



Heterogeneous Correlation Modeling Based on the Wavelet Diagonal Assumption and on the Diffusion Operator

OLIVIER PANNEKOUCKE

CNRM/GAME, Météo-France/CNRS, Toulouse, France

(Manuscript received 4 September 2008, in final form 6 March 2009)

ABSTRACT

This article discusses several models for background error correlation matrices using the wavelet diagonal assumption and the diffusion operator. The most general properties of filtering local correlation functions, with wavelet formulations, are recalled. Two spherical wavelet transforms based on Legendre spectrum and a gridpoint spherical wavelet transform are compared. The latter belongs to the class of second-generation wavelets. In addition, a nonseparable formulation that merges the wavelets and the diffusion operator model is formally proposed. This hybrid formulation is illustrated in a simple two-dimensional framework. These three formulations are tested in a toy experiment on the sphere: a large ensemble of perturbed forecasts is used to simulate a true background error ensemble, which gives a reference. This ensemble is then applied to compute the required parameters for each model. A randomization method is utilized in order to diagnose these different models. In particular, their ability to represent the geographical variations of the local correlation functions is studied by diagnosis of the local length scale. The results from these experiments show that the spectrally based wavelet formulation filters the geographical variations of the local correlation length scale but it is less able to represent the anisotropy. The gridpoint-based wavelet formulation is also able to represent some parts of the geographical variations but it appears that the correlation functions are dependent on the grid. Finally, the formulation based on the diffusion represents quite well the local length scale.

1. Introduction

Data assimilation aims to estimate the most likely numerical representation of a real system from known observations. This state is called the analysis and corresponds to the initial state of a new forecast. To estimate the analysis is quite a difficult problem for the atmosphere where observations are heterogeneous in time and space, and also because these observations are affected by noise. A prediction/correction method is often used so that the analysis is designed as a background corrected from the observations. In practice, the background generally corresponds to the most recent short-term forecast.

Data assimilation introduces two new unknown quantities: the observation error covariance matrix and the background error covariance matrix. The background error covariance matrix, denoted by \mathbf{B} , plays a key role

in the data assimilation scheme, since it contributes to the filtering of the observation error and also because it spreads the correction from these observations. The \mathbf{B} matrix quantifies statistical information evolving with time along the analysis/forecast cycles. It is known that horizontal and vertical correlations vary geographically (Lönnberg 1988). In particular, horizontal scales tend to be broader in the tropics than at high latitudes because of atmospheric dynamics (Ingleby 2001). There are also different horizontal correlations at different levels, and different vertical correlations for different horizontal wavenumbers. The latter property is the fingerprint of “nonseparability” of statistics (Rabier et al. 1998; Ingleby 2001, see his Fig. 2), which can be interpreted as reflecting a shallow-water and a deep-water behavior for the large-scale and small-scale structures, respectively. Furthermore, \mathbf{B} is flow dependent: correlation scales depend on the meteorological situation and on data density (Bouttier 1993, 1994). Two major difficulties occur with \mathbf{B} . The first one concerns the estimation of the matrix. Indeed, there are no measurement tools to determine \mathbf{B} , as it exists for the temperature or for the surface

Corresponding author address: Olivier Pannekoucke, Météo-France CNRM/GMAP, 42 Av. Coriolis, 31057 Toulouse, France.
E-mail: olivier.pannekoucke@meteo.fr

pressure. The second one is the huge size of \mathbf{B} (square of the dimension of the state vector) so that it is not possible to represent it explicitly in the computer memory.

Thus, \mathbf{B} must be modeled. A first common model is obtained by assuming that the correlations are homogeneous and isotropic (Courtier et al. 1998; Gaspari and Cohn 1999). Moreover, this correlation model is able to represent the nonseparability (Rabier et al. 1998). The resulting correlation matrix is represented by a diagonal in spectral space, and in this case, there are no geographical variations of the local correlation functions. Thus, statistics are estimated over long periods lasting several months. It is worth noting that some flow dependence can be introduced in this model [e.g., by using nonlinear balances in multivariate formulation (Fisher 2003), which generalize the multivariate formulation proposed by Derber and Bouttier (1999)].

Other models of correlation functions able to construct heterogeneous formulation exist; for example, the models using recursive filters (Purser et al. 2003a,b), or on the generalized diffusion equation (Weaver and Courtier 2001; Weaver and Ricci 2003). It should be noted that some alternate ways to construct model diffusion based on Gaussian correlation exists as the iterated Laplacian method (Derber and Rosati 1989; Egbert 1994). This formulation is feasible even over complex domains (e.g., in ocean modeling with coasts). But in their current formalism, neither the recursive filters nor the diffusion operator are able to represent the nonseparability of statistics. However, an improvement of the spectral formulation, designed with the spherical wavelets, has been proposed by Fisher and Andersson (2001; see also Fisher 2003). This correlation model is heterogeneous and nonseparable. An equivalent wavelet diagonal assumption has been considered by Deckmyn and Berre (2005) for limited-area models. It is worth noting that wavelets have also been introduced to their compression properties by Auger and Tangborn (2004) (see also Tangborn 2004). Since it is now able to represent local information, it is feasible to introduce some flow dependence, but the issue of \mathbf{B} estimation remains.

The technique of the assimilation ensemble elaborated with perturbed observations (Houtekamer et al. 1996; Fisher 2003) is a practical method to estimate the background error statistics. This approach is partly related to the work of Evensen (1994), who proposed an ensemble-based method to solve the Kalman filter equations. It offers an explicit time evolution of \mathbf{B} where the flow-dependent covariances are estimated from the ensemble. Ensemble assimilation methods have been used to estimate the stationary component of \mathbf{B} (Fisher 2003; Belo Pereira and Berre 2006; Pannekoucke and Massart 2008), or to compute the statistics of the day in

variational schemes (Pannekoucke et al. 2007; Berre et al. 2007; Pannekoucke 2008). In these last studies, spatial averaging was applied to remove the sampling noise due to the small size of the ensemble (Houtekamer and Mitchell 1998; Lorenc 2003). It should be also mentioned that some analogous filtering effects have been investigated by Buehner and Charron (2007).

The underlying question of the paper is how to model the background error covariance matrix with the wavelets and the diffusion operator and how to estimate these two models. Section 2 recalls the spectral-based wavelet formulation of Fisher (2003) and properties offered by the wavelet formulation (and more generally by the representation in a frame): the ability of the formulation to represent heterogeneous correlations, the filtering of sampling noise, and the smoothing of the geographical variations of correlation functions. A second-generation wavelet formulation, presented in section 3, corresponds to a wavelet formulation in grid space. The correlation model that relies on the diffusion equation is described in section 4, with some recent results on the estimation of the local diffusion tensor. This correlation model can represent heterogeneous correlations, but in its current formulation the model is not able to incorporate the nonseparability. A hybrid formulation is proposed to mix the diffusion operator and the wavelets in order to construct a nonseparable formulation. In section 5, the horizontal components of the three formulations are tested within a simulated experimental framework, using a toy model on the sphere. This model has been constructed to provide "true" local correlation functions that are unknown in real applications. In that case, only the feasibility and the ability of each formulation to represent the geographical variations of correlations are illustrated. The filtering of sampling noise is not addressed in this framework. Conclusions and perspectives are given in section 6.

2. Diagonal assumption in spectrally based wavelet representation

a. Definition of the wavelet transform

The basic ingredients of a wavelet transform on the real line are the translation and the dilatation of a locally supported function. The former analyzes a signal in physical space while the latter analyzes the signal in spectral space. A wavelet transformation splits a signal into different levels of detail. Thus, it can be considered as a set of bandpass filters and it is possible to formulate the wavelet transformation in term of convolution. On the sphere S^2 (also on the circle), it is difficult to define a good dilatation operator since a dilatation at a point along a great circle will lead to a contraction at the antipode (Holshneider 1990, 1996). Despite these difficulties,

different wavelet transformations on the sphere exist (Rosca 2005; Antoine et al. 2002).

To cope with the dilatation problem, Freeden and Windheuser (1996; see also Freeden and Schreiner 1998) have proposed a wavelet transformation based on the Legendre spectrum. It consists of defining an appropriate set of bandpass filters ψ^j , given by their Legendre spectrum, and then convoluting it with the spherical function to decompose. This tool was introduced in data assimilation for covariance modeling by Fisher (see, e.g., Fisher and Andersson 2001), who made this choice because such a spectrally based wavelet (SBW) formulation is invariant with spherical rotation. This last point is a strong constraint for the isotropic representation of dynamics (e.g., with the quasi-spectral model and triangular truncation), but it is less true in a gridpoint model as those encountered in ocean modeling, with a heterogeneous grid and resolution.

The direct and the inverse SBW transforms are defined as follows. For an arbitrary field ε , the wavelet coefficients at scale j , denoted by ε_j , are computed as the convolution by the bandpass filter ψ^j : $\varepsilon_w^j = \psi^j \otimes \varepsilon$, where \otimes denotes the convolution product. The field ε can be obtained again as the sum of the details ε_w^j with $\varepsilon = \sum_j \psi^j \otimes \varepsilon_w^j$. In general, spherical convolution is defined according to the group of rotations of \mathbb{R}^3 , denoted $SO(3)$. This group is also the group of 3×3 orthogonal matrices with determinant one. But in this particular case, thanks to the addition theorem for spherical harmonics with appropriate normalization (Müller 1966), the convolution reduces to a spectrum product as $(\psi^j \otimes \varepsilon)_n^m = (1/\sqrt{2n+1})\psi_n^j \varepsilon_n^m$, where ψ_n^j is the Legendre spectrum of ψ^j and ε_n^m is the spherical harmonic spectrum of ε . This is similar to the convolution product in the Fourier representation in a one-dimensional circle.

