu)[Y_{n'}^{m'}(\lambda,\mu)]^*d\mu d\lambda = 2\pi\delta_n^{n'}\delta_m^{m'} \tag{2}$$

where [.]* denotes the complex conjugate and $\delta_k^{k'}$ is the Kronecker symbol ($\delta_k^{k'}=1$ if $k=k'$ and 0 otherwise). When $m=0$, P_n^0 are called the *Legendre polynomials*. In such a case the value of m is omitted in the notation ($P_n^0 \equiv P_n$). When spherical harmonics are used in the context of spectral transforms, as in this study, it is appropriate to include the normalization constant in the associated Legendre functions. In this case, we define the *normalized associated Legendre functions* as $\overline{P}_n^m(\mu) \equiv N_n^m P_n^m(\mu)$.

2.2 Direct spectral transform

In general, a smooth function on a sphere $\Psi(\lambda,\mu)$ can be represented by a series of spherical harmonics with coefficients Ψ_n^m:

$$\Psi(\lambda,\mu) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n}\Psi_n^m Y_n^m(\lambda,\mu) . \tag{3}$$

In this transformation, and recalling the definition of the spherical harmonics (Eq. 1), the dependence along the longitude is taken by the Fourier exponentials. For this reason, m is called the *zonal wavenumber*. In addition, n is called the *total wavenumber*.

Ψ_n^m are complex coefficients. If $\Psi(\lambda,\mu)$ is real, which is the case in atmospheric sciences, then the following property holds:

$$\Psi_n^m = (-1)^m[\Psi_n^{-m}]^* . \tag{4}$$

This suggests that it might be possible to define real coefficients using Fourier sine and cosine instead of the complex coefficients using Fourier exponentials. In this study, the complex representation is kept.

The transform Eq. (3) cannot be evaluated exactly and requires to be truncated at some reasonable number of terms. Several truncation methods exist and here the triangular truncation is adopted. In this case, the relationship between the spectral coefficients and the physical field is given by

$$\Psi(\lambda_j,\mu_k) = \sum_{n=0}^{N}\sum_{m=-n}^{n}\Psi_n^m \overline{P}_n^m(\mu_k)e^{im\lambda_j} \tag{5}$$

$$= \sum_{m=-N}^{N}\sum_{n=|m|}^{N}\Psi_n^m \overline{P}_n^m(\mu_k)e^{im\lambda_j} \tag{6}$$

where N is the degree of truncation (and should not be confused with the normalization factor N_n^m) and where j and k are respectively the longitude and latitude indices of the

target grid of the spectral transform. For practical purposes, Eq. (6) is split in a Fourier transform followed by a Legendre transform:

$$\Psi^m(\mu_k) = \sum_{n=|m|}^{N}\Psi_n^m \overline{P}_n^m(\mu_k) \tag{7}$$

$$\Psi(\lambda_j,\mu_k) = \sum_{m=-N}^{N}\Psi^m(\mu_k)e^{im\lambda_j} . \tag{8}$$

Note that usually, operational centers use fast Fourier transforms instead of Eq. (8). This is not the case in the code provided in the Supplement as the CPU time necessary to operate the Fourier transform is found sufficiently fast.

2.3 Inverse spectral transform

Given a 2-D field $\Psi(\lambda,\mu)$ on the sphere, the question is now how we can calculate its spectral coefficients Ψ_n^m. Using the orthogonality of the spherical harmonics, the inverse of Eq. (3) is given by

$$\Psi_n^m = \frac{1}{2\pi}\int_0^{2\pi}\int_{-1}^{1}\Psi(\lambda,\mu)[Y_n^m(\lambda,\mu)]^*d\mu d\lambda . \tag{9}$$

In order to solve this equation, one usually starts with a discrete 2-D field given on a grid $\Psi(\lambda_j,\mu_k)$ and looks to invert the sequence (Eqs. 7, 8). The inverse of the Fourier transform (Eq. 8) is given by:

$$\Psi^m(\mu_k) = \frac{1}{M}\sum_{j=0}^{M-1}\Psi(\lambda_j,\mu_k)e^{-im\lambda_j} \tag{10}$$

where M is the number of longitudes.

The inverse of the Legendre transform (Eq. 7) is much more difficult to implement. One aspect that has not been addressed yet is the specification of the target grid of the spectral transform. The direct transform (Eqs. 7, 8) and the inverse of the Fourier transform (Eq. 10) are applicable to any type of grid, in particular to the equally spaced (in degree) model grid. However, this is not true for the inverse of the Legendre transform. If one needs to calculate the exact inverse of Eq. (7) without loss of information, the choice of spectral grid is limited to the *Gaussian grid*. On this grid, the latitudes correspond to the roots of the Legendre polynomials $P_N(\mu_k)=0$, which are not equally spaced. The longitudes, on the other hand, are equally spaced. On the Gaussian grid, the inverse spectral transform is defined by the *Gaussian quadrature* or *Gauss-Legendre quadrature*.

Let $f(\mu)$ be a polynomial of degree $(2L-1)$. The Gaussian quadrature specifies that the integral of $f(\mu)$ between -1 and 1 can be solved exactly by the following quadrature (see the proof in Krishnamurti et al., 2006, Sect. 6.6; or Satoh, 2004,

Sect. 21.3):

$$\int_{-1}^{+1} f(\mu)d\mu = \sum_{l=1}^{L} w(\mu_l) f(\mu_l) \tag{11}$$

where $w(\mu_l)$ are the Gaussian weights which can be calculated analytically.

Using (11), the inverse of the Legendre transform (7) is given by:

$$\Psi_n^m = \sum_{k=1}^{K} w(\mu_k) \Psi^m(\mu_k) \overline{P}_n^m(\mu_k) \tag{12}$$

where K is the number of latitudes.

An equally spaced grid can also be a target grid for the spectral transform. While the Gaussian grid allows one to define an exact inverse transform, an accurate (but not exact) inverse transform can also be defined from the equally spaced grid. This is reported in the Appendix A2.

In the following, the target grid of the spectral transform will be simply denoted by the target.

2.4 Adjoint spectral transform

In variational assimilation where the **B** matrix is defined in the spectral space, the adjoint of the spectral transform (Eqs. 7, 8) is needed. Let us denote the spectral transform operator by **S**. To obtain the gradient of the cost function (see Sect. 3), one needs to define the adjoint of **S**, denoted by **S*** (that must not be confused with the conjugate operator $[.]^*$). One way to implement the adjoint is to (re)define the spectral transform as unitary operations such that $\mathbf{S}^{-1} = \mathbf{S}^*$. This is possible by re-writing the direct spectral transform (Eqs. 7, 8) as

$$\Psi^m(\mu_k) = \sqrt{\frac{1}{M}} \sum_{n=|m|}^{N} \Psi_n^m \overline{P}_n^m(\mu_k) \tag{13}$$

$$\Psi(\lambda_j, \mu_k) = \sum_{m=-N}^{N} \sqrt{w(\mu_k)} \Psi^m(\mu_k) e^{im\lambda_j} , \tag{14}$$

and the inverse spectral transform (Eqs. 10, 12) as

$$\Psi^m(\mu_k) = \sqrt{\frac{1}{M}} \sum_{j=0}^{M-1} \Psi(\lambda_j, \mu_k) e^{-im\lambda_j} \tag{15}$$

$$\Psi_n^m = \sum_{k=1}^{K} \sqrt{w(\mu_k)} \Psi^m(\mu_k) \overline{P}_n^m(\mu_k) . \tag{16}$$

However, this formulation requires the use of the Gaussian grid as the target grid of the spectral transform. If one uses a model defined on an equally spaced grid, as usually done in atmospheric modeling, a mapping operation will be necessary to map the analysis increments from the target grid to the model grid. As mapping operations introduce a degradation of the increments, it is better to find an alternative to the unitary definition of the spectral transforms. For this reason, the adjoint of the code of the direct spectral transform sequence (Eqs. 7, 8) has been written by hand in BASCOE. In terms of equation, the operator **S*** takes the form:

$$\hat{\Psi}^m(\mu_k) = \sum_{j=0}^{M-1} \hat{\Psi}(\lambda_j, \mu_k) e^{-im\lambda_j} \tag{17}$$

$$\hat{\Psi}_n^m = \sum_{k=1}^{K} \hat{\Psi}^m(\mu_k) \overline{P}_n^m(\mu_k) \tag{18}$$

where $\hat{\Psi}$ denote the adjoint variable of Ψ, i.e. $\hat{\Psi} = \partial J / \partial \Psi$.

2.5 Degree of truncation, target grid and aliasing

To complete the definition of the spectral operators, we have to define a degree of truncation N and the corresponding dimensions of the horizontal grid M and K. First, let us consider the unitary definition of the spectral transforms. In order to keep the information content equal in a sequence of direct and inverse Fourier transforms, one needs to impose that $M = 2(N + 1)$ (Krishnamurti et al., 2006, Chap. 7; Swarztrauber, Lecture 1). If one tries $M > 2(N + 1)$, a false representation, or *aliasing* is introduced. On the other hand, if $M < 2(N + 1)$, there is a loss of information during the spectral transform (i.e. the spectral representation has a higher resolution than the horizontal grid). Aliasing in the associated Legendre functions also exists but is much more difficult to derive. Associated Legendre functions can be obtained by a sum of Fourier coefficients in the latitude direction (Swarztrauber, Lecture 2). In this case, the aliasing is prevented if $K \leq N + 1$ and the equality is required for no loss of information in a sequence of direct and inverse Legendre transform.

If one bases their code on the non-unitary spectral transforms, the condition is less strict. It can be verified with the code provided in the Supplement of this paper that the following conditions must hold: $N \geq \max(K, M/2) - 1$. This allows one to implement longitutes and latitudes that have not necessarilly the same resolution, which is usually the case in atmospheric modeling.

In the case of the BASCOE horizontal grid of $2° \times 2°$ that includes the poles (as done in Sect. 6), the model resolution is 91 latitudes and 180 longitudes. If the target grid of the spectral transform is the BASCOE model grid, $K = 91$, $M = 180$ and the degree of truncation is fixed to $N = 90$. If the target grid is the Gaussian grid, a mapping operator **G** must be introduced to account for the transformation from the Gaussian grid to the BASCOE grid. In this case, the size of the Gaussian grid is $K = 90$ and $M = 180$, the degree of truncation is $N = 89$ and the dimensions of **G** are 91×90.