The formulation of Fisher (2003) is designed with band-limited ψ^j functions so that their spectrum representation is as follows. For a set of arbitrary chosen integers, $(N_j)_{j \in [0, J]}$, with $N_j < N_{j+1}$, the spectral coefficients of function ψ^j are given by, for $j \neq 0$ (n being the total wavenumber in the spherical case):

$$\frac{1}{\sqrt{2n+1}} \psi_n^j = \begin{cases} \sqrt{\frac{n - N_{j-1}}{N_j - N_{j-1}}} & \text{for } N_{j-1} \leq n < N_j, \\ \sqrt{\frac{N_{j+1} - n}{N_{j+1} - N_j}} & \text{for } N_j \leq n < N_{j+1}, \\ 0 & \text{otherwise.} \end{cases}$$

For $j = 0$, the definition is the same, except that the range $N_{j-1} \leq n < N_j$ is replaced by $0 \leq n < N_0$, for which $(2n + 1)^{-1/2} \psi_n^j = 1$.

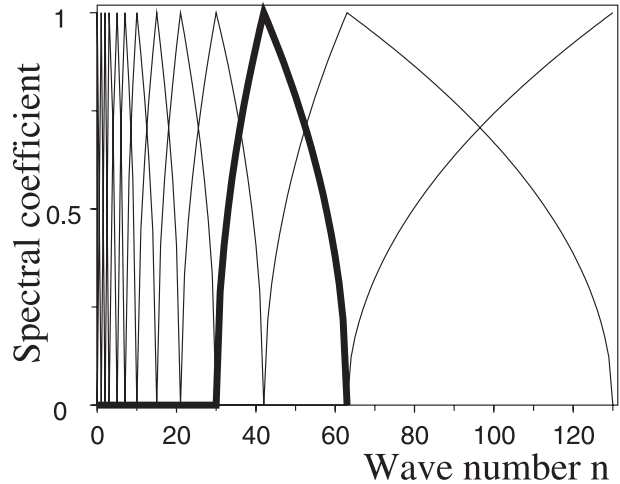


FIG. 1. Spectral coefficients $(2n + 1)^{-1/2} \psi_n^j$ of the wavelet bandpass filters ψ^j . The spectrum associated to the particular ψ^{10} function is shown in bold.

Note that these SBWs are not orthogonal and do not form a basis, but rather a frame (Fisher 2004; Daubechies 1992; Mallat 1999). In finite expansions, if \mathbf{W} represents the linear operator associated with this wavelet transformation, then \mathbf{W} is not invertible (in particular, \mathbf{W} is not square but rectangular). Of course, \mathbf{W} has a left inverse (the Moore–Penrose pseudoinverse) \mathbf{W}^{-1} so that $\mathbf{W}^{-1}\mathbf{W} = \mathbf{I}$ but $\mathbf{W}\mathbf{W}^{-1} \neq \mathbf{I}$, where \mathbf{I} represents the identity operator. Moreover, this particular kind of wavelets forms a tight frame, implying that $\mathbf{W}^{-1} = \mathbf{W}^T$.

Figure 1 illustrates the bandpass filters for the particular set $\{N_j\} = \{0, 1, 2, 3, 4, 5, 7, 10, 15, 21, 30, 63, 130\}$. As ψ^j are band limited, the wavelet coefficients ε_w^j can be stored on a coarse grid in physical space, associated with an adapted triangular truncation. By convention, the smaller the truncation, the smaller the index j . Note that the truncation for the two last scales is the truncation of the full-resolution grid. Thus, the number of wavelet coefficients is at least 2 times greater than the degree of freedom of the gridpoint representation (which is why \mathbf{W} is rectangular). The wavelet function associated with a given position x on the sphere is $\psi_x^j(x') = (\psi^j \otimes \delta_x)(x')$, where δ_x is the Dirac distribution at point x . Wavelet functions are such that the larger the spectral band, the more localized in physical space it is, constrained by Heisenberg’s uncertainty principle. Figure 2 represents an example of SBW expansion of a spherical field (Fig. 2a). There is one resolution for each scale, with a low resolution for the large scales (Fig. 2b) and a higher resolution for the small scales (Fig. 2d).

The choice of $\{N_j\}$ is a degree of freedom for the formulation and can be optimized in order to match geographical variations of local correlation from a given set

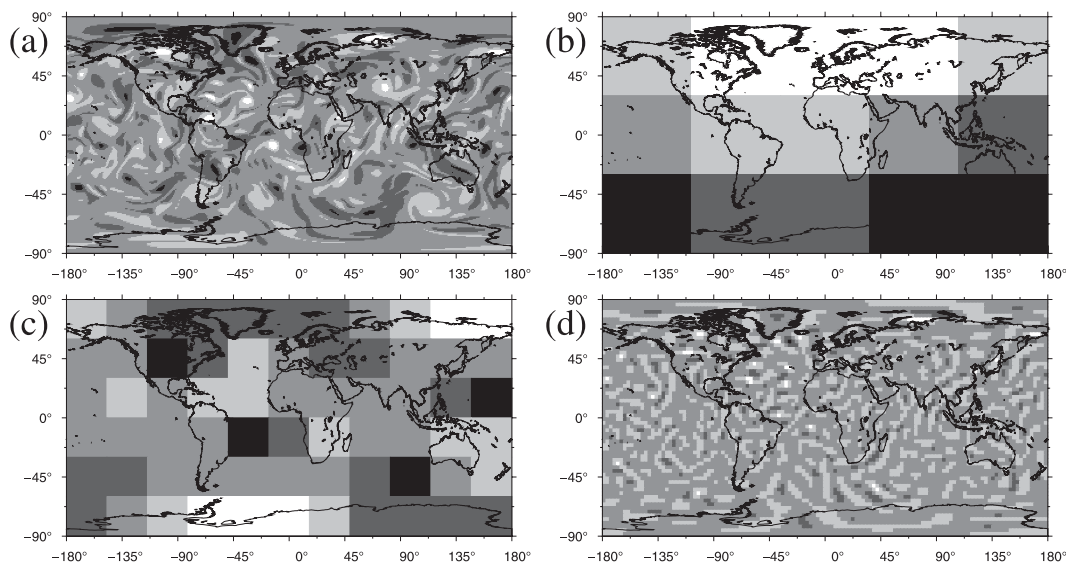


FIG. 2. (a) Illustration of a wavelet expansion of a field. Some wavelet coefficients are represented for only few scales: (b) the large scale $j = 1$, (c) the medium scale $j = 4$, and (d) the small scale $j = 10$. In (b)–(d), each gray bloc is a wavelet coefficient.

of statistics. For instance, it is possible to specify $\{N_j\}$ so that the modeled correlation length scale corresponds to the diagnosed length scale under least squares criterion. This discrete optimization problem can be solved by using metaheuristic algorithms such as simulated annealing. Note that for orthogonal wavelets, some algorithms are available for selecting the best basis among a dictionary of basis (in particular dictionary of wavelet packets; see Coifman and Wickerhauser 1992). In this framework, it is also possible to select the best approximation of the Karhunen–Loève basis (an orthonormal basis for which a covariance matrix is represented by a diagonal) (Mallat et al. 1998; appendix E in Pannekoucke 2008).

However, the choice of $\{N_j\}$ constraints, in correlation modeling, the resolution of the variation of vertical correlations with horizontal scale (spectrally defined): the larger bands $[N_j, N_{j+1}]$ are lower in the spectral resolution of the vertical–horizontal nonseparability (Fisher and Andersson 2001). Because of some atmospheric balances, (e.g., shallow-water behavior in the large scales and deep-water behavior in the small scales with a fast transition in the low wavenumbers) it appears better to construct a fine (coarse) representation of low (high) wavenumbers (i.e., the spectral support of the bandpass filters is larger in the small scales than in the large scales).

b. Formalism of diagonal assumption in wavelet space

Similarly to the diagonal formulation in spectral space (Courtier et al. 1998), the wavelet equivalent is that the background correlation tensor is represented by a di-

agonal in wavelet space. Thus, \mathbf{B} is expanded as the product of $\mathbf{B} = \Sigma_g \mathbf{C} \Sigma_g^T$ where Σ_g^2 is the diagonal matrix of \mathbf{B} and \mathbf{C} is the correlation matrix. With the previous notation, the wavelet formulation of the matrix \mathbf{B} is specified following Fisher (2003) and Deckmyn and Berre (2005) as $\mathbf{B}_{dw} = \Sigma_g \mathbf{C}_{dw} \Sigma_g^T$ with

$$\mathbf{C}_{dw} = \Lambda \Sigma_s \mathbf{W}^{-1} \mathbf{D}_w \mathbf{W}^{-T} \Sigma_s^T \Lambda^T, \quad (1)$$

where Σ_s is a diagonal normalization by the spectral standard deviation, $\mathbf{D}_w^{1/2}$ is a diagonal of wavelet standard deviations, and Λ is a diagonal normalization by standard deviations of $\Sigma_s \mathbf{W}^{-1} \mathbf{D}_w \mathbf{W}^{-T} \Sigma_s^T$. Thereafter, this normalization is not recalled, but is applied. Note that a nonseparable formulation can be constructed following Fisher and Andersson (2001; see also Fisher 2003). In that correlation model, the matrix \mathbf{D}_w is a block diagonal matrix whose blocks $\mathbf{D}_{w,\alpha}$ are the vertical covariances associated with wavelet index α , where $\alpha = (j, \mathbf{x})$ is a coupling of scale and position. Similarly to the spectral formulation, the dependence on frequency is ensured by the dependence on the scale j . But unlike the spectral formulation, the dependence on position \mathbf{x} involves a geographical variation of the vertical correlation by scale.