2.6 Spectral representation of spatial correlations

Let $\epsilon(\Omega)$ be an error field on the surface sphere where $\Omega = (\lambda, \mu)$. Let us assume that the error field is unbiased and normalized, i.e. $\langle \epsilon(\Omega)^2 \rangle = 1$ where $\langle \, \rangle$ denotes the mathematical expectation. Let us assume that the correlations are homogeneous and isotropic. This means that the correlations between the points Ω and Ω' depends only on the distance d between the two points. Gaspari and Cohn (1999) have shown that, on a sphere of radius A, d corresponds to the *chordal* or *geodetic* distance between the two points

$$d \equiv d(\theta) = A\sqrt{2(1-\cos\theta)} \,.^{\dagger} \tag{19}$$

where θ is their angular separation as defined in Eq. (A14). It follows that the correlations between the two points might be expressed as a function of θ:

$$\langle \epsilon(\Omega)[\epsilon(\Omega')]^* \rangle = f(\theta) \,. \tag{20}$$

The concept of homogeneous and isotropic correlations on the surface of a sphere is however not independent as noted by Gaspari and Cohn (1999), and needs to be somewhat revised. Isotropy can be obtained by translation invariance back to the same point along great circle with translations in any directions – so basically the correlation function only needs to be homogeneous. Also as discussed in the same paper, any homogeneous correlation function defined in $\mathbb{R}^3$ is also homogeneous on any continuous manifold embedded in $\mathbb{R}^3$, such as the surface of a sphere. Thus an homogeneous (translation invariant) correlation function on the surface of a sphere can be defined from a correlation using as the chordal distance d defined in $\mathbb{R}^3$ (and not the great circle distance), and which has the consequence to be implicitly periodic from all directions.

Let us come back to Eq. (20). As homogeneous (and isotropic) correlations over the sphere are invariant with rotation, let us suppose that one of the two points is at the north pole. Then θ is the co-latitude angle, i.e. $\theta \equiv \frac{\pi}{2} - \phi$. Consequently, we have $\cos\theta = \mu$. In this configuration, the correlations between the two points are independent of the longitude. If f_n^m represents the spectral coefficients of $f(\theta)$, we have $f_n^m = 0$ for $m \neq 0$. It follows that

$$f(\theta) = \sum_{n=0}^{N} \sum_{m=-n}^{n} f_n^m Y_n^m(\lambda, \cos\theta) \tag{21}$$

$$= \sum_{n=0}^{N} f_n \overline{P}_n^0(\cos\theta) \tag{22}$$

where $f_n \equiv f_n^0$. Boer (1983) and Gauthier et al. (1993) have shown that the correlations are represented by a diagonal matrix on a spherical harmonics basis, i.e.

$$\langle \epsilon_n^m [\epsilon_{n'}^{m'}]^* \rangle = b_n \delta_n^{n'} \delta_m^{m'} \tag{23}$$

---

† And not $2\sqrt{(1-\cos\theta)}$ as reported in Gaspari and Cohn (1999, Eq. 2.33).

where $b_n \equiv f_n/N_n^0$ (see also the proof in Appendix A4).

3 Variational assimilation and control variable transform

Variational methods (three-dimensional – 3D-Var – and four-dimensional – 4D-Var) aim at calculating the model state that minimizes the objective function $J(\mathbf{x})$ (Talagrand, 1997):

$$J(\mathbf{x}) = J^{\mathrm{b}}(\mathbf{x}) + J^{\mathrm{o}}(\mathbf{x}) \tag{24}$$

where

$$J^{\mathrm{b}}(\mathbf{x}) = \frac{1}{2}[\mathbf{x} - \mathbf{x}^{\mathrm{b}}]^T \mathbf{B}^{-1}[\mathbf{x} - \mathbf{x}^{\mathrm{b}}] \tag{25}$$

$$J^{\mathrm{o}}(\mathbf{x}) = \frac{1}{2}[\mathbf{y} - H(\mathbf{x})]^T \mathbf{R}^{-1}[\mathbf{y} - H(\mathbf{x})] \,, \tag{26}$$

where $\mathbf{x}^{\mathrm{b}}$ and $\mathbf{B}$ are respectively the background model state (i.e. the initial guess) and its associated error covariance matrix; $\mathbf{y}$ and $\mathbf{R}$ are respectively the observational vector and its associated error covariance matrix; $\mathbf{x}$ is the model state vector and H is the non-linear observation operator that projects the model state in the observation space. In the 3D-Var case the observations and the model state correspond to the same time. In the 4D-Var case, observations span over an assimilation time window, the model state $\mathbf{x}$ is defined at the beginning of that window (or initial time) and the observation operator H includes an evolution model operator that projects the model initial state $\mathbf{x}$ at the observation time.

The minimization of Eq. (24) is usually done with a quasi-Newton minimizer that requires the knowledge of the gradient of the objective function:

$$\nabla_{\mathbf{x}} J = \frac{\partial J}{\partial \mathbf{x}} = \nabla_{\mathbf{x}} J^{\mathrm{b}} + \nabla_{\mathbf{x}} J^{\mathrm{o}} \tag{27}$$

$$\nabla_{\mathbf{x}} J^{\mathrm{b}} = \mathbf{B}^{-1}[\mathbf{x} - \mathbf{x}^{\mathrm{b}}] \tag{28}$$

$$\nabla_{\mathbf{x}} J^{\mathrm{o}} = \left(\frac{\partial H(\mathbf{x})}{\partial \mathbf{x}}\right)^T \mathbf{R}^{-1}[\mathbf{y} - H(\mathbf{x})] \,. \tag{29}$$

As the typical dimension of $\mathbf{x}$ is around 10^6, the matrix $\mathbf{B}$ is of size 10^{12}. This is far too large to be stored and inverted by modern computers. Moreover, the specification of the elements of such a matrix requires a huge amount of a priori information, more than available (Dee, 1995). For those reasons, it is necessary to reduce the problem. Here, we follow the same strategy as developed at ECMWF in late nineties (Courtier et al., 1998). A review on the formulation of background error covariance matrix in meteorology might be found in Bannister (2008a,b).

In order to avoid the problem of inverting $\mathbf{B}$, a control variable transform is introduced:

$$\mathbf{L}\chi = \underbrace{\mathbf{x} - \mathbf{x}^{\mathrm{b}}}_{\delta\mathbf{x}} \tag{30}$$

where χ is a new control variable, $\delta \mathbf{x}$ is the analysis increment and $\mathbf{L}$ is the square root of $\mathbf{B}$:

$$\mathbf{B} = \mathbf{L}\mathbf{L}^T . \tag{31}$$

The objective function thus becomes:

$$J^b(\chi) = \frac{1}{2}\chi^T\chi \tag{32}$$

$$J^o(\chi) = \frac{1}{2}[\mathbf{d} - H(\mathbf{L}\chi)]^T\mathbf{R}^{-1}[\mathbf{d} - H(\mathbf{L}\chi)] \tag{33}$$

where

$$\mathbf{d} = \mathbf{y} - H(\mathbf{x}^b) \tag{34}$$

is the innovation vector. The background term (Eq. 32) can be easily calculated. If the observation operator H is linear, the term $\mathbf{d} - H(\mathbf{L}\chi)$ in Eq. (33) is equal to the term $\mathbf{y} - H(\mathbf{x})$ in Eq. (26). As a result, $J^o(\mathbf{x}) = J^o(\chi)$. If H is non-linear but the difference $\mathbf{x} - \mathbf{x}^b$ is small, then H can be linearized. It follows that the calculation of J^o is practically obtained by the resolution of Eq. (26), which is much easier to implement numerically than Eq. (33).

The gradient of the objective function is now given by:

$$\nabla_\chi J^b = \chi \tag{35}$$

$$\nabla_\chi J^o = \mathbf{L}^* \left(\frac{\partial H(\mathbf{x})}{\partial \mathbf{x}}\right)^T \mathbf{R}^{-1}[\mathbf{d} - H(\mathbf{L}\chi)] \tag{36}$$

where $\mathbf{L}^*$ is the adjoint of the operator $\mathbf{L}$. Again, the calculation of the background term (Eq. 35) of the gradient is no longer an issue and it can be verified that the gradient of the observation term (Eq. 36) corresponds to the operator $\mathbf{L}^*$ applied to the results of Eq. (29), i.e. $\nabla_\chi J^o = \mathbf{L}^* \nabla_\mathbf{x} J^o$. With χ as the control variable, the iterative sequence of operations necessary to minimize J is:

1. Initialization: read $\mathbf{x}^b$ and $\mathbf{y}$, set $\chi = 0$

2. Calculate J^b by Eq. (32)

3. Get $\mathbf{x} = \mathbf{L}\chi + \mathbf{x}^b$ and calculate J^o by Eq. (26)

4. Get $J = J^b + J^o$

5. Calculate $\nabla_\chi J^b$ by Eq. (35)

6. Calculate $\nabla_\mathbf{x} J^o$ by Eq. (29)

7. Get $\nabla_\chi J = \nabla J^b + \mathbf{L}^* \nabla_\mathbf{x} J^o$

8. Provide χ, J and $\nabla_\chi J$ to the minimizer and update χ

9. If convergence, get $\mathbf{x} = \mathbf{L}\chi + \mathbf{x}^b$ and start a forecast until the next analysis time. Otherwise, go to step 2.

4 Formulation of L and its adjoint L*

By definition, the background error covariance matrix $\mathbf{B}$ is the product of a standard deviation error matrix $\mathbf{\Sigma}$ (diagonal) and a correlation matrix $\mathbf{C}$ (non diagonal):

$$\mathbf{B} = \mathbf{\Sigma}\mathbf{C}\mathbf{\Sigma} \tag{37}$$

or

$$\mathbf{L} = \mathbf{\Sigma}\mathbf{C}^{1/2} . \tag{38}$$

In this study, we will only deal with univariate correlations such that no species-species correlations are considered in BASCOE. As the matrix $\mathbf{C}^{1/2}$ is huge, it is necessary to make assumptions and/or transformations to allow its implementation in a variational system (see Sect. 1). In the NWP community, it was suggested to build the correlation matrix in the spectral space (see e.g. Courtier et al., 1998). As mentioned in Sect. 2.6, assuming homogeneous and isotropic horizontal correlations at the surface of a sphere allows one to get a diagonal correlation matrix in the spectral space. This means that considering a 3-D model on the sphere, the spatial correlations are then represented by a block diagonal matrix. Let $\mathbf{\Lambda}^{1/2}$ denote the spectral representation of $\mathbf{C}^{1/2}$ and $\mathbf{S}$ denote the spectral transform operator. Hence, the transformation $\mathbf{C}^{1/2}$ can be rewritten as $\mathbf{S}\mathbf{\Lambda}^{1/2}$, so that $\mathbf{L}$ becomes:

$$\mathbf{L} = \mathbf{\Sigma}\mathbf{G}\mathbf{S}\mathbf{\Lambda}^{1/2} . \tag{39}$$

The operator $\mathbf{G}$ has been introduced in Eq. (39) to account for the transformation from the target grid of the spectral transform to the model grid. In the case where the analysis increments are calculated on the Gaussian grid, the transformation $\mathbf{G}$ is represented by a weighting average over the latitudes. In the case where the spectral grid is the model grid, no transformation is implemented ($\mathbf{G}$ is the identity matrix).