One of the consequences of the diagonal assumption in a frame is that the diagonal matrix \mathbf{C}_d is related to the initial matrix \mathbf{C} following a relation of the following form:

$$\mathbf{C}_d(x', x) = \sum_{z, z'} \mathbf{c}(z', z) \Phi_{(x', x)}(z', z),$$

where the weight matrix $\Phi_{(x',x)}$ is function of the frame (Pannekoucke et al. 2007). This means that C_d is a spatial average of C . In the spectral basis, the weight only depends on the distance between x' and x . Thus, the diagonal assumption in spectral space C_{ds} can be seen as a global spatial average of C , so that the resulting tensor is homogeneous (and isotropic according to Gaspari and Cohn 1999). In the wavelet representation, the weight $\Phi_{(x',x)}$ is localized around the position (x',x) , because the wavelets are localized in both space and spectrum. Here C_{dw} can be seen as a local spatial average of C . Then, this formulation is able to represent geographical variations of the local correlation functions. However, as the wavelets ψ_x^j are radial (i.e., isotropic) functions and thus not polarized (i.e., anisotropic or directional; see Antoine et al. 2002), the local average at each scale tends to damped anisotropy.

The result of this averaging is that the geographical variations of the local correlation functions, modeled with the wavelet diagonal assumption, are smooth, associated with a filtering of sampling noise (Pannekoucke et al. 2007). Figure 3 mimics the estimation of wavelet statistics at different scales and for two points A and B . For the large scales, the spatial average is done over a large area (horizontally hashed area for A and vertically hashed area for B). This eliminates small-scale contribution and thus damps a part of the sampling noise. At the opposite extreme, for small scales, the averaging is done over a local area near the point and they are more sensitive to sampling noise. The statistics computed at two different points vary slowly for the large scale: areas around A and B overlap. Then the statistics vary more rapidly in the medium scales: areas around A and B overlap slightly. Finally, they are independent in the small scales: areas around A and B are disjointed.

To illustrate the diagonal assumption in wavelet space and to give another point of view of why sampling noise is eliminated in this formulation, a one-dimensional example is now introduced.

c. Example in a one-dimensional framework with varying length scale

Following Pannekoucke et al. (2007), a simple 1D analytical heterogeneous correlation matrix on the circle is obtained from a change of variable. This procedure is well known as it offers a way to construct flow-dependent correlation (e.g., with a semigeostrophic change of variable; Desroziers 1997). In the present construction, a 1D Schmidt transformation of parameter c (or c stretching) is used (Courtier and Geleyn 1988). It is defined, on the circle of radius a , by

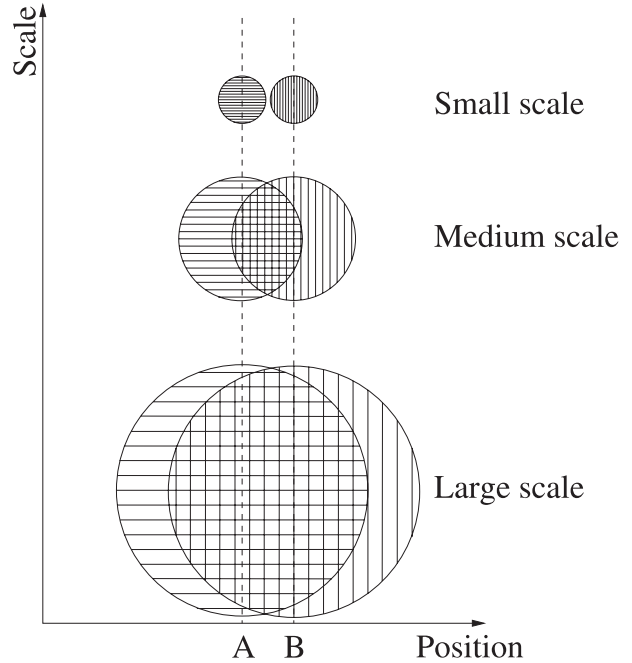


FIG. 3. Sketch of the local spatial average at different scales with the diagonal wavelet assumption (see text for details).

$$u(x) = a \left\{ \pi - 2 \arctan \left[c^{-1} \tan \left(\frac{\pi}{2} - \frac{1}{2a} x \right) \right] \right\},$$

where x is the geographical position, with x/a varying from 0° to 360° . The change of variable is applied to a homogeneous Gaussian correlation tensor:

$$C_h(x, y) = \exp \left[\frac{-(x - y)^2}{2L_h^2} \right],$$

where $L_h = 250$ km denotes the length scale. The resulting correlation tensor is thus

$$C(x, y) = C_h[u^{-1}(x), u^{-1}(y)].$$

For numerical applications, $a = 6400$ km (the radius value of the earth), $c = 2.4$, and $L_h = 250$ km. The circle is regularly discretized at truncation $T = 130$ and the discretization of the correlation tensor C is denoted by C . As shown in Fig. 4a, the correlation functions of C are more (less) spread near 0° (360°) than the original Gaussian correlations (not shown).

The wavelet transform considered on the circle is equivalent to the wavelet presented in section 2a but replaces the Legendre spectrum with a cosine spectrum. This class of circular wavelets also forms a tight frame ($W^T = W^{-1}$).

Now, applying the diagonal assumption in wavelet space (spectral normalization is not considered here)

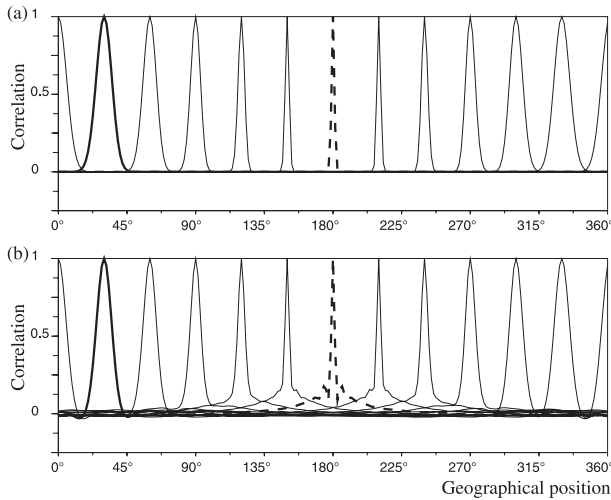


FIG. 4. Representation of some local correlation functions of (a) the heterogeneous correlation matrix \mathbf{C} and (b) its modeled version based on the diagonal assumption \mathbf{C}_{dw} . In both (a) and (b), each solid line curve represents a local correlation function. The bold solid line and the bold dashed line are two particular correlation functions.

consists, in a first step, in computing the diagonal operator \mathbf{D}_w . In this simple one-dimensional framework, where the matrices can be fully represented in computer memory, \mathbf{D}_w is explicitly computed as $\mathbf{D}_w = \text{diag}(\mathbf{W}\mathbf{C}\mathbf{W}^T)$. Then the computation of \mathbf{C}_{dw} is achieved according to $\mathbf{C}_{dw} = \mathbf{W}^{-1}\mathbf{D}_w\mathbf{W}^{-T}$. Some of the correlation functions of \mathbf{C}_{dw} are shown in Fig. 4b.

Geographical variations of correlation functions are well represented. Modeled correlation functions are largely spread near 0° —see for instance the bold solid line correlation function near 30° in Fig. 4b, which is quite similar to the original in Fig. 4a—and they are well localized near 180° as shown by the bold dashed line correlation function. Some defects are still present, visible especially on the bold dashed line correlation in Fig. 4b. This can be partly corrected by adapting the set of integers $\{N_j\}$.

The diagonal of \mathbf{D}_w is represented in Fig. 5b. In this panel, each box represents a wavelet variance with a gray level depending on the variance magnitude. A wavelet variance is associated to a correlation around a certain position and frequency. This position can be viewed as the middle abscissa of a box while the frequency is related to the scale j . In accordance with the definition of \mathbf{C} and Fig. 4a, large scales are excited in the vicinity of 0° while small scales are excited near 180° .

The local correlation function related to a position x is computed as $\mathbf{f}_x = \mathbf{C}_{dw}\delta_x$, with δ_x the Dirac distribution. For the present tight frame, $\mathbf{C}_{dw} = \mathbf{W}^{-1}\mathbf{D}_w\mathbf{W}$ and \mathbf{f}_x are obtained using a three-step process. The first step is the

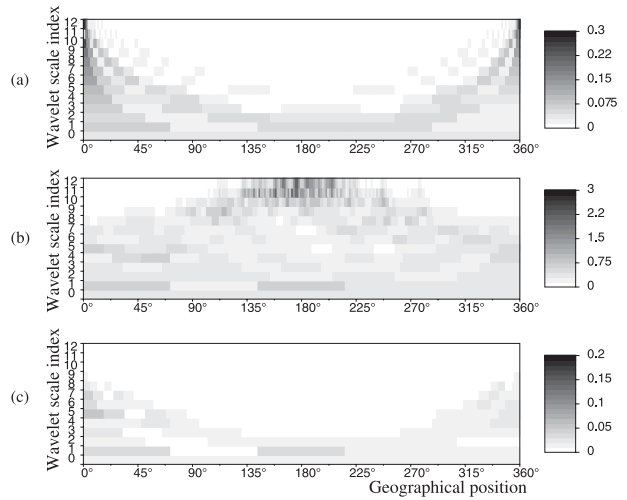


FIG. 5. Illustration of the construction of a correlation function in the wavelet transformation: (a) wavelet transform $\mathbf{W}\delta_{0^\circ}$ of the Dirac function δ_{0° , (b) wavelet standard deviation that corresponds to the diagonal of the diagonal matrix \mathbf{D}_w , and (c) the pre-correlation function $\mathbf{D}_w\mathbf{W}\delta_{0^\circ}$ that can be considered as the wavelet coefficients of \mathbf{f}_{0° , the correlation function at point 0° , modeled with the diagonal assumption. In (a)–(c), white (black) stands for low (large) values.

wavelet transformation $\mathbf{W}\delta_x$ of the Dirac distribution δ_x . The wavelet coefficients of $\mathbf{W}\delta_{0^\circ}$ are represented in Fig. 5a. It constitutes a particular pattern called the “cone of influence”; wavelet coefficients are negligible outside a conelike area. The second step is the element wise product of the wavelet coefficients $\mathbf{W}\delta_x$ with the wavelet variances \mathbf{D}_w . Here \mathbf{D}_w is represented in Fig. 5b and $\mathbf{D}_w\mathbf{W}\delta_{0^\circ}$ is represented in Fig. 5c. It can be considered as the wavelet coefficients of \mathbf{f}_{0° (this is not rigorously the case as $\mathbf{W}\mathbf{W}^{-1} \neq \mathbf{I}$). The last step is the return to physical space with \mathbf{W}^{-1} .

The former two steps can help us to understand (through a spatial average interpretation) why diagonal wavelet assumption is able to filter sampling noise. When statistics are estimated from an ensemble (e.g., an ensemble of perturbed forecasts) the finite and generally small size of the ensemble leads to sampling noise (Houtekamer and Mitchell 1998; Lorenc 2003). But when applying the diagonal assumption in wavelet space, and following the previously described steps, it appears that only the wavelet variances located in the cone of influence are taken into account. This results in information beyond a certain distance not contributing to the correlation function model. This can be interpreted as a filtering of the raw local correlation function.

An illustration of this filtering is given in Fig. 6a, which shows an example of a correlation function, relative to 0° estimated with 10 members (solid line) compared with the true correlation function (dashed line). In this case,

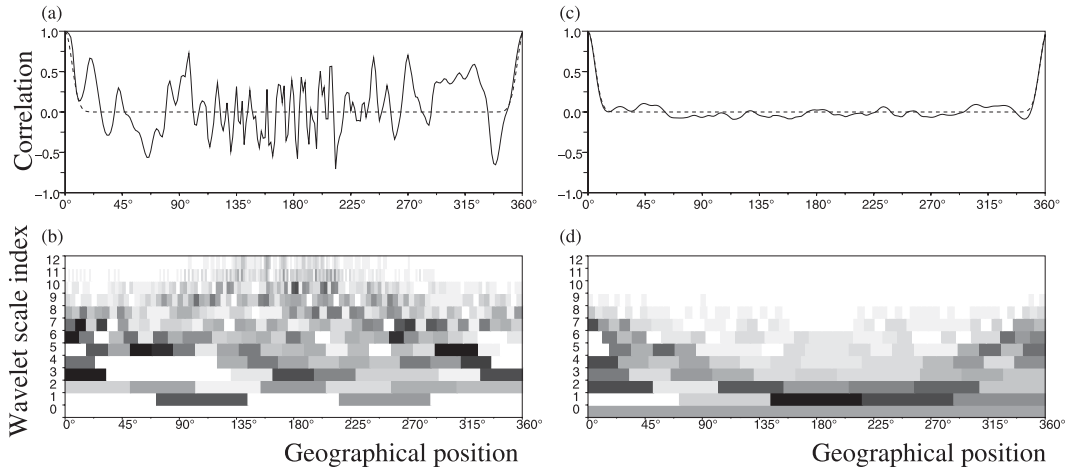


FIG. 6. Effects of sampling noise on correlation estimation view in wavelet representation (see the text for details).

each background error has been randomly generated as $\epsilon_b = \mathbf{C}^{1/2}\zeta$, where ζ is a sample of a Gaussian random vector with zero mean and the identity \mathbf{I} as a covariance matrix (Fisher and Courtier 1995). Thus, with the expectation function \mathbb{E} , the covariance matrix of the generated ϵ_b is $\mathbb{E}(\epsilon_b \epsilon_b^T) = \mathbf{C}^{1/2} \mathbb{E}(\zeta \zeta^T) \mathbf{C}^{1/2} = \mathbf{C}$. The covariance matrix estimated from this ensemble of simulated error ϵ_b^k is $\mathbf{C}_e = (1/N) \sum_k \epsilon_b^k \epsilon_b^{kT}$, with $N = 10$. The correlation matrix estimated is obtained after a normalization of \mathbf{C}_e by its standard deviations. The sampling noise is particularly visible from 45° to 315° where the true correlation is zero while its estimation oscillates with a magnitude of 0.5. Note that the frequency of these oscillations is varying along in space, with a low-frequency oscillation near 45° and a high-frequency oscillation near 180° . This is directly related to the local correlation itself: local correlation functions are more spread near 45° than near 180° . Figure 6b shows the absolute value of the wavelet coefficients of the estimated correlation function. It accurately represents this variation of local frequency with no energy in the small scales near 45° (coefficients at scale $j = 12$ are negligible) and some energy in the small scales near 180° . Now applying the diagonal assumption in wavelet space will damp all coefficients outside the cone of influence at point 0° , as shown in Fig. 6d, which corresponds to the absolute value of the wavelet coefficient of the local correlation function modeled with wavelet assumption. This last correlation function is represented in Fig. 6c (solid line) and its estimation is done from the same 10 members used in Fig. 6a. It is clear that the sampling noise has been reduced.

As seen previously, the wavelet formulation offers interesting properties: it is able to represent horizontal heterogeneity (see Fig. 4), it can filter a part of the

sampling noise, and it can represent the nonseparability. However, the size of the wavelet coefficient space is larger than in the initial space, this spectrally based wavelet cannot be used in complex geometry and because the wavelet functions are isotropic, it is not easy to get a good representation of anisotropy. To study some of these points, another wavelet formulation is now introduced.

3. Diagonal assumption in second-generation wavelets

The previous SBW formulation is well adapted to atmospheric models, using a triangular spectral representation. However, it could be interesting to extend this approach to other fields such as ocean models where the invariance by spherical rotation is no longer valid. As a preliminary example, a gridpoint-based wavelet transformation is now presented. It corresponds to another class of wavelet transformation called second-generation wavelets (SGWs; Sweldens 1995, 1998).

The basis of this construction is first a hierarchical mesh on the sphere, and second a scheme of prediction/correction where the wavelet coefficients appear as the correction.

a. Subdividing the sphere and splitting

A hierarchical subdivision of the sphere is first described. It relies on recursive partitioning of the sphere into geodesic spherical triangles. Starting from an icosahedron (a convex polyhedron composed of 20 triangular faces with 12 vertices) at each iteration, 4 triangles are generated from each triangle of the previous subdivision. Figure 7 represents the initial grid (Fig. 7a) and the first three subdivisions (Figs. 7b–d).

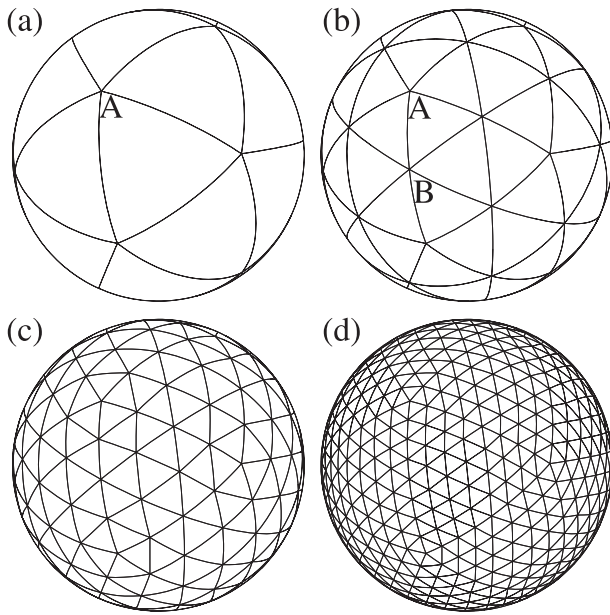


FIG. 7. Representation of grid subdivisions. (a) The initial grid, formed from the projection on a sphere of the icosahedron, is subdivided along a recursive process after (b) one, (c) two, and (d) three iterations.

The degree of a point is the number of connection it has in the mesh. The point *A* in Figs. 7a,b is of degree 5, while the point *B* in Fig. 7b is of degree 6. Thus, the grid is not homogeneous. It is worth noting that this kind of mesh on the sphere has been already used by Sadourny et al. (1968) and also by Williamson (1968) for numerical modeling. Recently the use of this kind of mesh has been extended, for example, with an adaptive wavelet collocation method (Mehra and Kevlahan 2008).

The set of all vertices p_k^j , after j subdivisions is denoted $S^j = \{p_k^j\}_{k \in K^j}$, where K^j is an index set. Here S^0 corresponds to the original icosahedron. Since $S^j \subset S^{j+1}$, we so let $K^j \subset K^{j+1}$. The point p_k^{j+1} for $k \in K^j$ is defined by $p_k^{j+1} = p_k^j$. The cardinal of K^j is $\text{Card}(K^j) = 10 \times 4^j + 2$. The index set of the vertices added when going from level j to level $j + 1$ is denoted $M^j = K^{j+1} - K^j$.