The adjoint operator $\mathbf{L}^*$ is given by

$$\mathbf{L}^* = \mathbf{\Lambda}^{1/2*}\mathbf{S}^*\mathbf{G}^*\mathbf{\Sigma} . \tag{40}$$

Recall that $\mathbf{\Sigma}$ is a diagonal matrix, so the transpose sign $(.)^T$ has been omitted in the equations.

The block diagonal elements of $\mathbf{\Lambda}$ have the dimensions $L \times L$ where L is the number of vertical levels of our model. Each element of a block $\mathbf{\Lambda}_n^m$ takes the form

$$\Lambda_n^m(p, p') = \sqrt{b_n(p)b_n(p')}C_n^v(p, p') \tag{41}$$

where the coefficients b_n are those of Eq. (23) (Berre, 2000, see also the Appendix A4 and p are the level indices of the model) . This is the non-separable formulation of the spatial correlations. By this formulation of $\mathbf{\Lambda}$, the horizontal correlation coefficients b_n depend on the altitude. Moreover, the matrices $\mathbf{C}_n^v$ allow one to take into account different vertical correlations for different wavenumbers. In other words, features with large horizontal scale could have different vertical

correlations than features of small horizontal scale (Fisher, 2003). This specification is necessary in meteorology in order to get correct correlations between the mass and wind fields (Phillips, 1986; Bartello and Mitchell, 1992) while this non-separability has not yet been studied in chemistry data assimilation. Note that the vertical correlation matrices $\mathbf{C}_n^v$ are defined independently of the zonal wavenumber m, as found for the horizontal correlation coefficients b_n. This formulation of the background matrix is then very low demanding in computer resources.

Although this non-separability formulation could be implemented in the present formulation of the correlation matrix $\mathbf{\Lambda}$ in BASCOE, this has not been done in this study. Here, we have assumed that horizontal and vertical correlations are separable, i.e.

$$\mathbf{\Lambda}_n^m = b_n \mathbf{C}^v \qquad (42)$$

where the horizontal correlation coefficients are taken independent of the altitude and where the vertical correlation matrices are independent of the wavenumber n.

Usually, the b_n coefficients and the vertical correlations matrices $\mathbf{C}_n^v$ are estimated or *calibrated* in order to match the statistical a priori errors of an assimilation system. This can be done using methods that allow one to derive an ensemble of error fields from which these statistics are calculated. We will not describe all these calibration methods and will refer to Bannister (2008a, Sect. 5) for a complete review. It is important to note that the calibration method will require an inverse of the operation $\mathbf{S}$ in order to estimate the correlation spectra b_n. If one is using an equally spaced model grid, the exact inverse of $\mathbf{S}$ is not necessary and the method described in Appendix A2 can be used. Even more simpler is to interpolate the error fields on the Gaussian grid before the inversion.

In the current study, the correlation matrix of BASCOE has not been calibrated. Instead, a priori correlation functions have been used to build it. In the code provided in the Supplement, two functions have been coded for the horizontal correlations: a Gaussian function

$$f(\theta) = \exp\left(-\frac{1 - \cos\theta}{L^2}\right) \qquad (43)$$

and a second-order autoregressive (SOAR) function

$$f(\theta) = \left(1 + \frac{2\sqrt{1 - \cos\theta}}{L}\right) \exp\left(-\frac{2\sqrt{1 - \cos\theta}}{L}\right) \qquad (44)$$

where $L = L_h/A_0$, A_0 is the Earth's radius and L_h is a horizontal correlation length scale.

For the vertical correlations, again, two functions are coded: a *hat* function

$$C^v(p_i, p_j) = \begin{cases} 1 & \text{for} \quad i = j \\ 1/2 & \text{for} \quad |i - j| = 1 \\ 0 & \text{for} \quad |i - j| > 1 \end{cases} \qquad (45)$$

and a Gaussian function

$$C^v(p_i, p_j) = \exp\left(-\frac{1}{2}\left(\frac{|p_i - p_j|}{L_v}\right)^2\right). \qquad (46)$$

For the Gaussian case, the distance between two levels $|p_i - p_j|$ can be measured either in level indices or in kilometers. Moreover, a correlation length scale L_v must be provided in the same units as p.

Finally, the square root of $\mathbf{C}^v$, which is necessary to build $\mathbf{\Lambda}^{1/2}$ (see Eqs. 39 and 40), is obtained by a singular value decomposition (SVD).

5 Assimilation of a single pseudo-observation

The operators described in Sect. 4 have been implemented in Fortran and tested. This code is provided in a Supplement to this paper. It includes the code for the different operators that define $\mathbf{L}$ and $\mathbf{L}^*$, as well as two tests. All this material is introduced in Appendix B.

This section presents the results of an assimilation of a single pseudo-observation using the 3D-Var method. The goal is to evaluate the implementation of the $\mathbf{B}$ matrix by an assimilation for which the result can be calculated analytically. Here, we choose to assimilate a pseudo-observation with a value of 1.2 and located at a model grid point. The background field $\mathbf{x}^b$ has a constant value of 1. Both the observation and the background field have an error standard deviation of $\sqrt{0.02}$. With this configuration, the value of J after assimilation is expected to be $1/2$. The optimization is performed by the quasi-Newton algorithm M1QN3 (Gilbert and Lemarechal, 1989).

In the assimilation experiments discussed here, the model grid is equally spaced and two different kinds of systems have been used. In the first one, the spectral transform operates form the spectral space to the equally spaced model grid. In the second case, the spectral transform operates from the spectral space to the Gaussian grid and a mapping operation from the Gaussian grid to the model grid is used. In the following part of this section, the analyses obtained with the two systems will be denoted, respectively, LL (i.e. lat/lon grid) and GG (i.e. Gaussian grid).

Several model resolutions, locations of the observation and (horizontal and vertical) correlation functions have been considered in these experiments. Here are shown only the results obtained with specifics but representative configuration of the experiment. This configuration considers a Gaussian correlation function with correlation length scales L_h of 600 km horizontally and L_v of 3 model levels vertically. The resolution of the model is 120 longitudes by 60 latitudes by 31 levels. Accordingly, the degree of truncation is 59. Three locations of the observation have been considered, namely the Equator, 40° N and 80° N. In such a configuration, the theoretical analysis at the observation location is 1.1. From that point, the analysis should decrease according to a Gaussian

Table 1. Scores of the single-observation assimilations using two different target grids of the spectral transform: the model lat/lon grid (LL) and the Gaussian grid (GG). The experiments have been performed considering three latitudes of the observations: the Equator, 40° N and 80° N. We report the value of the objective function J for the analysis, the analysis at the observation location $H(\mathbf{x})$, the analysis correlation length scales obtained by fitting the analysis cross-sections (see Fig. 1) by a Gaussian function, and the root mean square (RMS) between the analysis and the expected Gaussian function.

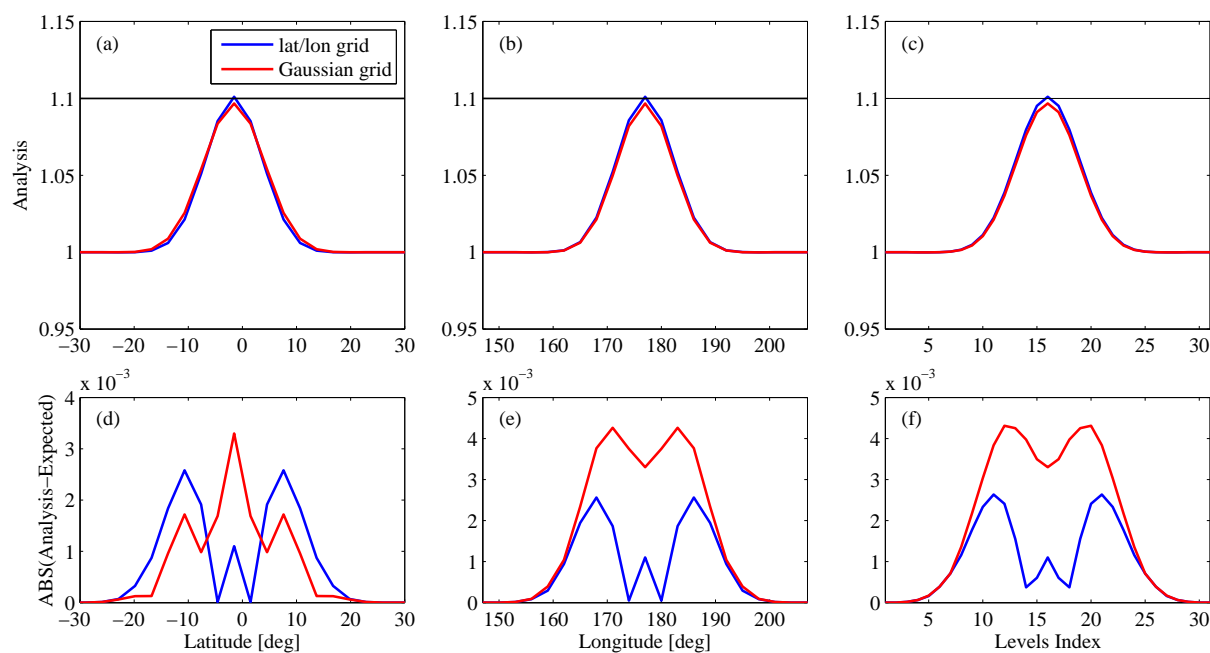|	Expected	Equator		40° N		80° N	
		LL	GG	LL	GG	LL	GG
J	0.5	0.5029	0.5208	0.5020	0.5200	0.5013	0.5166
$H(\mathbf{x})$	1.1	1.1015	1.0966	1.101	1.0971	1.0999	1.0967
L_h^{lat}	600	576	622	576	619	581	622
L_h^{lon}	600	575	574	576	575	601	607
L_v	3	2.88	2.87	2.88	2.87	2.9	2.9
RMS_{lat}	0	3.0×10^{-5}	2.6×10^{-5}	2.9×10^{-5}	2.0×10^{-5}	2.3×10^{-5}	2.0×10^{-3}
RMS_{lon}	0	3.1×10^{-5}	1.2×10^{-4}	3.7×10^{-5}	1.3×10^{-4}	7.3×10^{-5}	4.0×10^{-4}
RMS_{lev}	0	5.4×10^{-5}	2.0×10^{-4}	5.2×10^{-5}	1.7×10^{-4}	4.7×10^{-5}	1.6×10^{-3}

Fig. 1. Cross-sections of the analyses obtained by the assimilation of a single pseudo-observations located at the Equator using Gaussian spatial correlation function and two target grids of the spectral transform: the lat/lon grid (blue line) and the Gaussian grid (red line). The correlation length scales are $L_h = 600$ km horizontally and $L_v = 3$ model levels vertically. The three top panels show the cross-section of theses analyses along the latitude **(a)**, the longitude **(b)** and the altitude **(c)**, all crossing the observation location. The horizontal black lines at value 1.1 indicate the theoretical analyses at the observation location. The bottom panels display the differences between the analysed cross-sections and the expected cross-sections.

bell curve with a standard deviation that corresponds to the correlation length scales, i.e. 600 km horizontally and 3 levels vertically.