Thus at a given resolution $j + 1$ (high resolution) the set K^{j+1} can be split into two sets K^j and M^j so that $K^{j+1} = K^j + M^j$. If a field ε is represented by its discretization on a grid S^{j+1} as $\varepsilon^{j+1} = (\varepsilon_k^{j+1})_{k \in K^{j+1}}$, then its low-resolution version is defined as $\varepsilon^j = (\varepsilon_k^j)_{k \in K^j}$ so that

$$\forall k \in K^j, \quad \varepsilon_k^j = \varepsilon_k^{j+1}. \tag{2}$$

b. Wavelet coefficients as corrections to a prediction

Now, the idea consists in predicting the higher-resolution field ε^{j+1} from the low-resolution field ε^j . Following

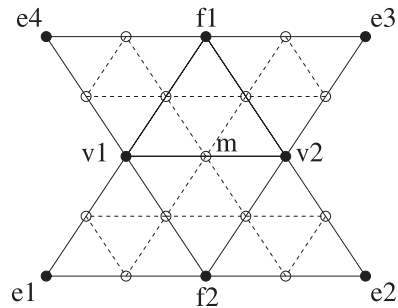


FIG. 8. The members of the index sets used in the prediction \mathcal{P} based on the butterfly scheme, with m in M^j and with $\mathcal{K}_m = \{v_1, v_2, f_1, f_2, e_1, e_2, e_3, e_4\}$ included in K^j . Circles and dashed lines represent the next finer level of resolution $j + 1$, while black points and solid lines represent the grid at level resolution j .

Eq. (2), only the value of points in M^j remains to be found. Schröder and Sweldens (1995a,b) have used a butterfly scheme, a technique inherited from computer-aided geometric design (Dyn et al. 1990), where what is wanted is to construct a smooth C^1 surface out of a control polyhedron. For each m in M^j , the butterfly basis uses a stencil of eight neighboring points $\mathcal{K}_m \subset K^j$ as illustrated in Fig. 8. The prediction is $\mathcal{P}(\varepsilon^j) = \sum_{k \in \mathcal{K}_m} w_{k,m}^j \varepsilon_k^j$, with the particular weight (independent of j): $w_{v_1} = w_{v_2} = 1/2, w_{f_1} = w_{f_2} = 1/8$ and $w_{e_1} = w_{e_2} = w_{e_3} = w_{e_4} = -1/16$. The wavelet coefficients $\varepsilon_w^j = (\varepsilon_{wk}^j)_{m \in M^j}$ at scale j are thus defined as the misfit between the true values of ε^{j+1} and the predicted ones:

$$\varepsilon_{wk}^j = \varepsilon_m^{j+1} - \sum_{k \in \mathcal{K}_m} w_{k,m}^j \varepsilon_k^j. \tag{3}$$

This expansion could be sufficient, but an additional step called lifting is added, after the computation of ε_w^j , in order to ensure that wavelet functions are of zero mean, which is realized as follows: for a given j and a given k in K^j , the scaling function ϕ_k^j is defined as the inverse of the wavelet coefficients ϕ_w so that $\phi_w^j = \delta_{j,j'} \delta_{k,k'}$, where $\delta_{i,j}$ stands for the Kronecker symbol. A wavelet function ψ_m^j is defined from these ϕ_k^j functions by

$$\psi_m^j = \phi_m^{j+1} - \sum_{k \in \mathcal{K}_m = \{v_1, v_2\}} s_{k,m}^j \phi_k^j, \tag{4}$$

where $s_{k,m}^j = I_m^{j+1}/2I_k^j$ with $I_k^j = \int_{S^2} \phi_k^j d\omega$ the integral over the sphere S^2 with its classical measure $d\omega$ (see the appendix for the computation of $I_{j,k}$). Finally, the wavelet coefficient is defined as

$$\forall m \in M^j: \forall k \in \mathcal{K}_m = \{v_1, v_2\}, \quad \varepsilon_k^j = \varepsilon_k^j + s_{k,m}^j \varepsilon_{wk}^j. \tag{5}$$

Thus, the direct wavelet transformation consists in computing recursively the wavelet coefficients ε_w^j . In practice this recursion is initialized by considering that the discrete signal samples ε belong to S^J , and is stopped when $j = 0$. Each level is divided into two steps, with a first step corresponding to the subsampling, described by Eqs. (2) and (3), and a second step, corresponding to the lifting, described by Eq. (5). Note that these wavelets are not radial (they do not have an axis of symmetry). This is related both to the butterfly scheme and to the directional definition of the wavelet functions ψ^j in Eq. (4). However, this anisotropy is not very strong: the wavelet functions of large scales are quasi isotropic (not shown here).

The inverse wavelet transform is also computed under a recursive form but in the reverse, from scale $j = 0$ to $j = J$. Each level is divided into two steps with a first step described by the opposite of Eq. (5):

$$\forall m \in M^j : \forall k \in \mathcal{K}_m = \{v_1, v_2\}, \quad \varepsilon_k^j = \varepsilon_k^{j+1} - s_{k,m}^j \varepsilon_{wk}^j, \quad (6)$$

and then the upsampling described by the opposite of Eq. (3):

$$\varepsilon_m^{j+1} = \varepsilon_{wk}^j + \sum_{k \in \mathcal{K}_m} w_{k,m}^j \varepsilon_k^j, \quad (7)$$

and by the reverse of Eq. (2):

$$\forall k \in K^j, \quad \varepsilon_k^{j+1} = \varepsilon_k^j. \quad (8)$$

The linear operator associated to the direct (the inverse) wavelet transformation is denoted by $\mathbf{W}(\mathbf{W}^{-1})$.

The mathematical background of this wavelet decomposition is given next. In terms of functional analysis, the space V^{j+1} , engendered by scaling functions ϕ^{j+1} (which forms a biorthogonal basis, or a Riesz basis, of V^{j+1}), is included in the space of square-integrable functions on the sphere $\mathcal{L}^2(S)$, while the wavelet functions ψ^j engender a subspace $W^j \subset V^{j+1}$ so that $V^{j+1} = V^j \oplus W^j$ (ψ^j forms a Riesz basis of W^j). Thus, V^J is decomposed along the recursive process into $V^J = V^0 \oplus W^0 \oplus W^1 \oplus \dots \oplus W^{J-1}$. Moreover, the closure of $\cup_j V^j$ is $\mathcal{L}^2(S)$. This kind of construction is called a multi-resolution analysis and is important for orthogonal wavelets. In this case, Riesz bases are replaced by orthogonal bases, with additional constraints (Daubechies 1992; Mallat 1999).

c. Formalism of diagonal assumption in SGW space

The correlation formulation retained in this paper is similar to the one presented in section 2b. In this case, the wavelet transformation is replaced, in Eq. (1) by the

SGW transformation (with the same notation \mathbf{W}). However, the normalization with the spectral standard deviation is not used.

What is attractive about the SGW is that the dimension of the wavelet coefficient space is equal to the dimension of the initial state. The wavelets are compactly supported. Then, as the wavelets are a little directional, they should be more adapted to represent a part of the anisotropy of the local correlation function (at least in the small scales where the wavelet functions are directional, but not in the large scales where the wavelet functions are quasi isotropic). But this might depend on the direction of the local correlation anisotropy compared with the direction of the wavelet. Furthermore, it is possible to construct second-generation wavelets that take into account boundaries. Note that algorithms to compute SGW transform are fast: the computational complexity is $O(N)$, where N is the number of grid points (Schröder and Sweldens 1995a).

An example of this SGW formulation, in a spherical toy framework, is presented in section 5.

4. Covariance modeling based on diffusion operator

a. Formalism of the correlation modeling with a diffusion operator

Another way to construct a heterogeneous correlation matrix has been introduced by Weaver and Courtier (2001) and is based on the generalized diffusion equation. It is noticeable that on a plane, a particular solution of the diffusion equation is

$$\partial_t \eta = \nabla \cdot (\boldsymbol{\nu} \nabla \eta), \quad (9)$$

with the constant diffusion tensor $\boldsymbol{\nu}$ (corresponding to the homogeneous diffusion equation) is the Gaussian function:

$$G^{\nu,t}(\mathbf{x}, \mathbf{x}') = \frac{1}{2\pi|\boldsymbol{\Gamma}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x} - \mathbf{x}')^T \boldsymbol{\Gamma}^{-1}(\mathbf{x} - \mathbf{x}')\right],$$

with

$$\boldsymbol{\Gamma} = 2t\boldsymbol{\nu}, \quad (10)$$

where t is the time of integration and $|\boldsymbol{\Gamma}|$ is the determinant of $\boldsymbol{\Gamma}$. The expansion of $\boldsymbol{\Gamma}^{-1}$ in the 2D case according to

$$\boldsymbol{\Gamma}^{-1} = \begin{pmatrix} \frac{1}{L_x^2} & \frac{1}{L_{xy}} \\ \frac{1}{L_{xy}} & \frac{1}{L_y^2} \end{pmatrix}$$

gives rise to the length scale $L_x(L_y)$ along axis $x(y)$. This length scale corresponds to the differential length scale (Daley 1991, p. 110), and it can be approximated under assumption of correlation behavior in the vicinity of small separation distance (Pannekoucke et al. 2008).

The idea is then to construct Gaussian-like correlation functions as the result of the diffusion of a Dirac distribution. By using this approach, it is feasible to generate heterogeneous correlation functions when the local diffusion tensor $\nu(\mathbf{x})$ varies with the position $\mathbf{x} = (x, y)$. The correlation tensor is then modeled as

$$\mathbf{C} = \mathbf{\Lambda} \mathbf{L}^{1/2} \mathbf{E}^{-1} \mathbf{L}^{T/2} \mathbf{\Lambda}^T,$$

where $\mathbf{L}^{1/2}$ is a half-time integration (or the propagator) from initial time 0 and final time $t/2$, \mathbf{E} is a local metric tensor, and (similarly to the wavelet formulation) $\mathbf{\Lambda}$ is a normalization so that \mathbf{C} is a correlation matrix. Note that in the homogeneous sphere of triangular truncation T , the inverse metric \mathbf{E}^{-1} is expressed by $\mathbf{E}^{-1} = (T + 1)^2 / 4\pi a^2 \mathbf{I}$, where a is the radius of the sphere and \mathbf{I} denotes the identity operator.

For practical applications, a numerical time integration of the diffusion Eq. (9) is achieved in a quasi-spectral approach (triangular truncation T), with an Euler time scheme of time step defined by $dt \leq (dx)^2 / (\text{CFL} \nu_{\max})$, where dx is the regular step associated to truncation T and the Courant–Friedrichs–Lewy condition of stability is set to $\text{CFL} = 4$ (Weaver and Courtier 2001). Here ν_{\max} corresponds to the maximum of the spectrum of all the local diffusion tensor. The convention is that local diffusion tensors are defined for time $t = 1$ in Eq. (10). Thus, $\mathbf{L}^{1/2}$ corresponds to a time integration over the interval $[0, 1/2]$.

b. Estimation of the local diffusion tensor

The question of how to estimate this local diffusion tensor remains. A first possibility is to estimate the local tensor from the dynamical aspects of the geophysical flow considered (A. Weaver 2007, personal correspondence). The underlying assumption is that the spread of the local correlation function is only related to the dynamics. But this is not entirely satisfactory as it is known that background error correlation functions are also influenced by the data network (Bouttier 1994).

Another possibility is to estimate the local diffusion tensor from ensemble statistics (Pannekoucke and Massart 2008). This estimation relies on the estimation of the local anisotropic tensor, under the assumption of a local homogeneous diffusion. This means that the local correlation function $\rho(\mathbf{x}, \cdot)$ is considered as being locally Gaussian in the limit of small separation distance:

$$\rho(\mathbf{x}, \mathbf{x} + \delta\mathbf{x}) \equiv \exp\left(-\frac{1}{2} \delta\mathbf{x}^T \mathbf{\Gamma}^{-1} \delta\mathbf{x}\right),$$

where $\delta\mathbf{x} = (\delta x, \delta y)$ is a small displacement. Here $\mathbf{\Gamma}^{-1}$ can be numerically estimated as the least squares solution from an ensemble of $K \geq 3$ position $\delta\mathbf{x}_k = (\delta x_k, \delta y_k)$ associated with numerical correlation values $\rho_k = \text{cor}(\mathbf{x}, \mathbf{x} + \delta\mathbf{x}_k) > 0$. Thus, local length scale and anisotropy play an important role in this approach (contrary to the wavelet diagonal assumption where the length scale is not necessary to estimate the formulation). Note that the question of how to estimate parameters of the generalized diffusion equation is not yet answered.

c. Nonseparable hybrid formulation based on the diffusion equation and the wavelets

A three-dimensional diffusion operator is constructed similarly (Weaver and Courtier 2001). In this case the local diffusion tensor ν is a 3×3 matrix. In the formulation, the diffusion along the vertical is independent of the horizontal features; even if vertical diffusion coefficients may depend on the horizontal diffusion coefficients as when a change of vertical coordinate is used. This leads to a separable formulation. This can be improved in the particular case of a homogeneous diffusion by level. For that situation, a three-dimensional correlation tensor can be modeled following Rabier et al. (1998), where horizontal modal variances are replaced at each level by the spectrum of homogeneous diffusion correlations for the level considered. The result is a nonseparable formulation. But, this situation is of less interest as the diffusion is not used optimally to construct heterogeneous horizontal correlation functions.

Inspired by the wavelet formulation of Fisher, extending the spectral formulation, a nonseparable heterogeneous formulation can be constructed with a diffusion operator by replacing the vertical modal correlation in Rabier et al. (1998) with a wavelet representation. The diffusion correlation matrix should be defined by $\mathbf{C} = \mathbf{\Lambda} \mathbf{C} \mathbf{\Lambda}^T$, where

$$\tilde{\mathbf{C}} = \mathbf{L}^{1/2} \mathbf{E}^{-1/2} \mathbf{W}^{-1} \mathbf{V} \mathbf{W}^{-T} \mathbf{E}^{-T/2} \mathbf{L}^{T/2}, \quad (11)$$

where \mathbf{L} is the propagator of the diffusion at each level, \mathbf{W} is the wavelet transform at each level, and \mathbf{V} corresponds to a block diagonal matrix whose blocks \mathbf{V}_α are the vertical correlations associated with wavelet index α , where $\alpha = (j, \mathbf{x})$ is a coupling of scale and position depending on the wavelet transform used. The matrix \mathbf{V} can be estimated from an ensemble of background errors in two steps. The former is the estimation of the

local diffusion tensor at each level, so that $\mathbf{L}^{1/2}$ is now known. The latter is the estimation or the parameterization of $\mathbf{V}_\alpha = (V_\alpha^{l,l'})$, where l and l' denote the level index. The direct estimation from an ensemble is hard to achieve: it is estimated from the ensemble of normalized background errors $\tilde{\epsilon}_b = \Sigma_b^{-1} \epsilon_b$, where Σ_b is the diagonal matrix of standard deviations and ϵ_b a background error, as follows. The correlation between levels l and l' is $V_\alpha^{l,l'} = \text{cor}(\epsilon_{w_\alpha}^l, \epsilon_{w_\alpha}^{l'})$, where $\epsilon_w = \mathbf{W}\mathbf{E}^{1/2}\mathbf{L}^{-1/2}\tilde{\epsilon}_b$, and $\epsilon_{w_\alpha}^l$ is the value of ϵ_w at wavelet index α and level l . However, these relations imply that to invert the propagator \mathbf{L} leads to an ill-posed problem. To overcome this issue, a parameterization should be used. This can be based on the climatology. But this solution is not able to represent any flow dependences, and other parameterizations still need to be found.

With this formulation, the resulting correlation model will benefit from the advantages of both formulations: a good representation of length scale and anisotropy because of the diffusion, and nonseparability because of scale analysis provided by wavelets.

d. Example in a two-dimensional framework of the hybrid formulation

To illustrate the previous formulation in Eq. (11) a simple analytical model is now presented. This is defined as a three-level vertical extension of the 1D framework described in section 2c, and the same wavelet transform is used. Level 1 (level 3) corresponds to the top (bottom) level. The horizontal correlations are modeled by using a diffusion operator. The local diffusion coefficients at level k are computed as $\nu_k(x) = Lp_k^2(x)/2$. Here $Lp_k(x)$ is the length scale field of a heterogeneous correlation tensor obtained as a c_k stretching of a homogeneous Gaussian correlation tensor of length scale L_{hk} . For numerical applications, the stretching coefficients are equal to $c_1 = 1.2$, $c_2 = 1.8$, and $c_3 = 2.4$; the homogeneous length scale values such as $L_{h1} = 1000$ km, $L_{h2} = 625$ km, and $L_{h3} = 250$ km. The length scale field for each level is represented in Fig. 9a (bold lines). Long (short) length scales are found near 0° (180°). The vertical correlation matrices are defined in wavelet space. With the notations of the previous paragraph, the vertical correlation matrix for the couple $\alpha = (j, \mathbf{x})$ is arbitrarily defined by $\mathbf{V}_\alpha = \gamma_\alpha \mathbf{T} + (1 - \gamma_\alpha)\mathbf{I}$, where \mathbf{T} is the symmetric Toeplitz matrix of first row (1, 0.8, 0.5) and $\gamma_\alpha \in [0, 1]$ is a mixing ratio. The \mathbf{T} matrix mimics a strong correlation between the three levels. Here \mathbf{V}_α is a mixing between a strong vertical correlation (when $\gamma_\alpha = 1$) and no vertical correlation (when $\gamma_\alpha = 0$). Thus, the vertical correlation increases with γ_α . The coefficients γ_α used here are represented in Fig. 9b. For the large scales (small j), the vertical correlation is deep (shallow) when

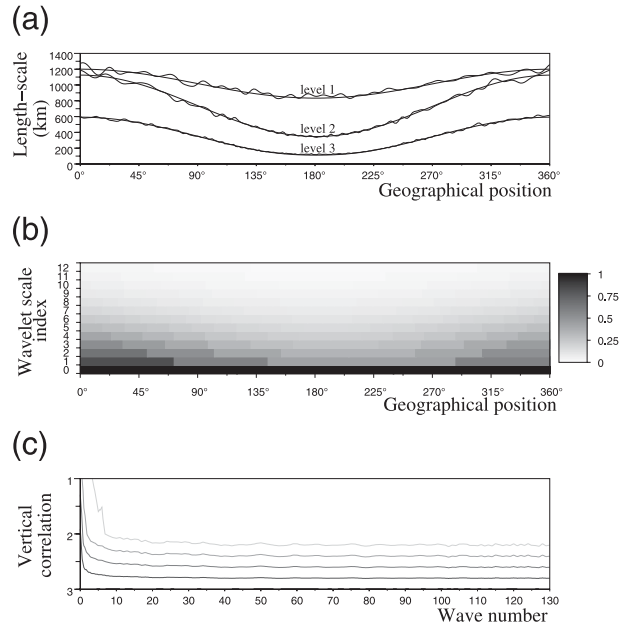


FIG. 9. Illustration of the hybrid formulation in Eq. (11): (a) length scale field per level: initial length scale field (bold lines) and estimated length scale field (thin lines); (b) the values of γ_α used in the example; and (c) the vertical correlations with level 3 as a function of horizontal wavenumber (see the text for details).

the horizontal length scale is long (short); while in the small scales (high j) the vertical levels are not correlated.