Figure 1 displays the results of the experiments that consider the observation at the Equator. It shows three cross-sections of the analyses (upper panels) – along the latitude, the longitude and the vertical directions, all of them crossing the observed location. The differences of the analysis cross-sections with the expected cross-sections are shown in the lower panels of Fig. 1. Both analyses present a maximum at the location of the observation. As the distance to the observation location increases, the analyses decrease with a Gaussian shape. The GG analysis has a value at the observation location which is significantly different than the expected value of 1.1. On the other hand, the LL analysis is (visually) closer to this expected value. This is confirmed by the Fig. 1d–f
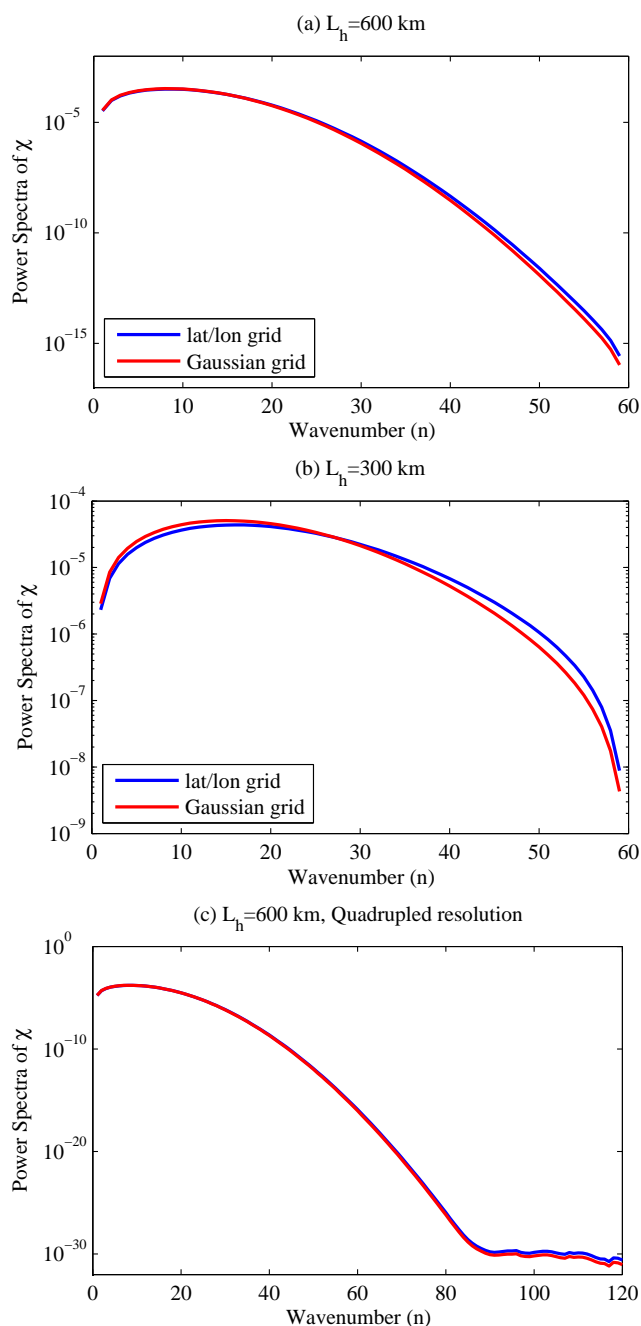
Fig. 2. Power spectra of the analysis increments χ at the observation level obtained using two different spectral grids – the lat/lon grid (blue line) and the Gaussian grid (red line). Panels **(a)** and **(b)** are obtained using a correlation length scales of $L_h = 600\,km$ **(a)** and $300\,km$, respectively. In both cases, the experiments are based on a model resolution of 120 longitudes by 60 latitudes and 31 levels. Results on panel **(c)** are obtained with $L_h = 600\,km$ and a model resolution of 240 longitudes by 120 latitudes and 31 levels.

where it can be seen that the LL analysis is usually closer to theoretical cross-sections than the GG analysis.

Table 1 reports several scores obtained by the different experiments and their expected theoretical values. For each (computed or theoretical) analysis, the table provides the value of the objective function J, the analysis value at the observation point $H(\mathbf{x})$ and the correlation length scales estimated by a fit of the cross-section lines in Figs. 1a–c by a Gaussian function. The table also provides the value of the root mean square (RMS) of the differences between the analyzed and the theoretical cross-section. Compared with the GG analysis, the values of J, $H(\mathbf{x})$ and RMS from the LL analysis are closer to the expected value. The correlation length scales obtained by both analyses are close to each other. However, correlation length scales found by the latitude and longitude cross-sections from the LL experiments are more consistent – they do not differ from more than 1 km. This is not the case for the GG experiments where the correlation length scales from the latitude cross-section is around 620 km while those from the longitude cross-section are around 575 km – around 45 km of differences.

We have also noted that the "fitted" correlation length scales are slightly different from the expected values, around 575 km instead of 600 horizontally and 2.88 levels instead of 3 vertically. The reason for these deviations remains unknown. This issue has not been investigated as we found satisfactory the results of the single assimilation experiments and real data assimilation (see Sect. 6).

Another way of comparing both analyses is to look at the power spectrum of their increments, which are displayed at the model level of the observation location (Fig. 2a). Recall that the analysis increments are the difference between the analysis and the background field and measure the impact of the observations on the analyses. We see that the information provided by the increments of both analyses differs according to the wavenumber. This is even clearer when the horizontal correlation length scale decreases up to 300 km, i.e. a value closer to the resolution of the grid in the region of the observation (Fig. 2b). Taking the LL experiments as references, Fig. 2a–b tell us how the mapping operation from the Gaussian grid to the model grid degrades the GG experiments. At low frequencies (i.e. high horizontal scales), the GG analysis provides more information but decreases more rapidly than the LL analysis, this later providing more information at high frequencies (i.e. small horizontal scales). Therefore, the mapping operation tends to overestimate the spatial correlations for high horizontal scale features and to underestimate the spatial correlations for small horizontal scale features.

Finally, the power spectra of the analysis increments of two additional LL and GG experiments (one each) are displayed on Fig. 2c. These experiments consider a model horizontal resolution four times higher (240 longitudes by 120 latitudes) than the previous experiments. Those experiments have been run in order to get the optimal size of the model resolution. Considering $L_v = 600\,km$, one sees that the increments does not provide additional information
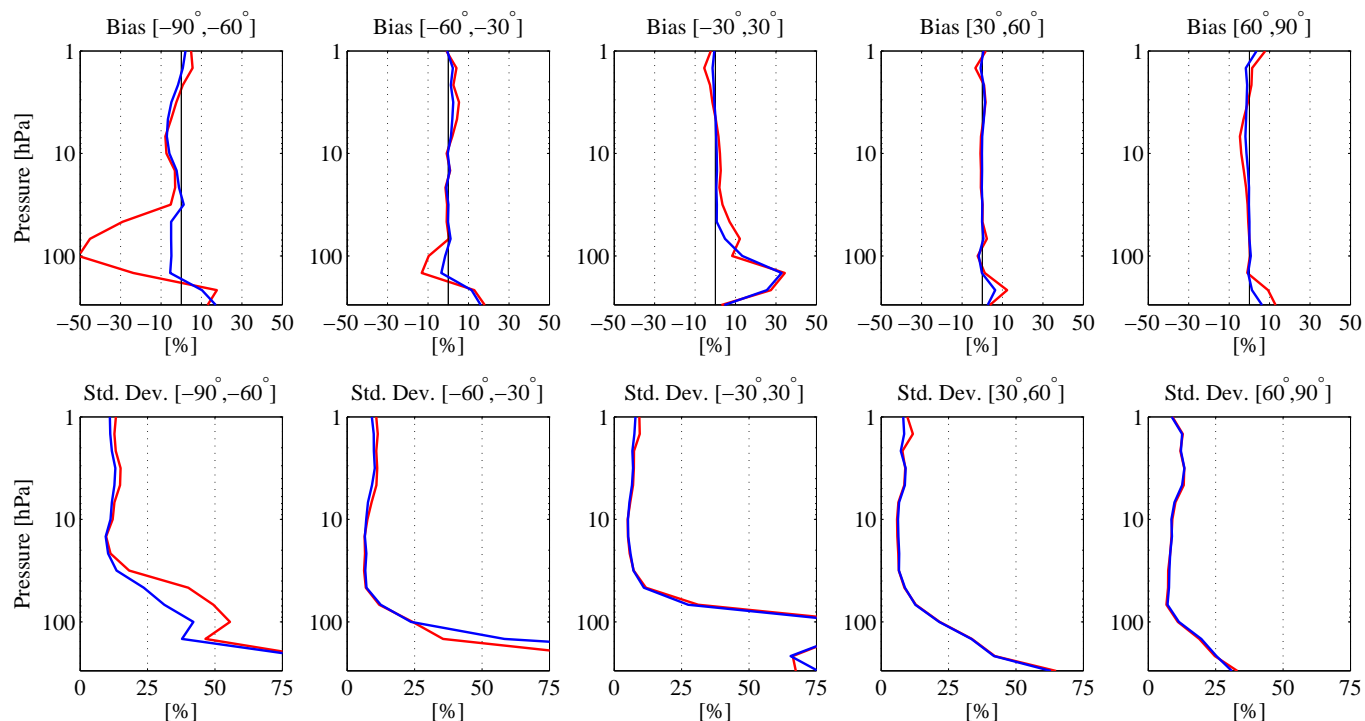
Fig. 3. Bias (top row) and standard deviation (bottom row) of the Observations-minus-Forecast (OmF) statistics between the analyses and the MIPAS data, in percent. Statistics are done at five latitude bands denoted in the title of each plot. Blue and red lines represent the statistics obtained, repectively, by the CORREL and DIAG runs.
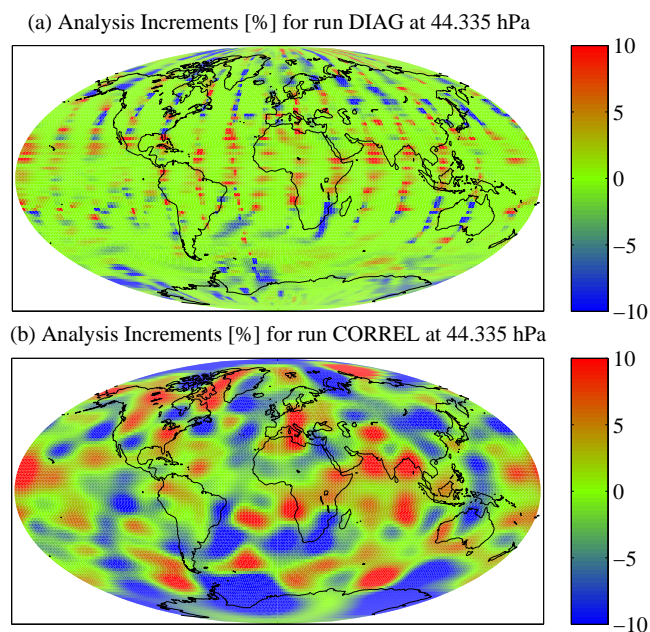

(a) Analysis Increments [%] for run DIAG at 44.335 hPa

(b) Analysis Increments [%] for run CORREL at 44.335 hPa

Fig. 4. Analysis increments of the runs DIAG **(a)** and CORREL **(b)** at the model level 21 (∼44 hPa) on 15 September 2003, in percent. Blue and red indicate, respectively, negative and positive increments while green indicates close to zero increments.

above wavenumber ∼85. As a consequence, a model resolution of about 180 longitudes by 90 latitudes appears to be optimal. While this result is valid for the LL and GG grid, it is restricted to the correlation length scales of 600 km. On the other hand, the the optimal resolution increase as L_v decrease. For example, with $L_v = 300$ km, one finds (not shown) an optimal truncation of ∼180 or a resolution of 360 longitudes by 180 latitudes. In the next section, assimilation of real data will then be performed using a model horizontal resolution of 180 longitudes by 90 latitudes with a correlation length scales of 600 km.

6 Assimilation of real data

In this section, the results of two chemical data assimilation experiments are presented. The first one includes correlations in the **B** matrix while the second one considers a diagonal **B**. The experiments have been performed by means of the 4D-Var Belgian Assimilation System for Chemical ObsErvations (BASCOE, Errera et al., 2008). In its usual configuration, this system considers 57 stratospheric species advected by the Flux-Form Semi-Lagrangian scheme (Lin and Rood, 1996), 200 chemical reactions and a parameterization of the physico-chemical processes due to the so-called Polar Stratospheric Clouds (PSCs). However, in this study, only the advection of ozone is considered (i.e. the chemical and PSC
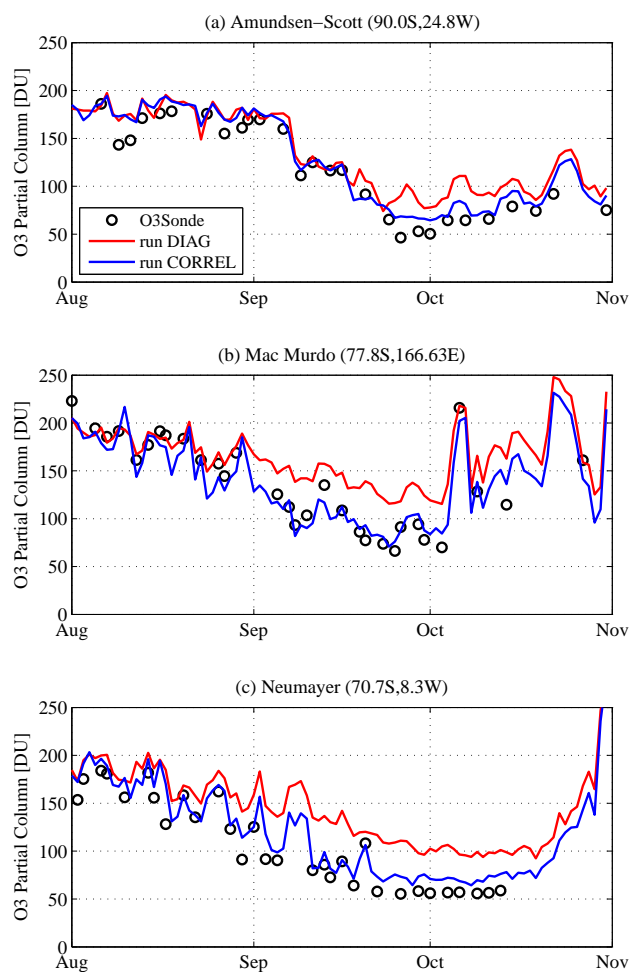
Fig. 5. Time series of the ozone partial column (10–100 hPa) between August and October 2003 obtained above three NDACC stations in Antarctica by the ozone sondes (black circles), and the runs DIAG (red line) and CORREL (blue line), in Dobson unit (DU).

schemes have been turned off) in order to reduce the CPU time. By doing this, it is assumed that ozone behaves like an inert tracer. This is a fair assumption by choosing an assimilation window of one day and by excluding observations above 1 hPa where the ozone time scale is shorter than the assimilation window.

The horizontal resolution is set to $2° \times 2°$ lat/lon grid. This is higher than past experiments based on $3.75° \times 5°$ lat/lon (Geer et al., 2006; Errera et al., 2008; Viscardy et al., 2010) or $2.5° \times 3.75°$ lat/lon grid (Lahoz et al., 2011, and the BASCOE Near Real Time service – http://macc.aeronomie.be). The vertical grid is represented by the BASCOE usual 37 vertical levels from the surface to 0.1 hPa, these levels being a subset of the ECMWF levels. In these experiments, the dynamics is provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim analyses (Dee et al., 2011). The assimilated ozone data are taken by the

Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) ESA off-line level 2 profiles between August and October 2003, as in BASCOE past experiments discussed by Geer et al. (2006) and Errera et al. (2008). MIPAS is a limb fourier transform spectrometer onboard the Envisat platform operating between 2002 and 2012. Measuring in the infrared, limb scans are inverted to provide profiles of numerous trace gases, including ozone (Fischer et al., 2008).

Both experiments described in this paper consider 30 % standard deviation in the background fields. The experiment in which spatial correlations are introduced – denoted CORREL – considers a Gaussian correlation function in the horizontal and vertical directions with a correlation length scale $L_h = 600$ km and $L_v = 1$ level, respectively. Note that this combination of model horizontal resolution and horizontal correlation length scale is optimal as mentioned in Sect. 5. The experiment using a diagonal **B** is denoted DIAG.

Figure 3 shows the Observations-minus-Forecast (OmF) statistics for both runs and for the period of the 2003 Antarctic ozone hole (September–October 2003) and for five latitude bands as explained in Errera et al. (2008, Sect. 3.4). Generally, the biases of the OmF obtained by the CORREL run are lower than those obtained by the DIAG run. This is significant in the upper troposphere/lower stratosphere (UTLS) region, between 70 and 300 hPa, particularly so in the southern polar region. Regarding the standard deviations, the statistics of both runs are very similar except at the South Pole.

To understand the origin of the differences between both runs, their analysis increments on September 15 at ~44 hPa (i.e. at model level 21) are shown in Fig. 4. In the DIAG increments, the track of the satellite is clearly visible and the increments are confined around the observation locations. On the other hand, the CORREL increments are spread over a much larger area, especially in southern high latitudes, thanks to the non-diagonal nature of **B**. As shown by the OmF statistics, this improves the assimilation.

To assess the improvements at the southern high latitudes brought by the non-diagonal **B**, both experiments have been compared with ozonesonde observations at three NDACC (Network for the Detection of Atmospheric Composition Change) stations in Antarctica (Amundsen-Scott, Mac Murdo and Neumayer). Figure 5 presents the time series of the ozone partial column between 10 and 100 hPa during the formation of the ozone hole in the 2003 Antarctic winter, along with the corresponding partial column from the analyses. As expected, the CORREL run agrees much better with the observations than the DIAG run. As a result, CORREL exhibits a deeper ozone hole than DIAG which is in better agreement with observations provided by the Total Ozone Mapping Spectrometer (TOMS) (Fig. 6).

The comparison of both runs against ozone sondes in the UTLS has not revealed any "winner" or "looser" (not shown). The possible reasons are that MIPAS observations are of lower quality in the UTLS. Around 100 hPa, it has been
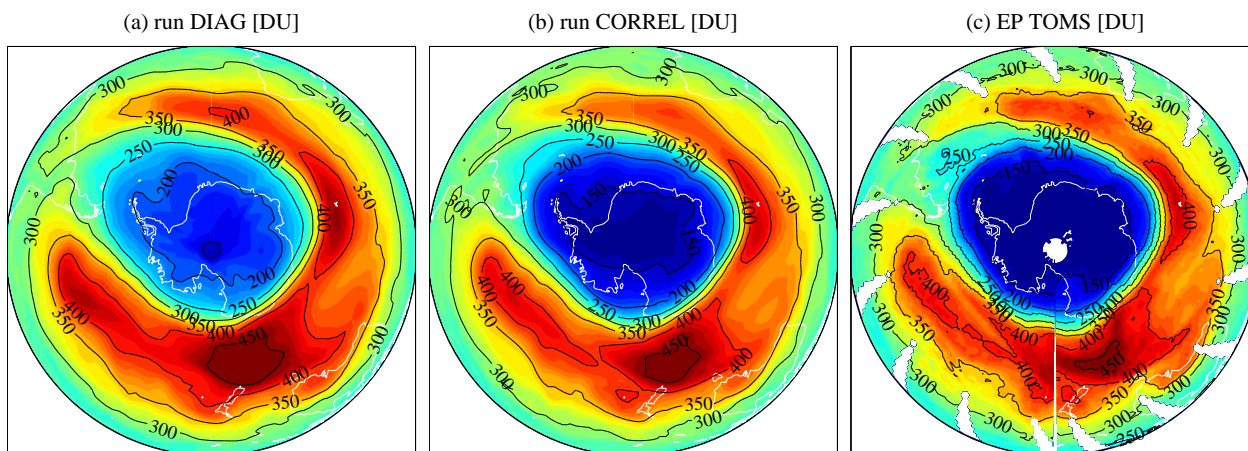
(a) run DIAG [DU]   (b) run CORREL [DU]   (c) EP TOMS [DU]

Fig. 6. Total ozone in the Southern Hemisphere on 1 October 2003 obtained by the runs DIAG **(a)** and CORREL **(b)**, and observed by the Total Ozone Measurement Satellite (TOMS) **(c)**, in Dobson unit (DU).

found that MIPAS O_3 values are 5 % to 25 % larger compared to the majority of the other datasets (Cortesi et al., 2007). Removing the bias in the MIPAS data could help to improve the analyses in this region. Nevertheless, this task has been left for future studies. Moreover, the parameters of **B** in the COR-REL run have not been optimized. For instance, nothing tells us that the background error standard deviation is 30 % of the first guess field, that the correlations have a Gaussian shape neither that the length scales are 600 km horizontally at every model level or 1 level vertically. In principle, a calibrated **B** should improve the analyses. Several methods have been developed by the NWP centers and are reviewed by Bannister (2008a, Sect. 5). This optimization effort has not been implemented in the present study. Indeed, we aimed to focus on the implementations of the correlations and not essentially on the BASCOE results. We nevertheless plan to present the optimization of the parameters of **B** in future papers.