To diagnose the vertical and horizontal components of this correlation model, a randomization technique has been used with a large ensemble ($N_e = 1600$). The horizontal length scales diagnosed for each level are represented in Fig. 9a (thin lines). In spite of the sampling noise, the estimated length scale fields accurately reproduced the initial length scale fields (bold lines). The nonseparability of the horizontal and the vertical correlations is illustrated in Fig. 9c. This panel represents the vertical correlation at level 3 versus the horizontal wavenumber. The result is that the vertical correlation is deep for low wavenumbers while it is shallow for high wavenumbers. This proves that the correlation model in Eq. (11) is nonseparable.

In the next section, a comparison between the wavelet formulations and the diffusion is proposed, but only the horizontal features are discussed.

5. Illustration with a toy ensemble of perturbed forecasts

This paper focuses on the representation of horizontal heterogeneity of the background error covariances. In real applications, the true background error correlation matrix is not known. This matrix is generally estimated

from a small ensemble and thus its estimation is affected by sampling noise (Houtekamer and Mitchell 1998; Lorenc 2003). Thus, it is not possible to compare the correlation model with the true correlation matrix, as done for the analytical example in section 2c (see the discussion about the example in Fig. 4) that corresponds to an infinite ensemble framework. To overcome these difficulties, a simple toy model on the sphere has been implemented to generate a large ensemble of perturbed forecasts, providing a correlation matrix of reference. The ensemble is large enough to be considered as infinite. In other words, this correlation matrix corresponds to the truth unknown in practice. The heterogeneous correlation functions produced with this approach are sufficiently complex to be considered as a test bed for the formulation presented in the previous sections. The model is first described and the construction and analysis of the ensemble are also detailed. The analysis of previous formulations is then presented.

a. Description of the model and the ensemble

The spherical model considered is the nonrotating nondivergent barotropic vorticity equation:

$$\partial_t \zeta + J(\Delta^{-1} \zeta, \zeta) = -\nu_2 \Delta^2 \zeta,$$

where ζ is the vorticity of the horizontal wind, t is time, J is the Jacobian operator, Δ is the horizontal Laplacian, Δ^2 is the horizontal bi-Laplacian, and $\nu_2 = 5 \times 10^{14} \text{ a}$ superviscosity coefficient (Yoden and Yamada 1993). The numerical integration is based on a quasi-spectral approach in a triangular truncation $T = 130$, with a classical $3/2$ rules for the dealiasing procedure (Boyd 2001). The time integration is done by a leapfrog scheme (Kalnay 2002) with Asselin's filter to damp numerical modes (Asselin's parameter is set to 1%; Asselin 1972).

An initial state of vorticity $\zeta^r(t = 0)$ has been randomly generated as follows: from an energy density spectrum $E(n) = n(1 + n/n_c)^{-7}$ (n denotes the wave-number) and normalized so that the total energy $E_{\text{tot}} = \sum_{n=0}^T E(n)$ is $E_{\text{tot}} = 300 \text{ m}^2 \text{ s}^{-2}$, and the enstrophy spectrum G_n has been computed according to $G_n = n(n+1)/a^2 E_n$, with the earth radius $a = 6400 \text{ km}$ (Boer 1983; Boer and Shepherd 1983). The cutoff wavenumber n_c is set to 20. Then, the spherical harmonic spectrum ζ_n^m of the vorticity field has been randomly generated with $\zeta_n^m = \sqrt{2G_n/(2n+1)} \xi_n^m$, where ξ_n^m is a sample of a complex Gaussian variable of mean 0 and standard deviation 1. The initial state $\zeta^r(t = 0)$ of reference is represented in Fig. 10a. It corresponds to a heterogeneous field with patterns randomly distributed. The 2D-characteristic length of patterns is given by the differential length scale (Daley 1991, p. 110)

$l = \sqrt{2a^2 \sum_n G_n / \sum_n n(n+1) G_n} = 216 \text{ km}$. The coast lines are represented to facilitate the discussion on localized structures, but no orography is present in the model.

Similarly, an ensemble of $N_e = 800$ independent perturbations $\delta \zeta_k(t = 0)$ (k denotes the number of the member) has been randomly generated with the energy spectrum $E_n' = n(1 + n/n_c)^{-4}$ normalized so that the total energy of each perturbation is $E_{\text{tot}}' = 30 \text{ m}^2 \text{ s}^{-2}$. An example of perturbation, $\delta \zeta_1(0)$ is represented in Fig. 10d. Similarly to the initial field, the 2D-characteristic length of the patterns is $l' = 113 \text{ km}$. The patterns of the perturbation are of smaller scales than in the initial field, and the magnitude of the perturbation is 26% of the magnitude of the reference $\zeta^r(0)$.

The reference $\zeta^r(t = 0)$ [respectively each member of the ensemble of perturbed initial state $\zeta_k(0) = \zeta^r(0) + \delta \zeta_k(0)$] has been integrated over $\tau = 54 \text{ h}$, to obtain $\zeta^r(\tau)$ [respectively $\zeta_k(\tau) = \zeta^r(\tau) + \delta \zeta_k(\tau)$]. The field $\zeta^r(\tau)$ is represented in Fig. 10b. The interaction of coherent structures have emerged during the integration from the initial patterns (McWilliams 1984; Yoden and Yamada 1993), with formation of long filaments of vorticity similar to some frontlike structure. Any perturbed forecasts exhibit a similar time evolution. Slight differences (e.g., resurgences of the initial independent perturbations) still exist however. The perturbation $\delta \zeta_1(\tau)$ is represented in Fig. 10e. Several dipoles appear as near the point at latitude -45° and longitude 0° , or near the point at latitude -45° and longitude -180° . Each dipole corresponds to a phase shift error. Moreover, the error magnitude is now of the same order as the reference magnitude $\zeta^r(\tau)$. Subsequently, the statistics are computed at time τ .

Some simple diagnostics of the ensemble can be computed, such as the ensemble mean $\bar{\zeta}_k$ represented in Fig. 10c. The field is clearly of larger scales than the reference forecast; this is the fingerprint of slight nonlinearities. The standard deviation field $\sigma(\zeta_k)$ is reported in Fig. 10f. Large values of standard deviation are concentrated on coherent structures and on filaments, which is consistent with the fact that coherent structures are the dynamical part while they inhibit the nonlinear instability of the background flow (i.e., the incoherent part of the flow; Kevlahan and Farge 1997).

Other diagnostics of correlation functions can be computed as the local anisotropy (not shown here) or the local length scale. Figure 11 represents the zonal length scale L_x in Fig. 11a and the meridional length scale L_y in Fig. 11b. The length scale field is computed according to the method described section 4b. In that case, zonal and meridional displacement are the finest scale associated to the homogeneous resolution at truncation T . Both length scale fields are very organized and fingerprints of