7   Conclusions

There are several outcomes from this work. First, we have gathered together all the material published so far in order to provide a stand alone document describing the spherical harmonic representation of the background error covariance matrix (the **B** matrix) in the case of univariate assimilation. The motivation of a spherical harmonic representation of **B** is based on the fact that, if horizontal correlations are assumed homogeneous and isotropic, **B** can be represented in the spectral space by a block diagonal matrix with repeated matrices for each zonal wavenumber m, which makes its numerical computation feasible. Horizontal and vertical correlations can be implemented in two ways. The first one consists of providing correlation functions for the vertical and the horizontal directions (not necessarily the same) along with correlation length scales. In the second method, these correlation

matrices are calibrated using optimization methods such as those summarized by Bannister (2008a, Sect. 5). This latter method allows the implementation of non-separable spatial correlations. This has proved to be important in meteorology but has never been studied in chemistry.

The second outcome of this study is to provide the numerical Fortran code of the formulation of the **B** matrix (see the Supplement of this paper). During the course of this implementation, we realized that only the adjoint of the spectral transform operator was necessary for variational assimilation, not the inverse. The problem with the inverse operator is to restrict the target grid of the spectral transform to the (non-equally spaced) Gaussian grid. On the other hand, the adjoint operator does not suffer this limitation thus allowing the use of the (equally spaced) model grid. In this way, there is no need to implement a grid transformation from the Gaussian grid to the lat/lon (model) grid, which degrades the analyses.

The spherical harmonic representation of **B** has been implemented in the 4D-Var stratospheric chemical assimilation system BASCOE using the Michelson Interferometer for Passive Atmosphere Sounding (MIPAS) ozone observations between August and October 2003. Two experiments were performed: the one considering a spectral formulation of **B** and the other considering the diagonal **B** (i.e. spatially uncorrelated) implemented so far in BASCOE. In those experiments, the chemical and PSC schemes of BASCOE were turned off such that the influence of ozone observations on other modeled constituents has not been studied. At southern high latitudes and during the ozone hole period, the spatially correlated **B** allows a great improvement of the analyses with respect to the uncorrelated one. The correlations allow one to increase the size of the analysis increments which allow one to produce ozone fields in good agreement against ozonesondes and total ozone observations, which is not the case for the analyses calculated with the diagonal **B**.

In the future, two studies are planned. The first one is to adapt this formulation for BASCOE in the case where the chemistry module is turned on and to evaluate the improvements to constituents other than ozone. The second study consists of estimating the correlation matrices by one of the methods reviewed by Bannister (2008a, Sect. 5) for separable and non-separable spatial correlations. In the atmosphere, the constituent's concentrations can vary with height by several orders of magnitude and chemical regimes can completely change with altitude, for instance across the tropopause and stratopause. For those reasons, the horizontal and vertical lengths scales are likely to be strongly dependent on height. Thus a non-separable three-dimensional error correlation matrix could provide a better representation of the background error for chemical variables than a separable matrix.

Appendix A

Complement on spherical harmonic representation

A1 Normalization of spherical harmonics

Let us recall the definition of the spherical harmonics:

$$Y_n^m(\lambda, \mu) = N_n^m P_n^m(\mu) e^{im\lambda} \tag{A1}$$

where

$$P_n^m(\mu) = \frac{\left(1 - \mu^2\right)^{m/2}}{2^n n!} \frac{d^{n+m}}{d\mu^{n+m}} (\mu^2 - 1)^n \tag{A2}$$

are the associated Legendre functions. Since

$$\int_0^{2\pi} e^{i(m-m')\lambda} d\lambda = 2\pi \delta_m^{m'} \tag{A3}$$

and

$$\int_{-1}^{1} P_n^m(\mu) P_{n'}^m(\mu) d\mu = \left(\frac{2}{2n+1}\right) \left[\frac{(n+m)!}{(n-m)!}\right] \delta_n^{n'}, \tag{A4}$$

(where we note that associated Legendre functions are not orthogonal for different values of m), we have

$$\int_0^{2\pi} \int_{-1}^{+1} Y_n^m(\lambda, \mu)[Y_{n'}^{m'}(\lambda, \mu)]^* d\mu d\lambda =$$

$$(N_n^m)^2 \delta_n^{n'} \delta_m^{m'} \left(\frac{4\pi}{2n+1}\right) \left[\frac{(n+m)!}{(n-m)!}\right]. \tag{A5}$$

The normalization constants N_n^m take a different form following the choice of inner-product of the spherical harmonics. For practical reasons, there are several ways to normalize

the spherical harmonics

$$\int_\Omega Y_n^m(\Omega) Y_{[n']}^{[m']}{}^*(\Omega) d\Omega = \begin{cases} \frac{4\pi}{2n+1} \frac{(n+m)!}{(n-m)!} \delta_n^{n'} \delta_m^{m'} & \text{UN} \\ \frac{4\pi}{2n+1} \delta_n^{n'} \delta_m^{m'} & \text{Schmit} \\ \delta_n^{n'} \delta_m^{m'} & \text{ON} \\ 4\pi \delta_n^{n'} \delta_m^{m'} & \text{G4}\pi \\ 2\pi \delta_n^{n'} \delta_m^{m'} & \text{G2}\pi \end{cases} \tag{A6}$$

where the abbreviations have the following meaning: UN for *unormalized*, Schmit for *Schmit semi-normalized*, ON for *orthonormalized*, G4π for *geodesy 4π* (as in Courtier et al., 1998) and G2π for *geodesy 2π* (as in this paper and in Boer, 1983). These different normalizations of the spherical harmonics define the value of the normalization constant:

$$N_n^m = \begin{cases} 1 & \text{UN} \\ \left[\frac{(n-m)!}{(n+m)!}\right]^{1/2} & \text{Schmit} \\ \left(\frac{2n+1}{4\pi}\right)^{1/2} \left[\frac{(n-m)!}{(n+m)!}\right]^{1/2} & \text{ON} \\ \sqrt{2n+1} \left[\frac{(n-m)!}{(n+m)!}\right]^{1/2} & \text{G4}\pi \\ \left(\frac{2n+1}{2}\right)^{1/2} \left[\frac{(n-m)!}{(n+m)!}\right]^{1/2} & \text{G2}\pi. \end{cases} \tag{A7}$$

For practical reasons, computer code implementations often introduce the normalization factor into the associated Legendre functions (which is done in the code provided in the Supplement). In this case, we define the normalized associated Legendre functions: $\overline{P}_n^m(\mu) = N_n^m P_n^m(\mu)$. Accordingly, the computation of spherical harmonics is decomposed into the product of two functions or two operations, a Fourier term and a Legendre term. For the normalized associated Legendre functions, we have

$$\int_{-1}^{1} \overline{P}_n^m(\mu) \overline{P}_{n'}^m(\mu) d\mu = \begin{cases} \frac{2}{2n+1} \frac{(n-m)!}{(n+m)!} \delta_n^{n'} & \text{UN} \\ \frac{2}{2n+1} \delta_n^{n'} & \text{Schmit} \\ \frac{1}{2\pi} \delta_n^{n'} & \text{ON} \\ 2\delta_n^{n'} & \text{G4}\pi \\ \delta_n^{n'} & \text{G2}\pi. \end{cases} \tag{A8}$$

One easy way to check the choice of normalization is to verify that $\overline{P}_n^0(\mu = 0) = N_n^0$, i.e.

$$\overline{P}_n^0(0) = \begin{cases} 1 & \text{UN} \\ 1 & \text{Schmit} \\ \left(\frac{2n+1}{4\pi}\right)^{1/2} & \text{ON} \\ \sqrt{2n+1} & \text{G4}\pi \\ \left(\frac{2n+1}{2}\right)^{1/2} & \text{G2}\pi. \end{cases} \tag{A9}$$

However, note that this check does not make any difference between UN and Schmit normalization.

Finally, let us mention the following property $\overline{P}_n^{-m} = (-1)^m \overline{P}_n^m$ (which is not valid for unormalized associated Legendre functions). It follows that $[Y_n^m]^* = (-1)^m Y_n^{-m}$.

A2   Inverse of spectral transform from non-Gaussian grid

It is worth noting that the inverse transform on other grids than the Gaussian grid can be made by a quadrature developed by Machenhauer and Daley (1972), which has the same numerical accuracy (although not exact) as the method based on the Gaussian grid (see also Swarztrauber, 1979; Swarztrauber and Spotz, 2000; Machenhauer, 1979, Sect. 4.6):

$$\Psi_n^m = \frac{1}{M} \sum_{k=1}^{K} \sum_{j=0}^{M-1} \Psi(\lambda_j, \phi_k) \overline{Z}_n^m(\phi_k) e^{-im\lambda_j} \tag{A10}$$

$$\overline{Z}_n^m(\phi_k) = \frac{1}{K} \sum_{r=0}^{N} \delta L_r(\phi_k) \int_{-1}^{1} L_r(\phi) \overline{P}_n^m(\sin\phi) d\mu \tag{A11}$$

$$\delta = \begin{cases} 1/2 & \text{for} \quad r = 0 \\ 1 & \text{for} \quad r \neq 0 \end{cases} \tag{A12}$$

and

$$L_r(\phi) = \begin{cases} \cos(r\phi) \begin{cases} m \quad \text{and} \quad r \quad \text{odd} \\ m \quad \text{and} \quad r \quad \text{even} \end{cases} \\ \sin(r\phi) \begin{cases} m \quad \text{even} \quad \text{and} \quad r \quad \text{odd} \\ m \quad \text{odd} \quad \text{and} \quad r \quad \text{even.} \end{cases} \end{cases} \tag{A13}$$

The integral on the right-hand side of Eq. (A11) can be computed exactly by a Gaussian quadrature.

The code available in the Supplement provides the routine to calculate the inverse operation $\mathbf{S}^{-1}$ from the Gaussian grid (but not from the equally spaced lat/lon grid). Such a routine can be found in the SPHEREPACK module available at www2.cisl.ucar.edu/resources/legacy/spherepack.

A3   Addition theorem of spherical harmonics

The derivation of the spectral representation of homogeneous and isotropic horizontal correlations requires the use of the addition theorem of spherical harmonics. Suppose we have two points on a unit sphere $\Omega = (\lambda, \mu)$ and $\Omega' = (\lambda', \mu')$. Let θ denote the great circle angular separation between those points which is defined by

$$\cos\theta = \sin\phi \sin\phi' + \cos\phi \cos\phi' \cos(\lambda' - \lambda) \tag{A14}$$

where $\phi = \sin^{-1}\mu$. The addition theorem of spherical harmonics states that (Jackson, 1998, Sect. 3.6):

$$\sum_{m=-n}^{n} Y_n^m(\Omega)[Y_n^m(\Omega')]^* = (N_n^0)^2 P_n(\cos\theta) \tag{A15}$$

or

$$\sum_{m=-n}^{n} Y_n^m(\Omega) \left[Y_n^m(\Omega')\right]^* = N_n^0 \, \overline{P}_n(\cos\theta) . \tag{A16}$$

Table A1. List of operators and matrix operations that define $\mathbf{L}$ and $\mathbf{L}^*$, and the routines in which they are coded.

Operator	corresponding routine
$\mathbf{S}$	S_DIRECT
$\mathbf{S}^*$	S_ADJ
$\mathbf{S}^{-1}$	S_INV[a]
$\mathbf{\Lambda}^{1/2}$	SQRT_CORREL
$\mathbf{G}$	SG2MG
$\mathbf{G}^T$	AD_SG2MG
$\mathbf{\Sigma}$	B_STDDEV

[a] $\mathbf{S}^{-1}$ is only coded for the transform from the Gaussian grid to the spectral space.

A4   Homogeneous and isotropic horizontal correlations

Let $\epsilon(\Omega)$ be a 2-D error field on the sphere as defined in Sect. 2.6 and ϵ_n^m its spectral coefficients. The physical representation of the correlations $\langle \epsilon(\Omega)[\epsilon(\Omega')]^* \rangle$ and its spectral counterpart $\langle \epsilon_n^m [\epsilon_{n'}^{m'}]^* \rangle$ are related by

$$\langle \epsilon(\Omega)[\epsilon(\Omega')]^* \rangle = \sum_{n=0}^{N} \sum_{m=-n}^{n} \sum_{n'=0}^{N} \sum_{m'=-n'}^{n'} \langle \epsilon_n^m [\epsilon_{n'}^{m'}]^* \rangle Y_n^m(\Omega)[Y_{n'}^{m'}(\Omega')]^* \tag{A17}$$

and by the inverse relationship (using a geodesy 2π norm)

$$\langle \epsilon_n^m [\epsilon_{n'}^{m'}]^* \rangle = \frac{1}{4\pi^2} \int_{\Omega} \int_{\Omega'} \langle \epsilon(\Omega)[\epsilon(\Omega')]^* \rangle Y_n^m(\Omega)[Y_{n'}^{m'}(\Omega')]^* d\Omega d\Omega' . \tag{A18}$$

Assuming homogeneous and isotropic covariances using Eq. (22), Eq. (A18) can be rewritten as

$$\langle \epsilon_n^m [\epsilon_{n'}^{m'}]^* \rangle = \frac{1}{4\pi^2} \sum_{n''}^{N} f_{n''} \int_{\Omega} \int_{\Omega'} \overline{P}_{n''}(\cos\theta) Y_n^m(\Omega)[Y_{n'}^{m'}(\Omega')]^* d\Omega d\Omega' \tag{A19}$$

Using the addition theorem of spherical harmonics, Eq. (A19) becomes

$$\langle \epsilon_n^m [\epsilon_{n'}^{m'}]^* \rangle = \frac{1}{4\pi^2} \sum_{n''}^{N} f_{n''} (N_{n''}^0)^{-1}$$
$$\int_{\Omega} \int_{\Omega'} Y_n^m(\Omega) \sum_{m''=-n''}^{n''} [Y_{n''}^{m''}(\Omega)]^* Y_{n''}^{m''}(\Omega')[Y_{n'}^{m'}(\Omega')]^* d\Omega d\Omega' \tag{A20}$$

which yields

$$\langle \epsilon_n^m [\epsilon_{n'}^{m'}]^* \rangle = b_n \delta_n^{n'} \delta_m^{m'} ; \quad b_n \equiv f_n / N_n^0 \tag{A21}$$

as a result of the orthogonality of the spherical harmonics. Equation (A21) is a necessary condition when correlations are homogeneous and isotropic. By substituting Eq. (A21) in (A17) and using the addition theorem of spherical harmonics, one finds Eq. (22), thus proving that Eq. (A21) is also sufficient.

Let us see how this result is extended in three dimensions. Let $\zeta(\Omega, p)$ be a 3-D error field on the sphere where p denotes the altitude. Again, we assume that ζ is unbiased and

Table A2. List of precalculated variables and their Fortran counterpart.

Variable names	corresponding Fortran names
μ	mu
λ	lambda
w	wg
P_n^m	Pnm
$e^{im\lambda}$	Eiml
$\mathbf{G}$	G
$\mathbf{G}^T$	G_TRANSP
$(\mathbf{C}^{\mathrm{v}})^{1/2}$	Cv
b_n	Values stored in Ch
$\boldsymbol{\Sigma}$	bg_std

normalized such that $\langle \zeta(\Omega, p)^2 \rangle = 1$. Equation (A21) obtained in two dimensions is generalized in three dimensions as follows:

$$\langle \zeta_n^m [\zeta_{n'}^{m'}]^* \rangle = \sqrt{b_n(p)b_n(p')} C_n^{\mathrm{v}}(p, p') \delta_n^{n'} \delta_m^{m'} . \qquad \text{(A22)}$$

where $\mathbf{C}_n^{\mathrm{v}}$ is the vertical correlation matrix for the wavenumber n and the coefficients b_n are now dependent on the altitude.

If $\mathbf{C}_n^{\mathrm{v}}$ and $b_n(p)$ are independent of n and p, respectively, (A22) might be rewritten as

$$\langle \zeta_n^m(p) [\zeta_{n'}^{m'}(p')]^* \rangle = b_n C^{\mathrm{v}}(p, p') \delta_n^{n'} \delta_m^{m'} \qquad \text{(A23)}$$

and we can prove that the spatial correlations are separable. By definition, a matrix is said separable if it can be calculated by a tensor product of two sub-matrices. In this case, if the horizontal and vertical correlation matrices are denoted by $\mathbf{C}^{\mathrm{h}}$ and $\mathbf{C}^{\mathrm{v}}$, then the 3-D spatial correlation matrix $\mathbf{C}$ is defined by

$$\mathbf{C} = \mathbf{C}^{\mathrm{h}} \otimes \mathbf{C}^{\mathrm{v}} \qquad \text{(A24)}$$

where $\otimes$ denotes the Kronecker tensor product. In terms of matrix elements, and considering homogeneous and isotropic correlations, we have

$$C(\Omega, p, \Omega', p') = f(\theta) C^{\mathrm{v}}(p, p') . \qquad \text{(A25)}$$

Indeed, the correlations in the physical space might have the following spectral representation

$$\langle \zeta(\Omega, p)[\zeta(\Omega', p')]^* \rangle = \sum_{n=0}^{N} \sum_{m=-n}^{n} \sum_{n'=0}^{N} \sum_{m'=-n'}^{n'}$$
$$\langle \zeta_n^m(p) [\zeta_{n'}^{m'}(p')]^* \rangle Y_n^m(\Omega) [Y_{n'}^{m'}(\Omega')]^* . \text{ (A26)}$$

By using Eq. (A23), Eq. (A26) takes the form

$$\langle \zeta(\Omega, p)[\zeta(\Omega', p')]^* \rangle = C^{\mathrm{v}}(p, p') \sum_{n=0}^{N} \sum_{m=-n}^{n} b_n Y_n^m(\Omega)[Y_n^m(\Omega')]^* . \text{ (A27)}$$

and using the Eq. (A16), the previous equation can be rewritten as

$$\langle \zeta(\Omega, p)[\zeta(\Omega', p')]^* \rangle = C^{\mathrm{v}}(p, p') \sum_{n=0}^{N} N_n^0 b_n \overline{P}_n(\cos\theta) \text{ (A28)}$$

which leads to Eq. (A25). If spatial correlations are separable, Eq. (A28) is a necessary condition. To prove that the condition is sufficient, we substitute Eq. (A28) in

$$\langle \zeta_n^m(p) [\zeta_{n'}^{m'}(p')]^* \rangle = \frac{1}{4\pi^2} \int_\Omega \int_{\Omega'} \langle \zeta(\Omega, p)[\zeta(\Omega', p')]^* \rangle$$
$$Y_n^m(\Omega) [Y_{n'}^{m'}(\Omega')]^* d\Omega d\Omega' . \qquad \text{(A29)}$$

and using the orthogonality of the spherical harmonics and the addition theorem, we can find Eq. (A23).

Appendix B

Fortran code

The Supplement of this paper provides a code that allows calculation of the operations $\mathbf{L}$ and $\mathbf{L}^*$ along with two test programs. The operations related to the $\mathbf{B}$ matrix are given in bglib_sh.f90 (which means "Back-Ground LIBrary and Spherical Harmonics representation"). The length of this file is rather short. Moreover, we tried to use Fortran variable names that are close to those presented here. The correspondence between the different operators discussed in Sect. 3 and their coded counterpart are given in Table A1. A very important routine is PRECOMPUTE_BGLIB which precomputes values used by the different operators. They are described in Table A2. The Gaussian latitudes and weights such as the associated Legendre functions are precalculated by the freely available software SHTOOLS (available at: http://www.ipgp.fr/~wieczor/SHTOOLS/SHTOOLS.html) which needs to be installed before compiling the code. The vertical correlation matrix is calculated during the precomputing and its square root is calculated using singular value decomposition (SVD). The SVD uses LAPACK libraries which also need to be installed.

In order to allow the user to check if this code fulfills his needs without having to install the SHTOOLS library or LAPACK (while this later is usually installed), this code can also be compiled in a "demo" mode. In this case, the Gaussian latitudes and weights, the associated Legendre functions and the vertical correlation matrix are read in files also provided. However, this mode is only available for a size grid of $120 \times 60 \times 31$ in lon/lat/lev as for vertical Gaussian correlations with $L_{\mathrm{v}} = 3$ levels.

In addition to the library of the spectral representation of $\mathbf{B}$, the code is provided with two tests. The first one checks that the test of the adjoint is satisfied, i.e. $\langle \mathbf{x}, \mathbf{L}\chi \rangle = \langle \mathbf{L}^*\mathbf{x}, \chi \rangle$ for any value of $\mathbf{x}$ and χ. This is the case if we exclude round-off errors. This test is calculated by test_adj.f90.

The second test is the assimilation of a single pseudo-observation as presented in Sect. 5. For this test, a simulator is given (`simul_sh.f90`) and the quasi-Newton minimizer M1QN3 (`m1qn3.f90`, Gilbert and Lemarechal, 1989).

Information necessary to compile and run the code, and the conditions for its use, are provided in a README file given in the Supplement.

In terms of numerical performances, this code is very low demanding in computer resources (CPU and memory). On a single processor, the CPU time consumed by the operations $\mathbf{L}$ and $\mathbf{L}^*$ is lower than the time consumed by M1QN3. For the real assimilation of MIPAS ozone data as discussed in Sect. 6, the operations $\mathbf{L}$ and $\mathbf{L}^*$ take less than 1 % of the total CPU time. On an AMD Opteron 1.15 GHz processor and considering a model resolution of $180 \times 91 \times 37$ grid points, the total time necessary to calculate the operations $\mathbf{L}$ and $\mathbf{L}^*$ is $\sim$1.1 s ($\sim$0.55 s each).

Supplementary material related to this article is available online at: http://www.atmos-chem-phys.net/12/10015/2012/acp-12-10015-2012-supplement.zip.

References

Bannister, R. N.: 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, 2008a.

Bannister, R. N.: A review of forecast error covariance statistics in atmospheric variational data assimilation. II: modelling the forecast error covariance statistics, Q. J. R. Meteorol. Soc., 134, 1971–1996, 2008b.

Bartello, P. and Mitchell, H. L.: A continuous three-dimensional model of short-range forecast error covariances, Tellus A, 44, 217, doi:10.1034/j.1600-0870.1992.t01-2-00002.x, 1992.

Berre, L.: Estimation of synoptic and mesoscale forecast error covariances in a limited-area model, Mon. Weather Rev., 128, 644, doi:10.1175/1520-0493(2000)128<0644:EOSAMF>2.0.CO;2, 2000.

Boer, G. J.: Homogeneous and isotropic turbulence on the sphere, J. Atmos. Sci., 40, 154–163, 1983.

Boer, G. J. and Shepherd, T. G.: Large-scale two-dimensional turbulence in the atmosphere, J. Atmos. Sci., 40, 164–184, 1983.

Cortesi, U., Lambert, J. C., De Clercq, C., Bianchini, G., Blumenstock, T., Bracher, A., Castelli, E., Catoire, V., Chance, K. V., De Mazière, M., Demoulin, P., Godin-Beekmann, S., Jones, N., Jucks, K., Keim, C., Kerzenmacher, T., Kuellmann, H., Kuttippurath, J., Iarlori, M., Liu, G. Y., Liu, Y., McDermid, I. S., Meijer, Y. J., Mencaraglia, F., Mikuteit, S., Oelhaf, H., Piccolo, C., Pirre, M., Raspollini, P., Ravegnani, F., Reburn, W. J., Redaelli, G., Remedios, J. J., Sembhi, H., Smale, D., Steck, T., Taddei, A., Varotsos, C., Vigouroux, C., Waterfall, A., Wetzel, G., and Wood, S.: Geophysical validation of MIPAS-ENVISAT operational ozone data, Atmos. Chem. Phys., 7, 4807–4867, doi:10.5194/acp-7-4807-2007, 2007.

Courtier, P., Andersson, E., Heckley, W., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F., Fisher, M., and Pailleux, J.: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: formulation, Q. J. R. Meteorol. Soc., 124, 1783–1807, doi:10.1002/qj.49712455002, 1998.

Dee, D. P.: On-line estimation of error covariance parameters for atmospheric data assimilation, Mon. Weather Rev., 123, 1128, doi:10.1175/1520-0493(1995)123<1128:OLEOEC>2.0.CO;2, 1995.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc., 137, 553–597, doi:10.1002/qj.828, 2011.

Derber, J. and Rosati, A.: A global oceanic data assimilation system, J. Phys. Oceanogr., 19, 1333–1347, doi:10.1175/1520-0485(1989)019<1333:AGODAS>2.0.CO;2, 1989.

Dethof, A. and Hólm, E. V.: Ozone assimilation in the ERA-40 reanalysis project, Q. J. R. Meteorol. Soc., 130, 2851–2872, doi:10.1256/qj.03.196, 2004.

Elbern, H., Schwinger, J., and Botchorishvili, R.: Chemical state estimation for the middle atmosphere by four-dimensional variational data assimilation: system configuration, J. Geophys. Res., 115, D06302, doi:10.1029/2009JD011953, 2010.

Errera, Q., Daerden, F., Chabrillat, S., Lambert, J. C., Lahoz, W. A., Viscardy, S., Bonjean, S., and Fonteyn, D.: 4D-Var assimilation of MIPAS chemical observations: ozone and nitrogen dioxide analyses, Atmos. Chem. Phys., 8, 6169–6187, doi:10.5194/acp-8-6169-2008, 2008.

Fischer, H., Birk, M., Blom, C., Carli, B., Carlotti, M., von Clarmann, T., Delbouille, L., Dudhia, A., Ehhalt, D., Endemann, M., Flaud, J. M., Gessner, R., Kleinert, A., Koopman, R., Langen, J., López-Puertas, M., Mosner, P., Nett, H., Oelhaf, H., Perron, G.,

Remedios, J., Ridolfi, M., Stiller, G., and Zander, R.: MIPAS: an instrument for atmospheric and climate research, Atmos. Chem. Phys., 8, 2151–2188, doi:10.5194/acp-8-2151-2008, 2008.

Fisher, M.: Background error covariance modelling, in: Seminar on Recent developments in data assimilation for atmosphere and ocean, 8–12 September 2003, edited by: ECMWF, Shinfield Park, Reading, 45–64, 2003.

Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, Q. J. R. Meteorol. Soc., 125, 723–757, 1999.

Gauthier, P., Courtier, P., and Moll, P.: Assimilation of simulated wind lidar data with a Kalman filter, Mon. Weather Rev., 121, 1803, doi:10.1175/1520-0493(1993)121<1803:AOSWLD>2.0.CO;2, 1993.

Geer, A. J., Lahoz, W. A., Bekki, S., Bormann, N., Errera, Q., Eskes, H. J., Fonteyn, D., Jackson, D. R., Juckes, M. N., Massart, S., Peuch, V.-H., Rharmili, S., and Segers, A.: The ASSET intercomparison of ozone analyses: method and first results, Atmos. Chem. Phys., 6, 5445–5474, doi:10.5194/acp-6-5445-2006, 2006.

Gilbert, J.-C. and Lemarechal, C.: Some numerical experiments with variable storage quasi-Newton algorithms, Math. Prog., 45, 407–435, 1989.

Jackson, J. D.: Classical Electrodynamics, 3rd edn., Wiley & Sons, 1998.

Krishnamurti, T. N., Bedi, H. S., Hardiker, V. M., and Ramaswamy, L.: An Introduction to Global Spectral Modeling, 2nd revised and enlarged edition, Springer, 2006.

Lahoz, W. A., Errera, Q., Viscardy, S., and Manney, G. L.: The 2009 stratospheric major warming described from synergistic use of BASCOE water vapour analyses and MLS observations, Atmos. Chem. Phys., 11, 4689–4703, doi:10.5194/acp-11-4689-2011, 2011.

Lin, S.-J. and Rood, R. B.: Multidimensional flux-form semi-Lagrangian transport schemes, Mon. Weather Rev., 124, 2046–2070, 1996.

Machenhauer, B.: The spectral method, in: Numerical Methods used in Atmospheric Models, edited by: Kasahara, E. A., No. 17 in GARP Publication Series, WMO, 1979.

Machenhauer, B. and Daley, R.: A baroclinic primitive equation model with a spectral representation in three dimensions, Tech. Rep. 4, Institute for Theoritical Meteorology, Copenhagen University, 1972.

Monin, A. S. and Yaglom, A. M.: Statistical Fluid Mechanics: Mechanics of Turbulence, vol. 1, The MIT Press, 1971.

Monin, A. S. and Yaglom, A. M.: Statistical Fluid Mechanics: Mechanics of Turbulence, vol. 2, The MIT Press, 1975.

Panchev, S.: Random Functions and Turbulence, International Series of Monograph in Natural Philosophy, Pergamon Press, 1971.

Parrish, D. F. and Derber, J. C.: The National Meteorological Center's spectral statistical-interpolation analysis system, Mon. Weather Rev., 120, 1747–1763, 1992.

Phillips, N. A.: The spatial statistics of random geostrophic modes and first-guess errors, Tellus A, 38, 314, doi:10.1111/j.1600-0870.1986.tb00418.x, 1986.

Purser, R. J., Wu, W.-S., Parrish, D. F., and Roberts, N. M.: Numerical Aspects of the Application of Recursive Filters to Variational Statistical Analysis. Part I: Spatially Homogeneous and Isotropic Gaussian Covariances, Mon. Weather Rev., 131, 1524–1535, 2003.

Satoh, M.: Atmospheric Circulation Dynamics and Circulation Models, ISBN 978-3-540-42638-7, Springer, 2004.

Schwinger, J. and Elbern, H.: Chemical state estimation for the middle atmosphere by four-dimensional variational data assimilation: a posteriori validation of error statistics in observation space, J. Geophys. Res., 115, D18307, doi:10.1029/2009JD013115, 2010.

Swarztrauber, P. N.: On the spectral approximation of discrete scalar and vector functions on the sphere, SIAM J. Numer. Anal., 16, 934–949, doi:10.1137/0716069, 1979.

Swarztrauber, P. N.: Discrete Fourier transforms and related topics, available at: http://www.cisl.ucar.edu/css/staff/pauls/tutorials/index.html, Lecture 1, 2003.

Swarztrauber, P. N.: Computing on the sphere: Part I, available at: http://www.cisl.ucar.edu/css/staff/pauls/tutorials/index.html, Lecture 2, 2003.

Swarztrauber, P. N. and Spotz, W. F.: Generalized discrete spherical harmonic transforms, J. Comp. Phys., 159, 213–230, 2000.

Talagrand, O.: Assimilation of observations, an introduction, J. Meteorol. Soc. Jpn, 277, 191–209, 1997.

Viscardy, S., Errera, Q., Christophe, Y., Chabrillat, S., and Lambert, J.-C.: Evaluation of ozone analyses from UARS MLS assimilation by BASCOE between 1992 and 1997, IEEE J. Sel. Top. Appl., 3, 190–202, doi:10.1109/JSTARS.2010.2040463, 2010.

Weaver, A. and Courtier, P.: Correlation modelling on the sphere using a generalized diffusion equation, Q. J. R. Meteorol. Soc., 127, 1815–1846, doi:10.1002/qj.49712757518, 2001.