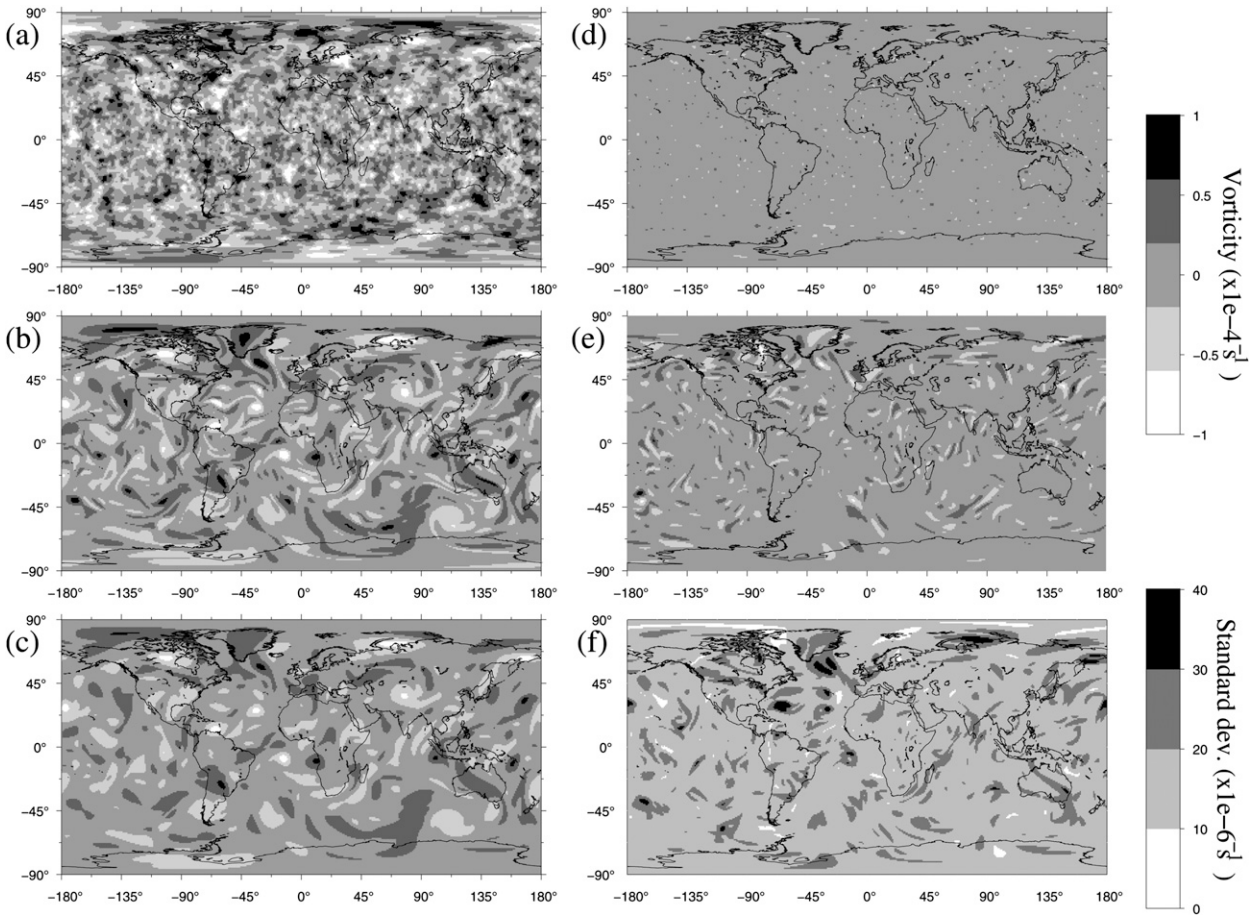


FIG. 10. Representation of (a) initial vorticity field (10^{-4} s^{-1}) and (b) the final field after the time integration of the nondivergent barotropic model. An example of (d) initial perturbation and (e) its nonlinear time evolution are represented. (c) The mean of the ensemble of perturbed forecasts is represented. (f) The standard deviation (in 10^{-6} s^{-1}) of this ensemble is represented. Coast lines are represented to facilitate the presentation but the model does not include any orography.

coherent structures are visible. Areas of strong anisotropy are visible; for example, over the northeast of Greenland, where L_x are almost 100 km while L_y are nearly 300 km. Both length scale fields appear to be intermittent, which is a difficulty for covariance modeling. Now this diagnosis is used to compare the different formulations introduced in previous sections.

b. Comparison of the different correlation models

The three correlation models presented in the preliminary sections are now compared. Each of them has been estimated from the previous ensemble of $N_e = 800$ perturbed forecasts $\zeta_k = \zeta^r + \delta\zeta_k$. Then an ensemble of $N_e = 800$ simulated background errors has been generating as $\epsilon_b = \mathbf{B}_{\text{model}}^{1/2} \xi$, where $\xi \sim \mathcal{N}(0, \mathbf{I})$ is a sample of a Gaussian random vector with zero mean and the identity as covariance matrix (Fisher and Courtier 1995). In this section, only qualitative results are pre-

sented in order to illustrate the feasibility of each method and their behavior. Quantitative results can be found in previous studies about the wavelets formulation and the diffusion operator (see e.g., Pannekoucke et al. 2007; Pannekoucke and Massart 2008; Pannekoucke 2008).

The diagnosed length scale associated with the SBW formulation is reported in Fig. 11c for L_x and Fig. 11d for L_y . The heterogeneity is accurately represented and the main structures present on the reference in Figs. 11a,b are retrieved. For instance, one can notice the long length scale pattern over the Himalayas (cf. Figs. 11a,c) and the short length scale area in West Africa (cf. Figs. 11b,d). But in this experiment, the SBW formulation misses the extremes in the length scale (Figs. 11a,b). Moreover, the range of magnitude variation of the modeled length scale is smaller than the original reference. On the reference, the order of the minimum length scale represented is 100 km while this minimum is now

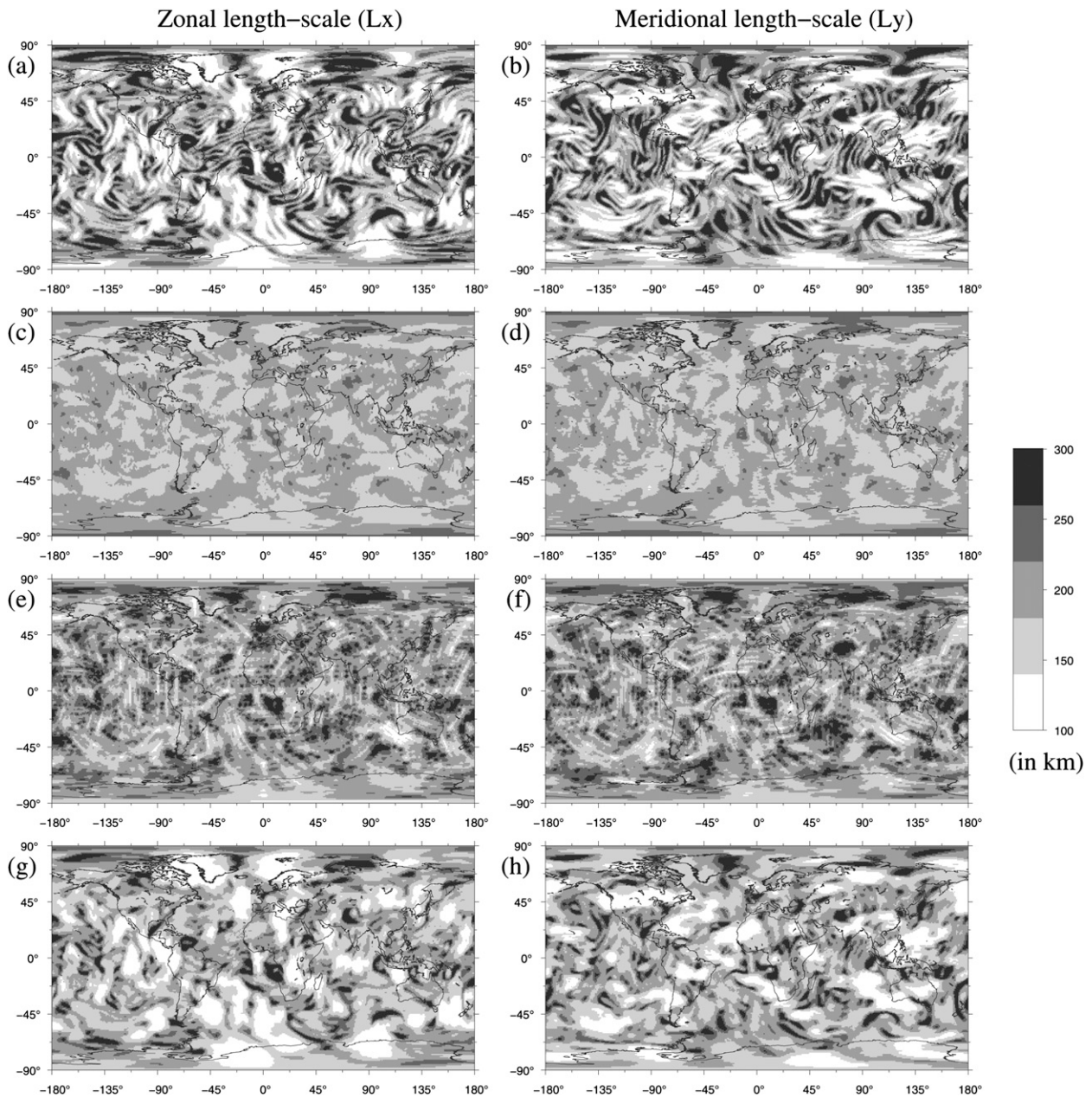


FIG. 11. Diagnosis of the length scale estimated from the raw ensemble or from a randomization: (a) raw zonal length scale L_x and (b) raw meridional length scale L_y ; the length scales modeled with SBW (c) L_x and (d) L_y ; the length scales modeled with SGW (e) L_x and (f) L_y ; and the length scale modeled with the diffusion equation (g) L_x and (h) L_y . [The length scale values low to 100 km (up to 300 km) are in white (black).]

up to 150 km for the wavelets. The height values are also reduced as they are nearby 300 km on the reference and 250 km with the model. This corresponds to the spatial average discussed in section 2b. Even if this spatial average is an advantage for an ensemble of small size (Pannekoucke et al. 2007), in this simulated example, the SBW formulations seems to fail. This should be further investigated. A first guess may be that the SBW

formulation is less appropriate to represent filament structures, which are particularly strong in this toy experiment. Similar to the bad representation of the anisotropy, this could be related to the isotropy of the SBW functions.

Now the SGW diagonal assumption is analyzed. The modeled length scale fields of this formulation are reported in Fig. 11e for L_x and in Fig. 11f for L_y . The

geographical variations of the length scale are present with a good range of length scale value. Nevertheless, the most visible defect is that the length scale fields are dependent on the grid. This might be attenuated by using a spectral normalization (Deckmyn and Berre 2005). This solution seems difficult to achieve as the hierarchical mesh is not adapted for spectral computation, results should depend on the interpolation strategy. Moreover, it appears that anisotropy is not well represented, even if there exists a part anisotropy represented such as over Himalaya.

Finally, the length scales of the model based on the diffusion equation are represented in Fig. 11g for L_x and in Fig. 11h for L_y . With this approach, the geographical variations of the local correlation functions are better represented. Both length scale maps are well correlated with their associated reference (Fig. 11a) for L_x or (Fig. 11b) for L_y . Representation of anisotropy is better than the two previous models (e.g., West Africa). Some slight known defects occur, such as the difficulty with catching fast oscillations at the south of South Africa (Pannekoucke and Massart 2008).

This simulated experiment has shown the various abilities of these formulations to represent geographical variations. In this framework, the correlation model based on the diffusion has been shown better than the SBW and the SGW formulations. Of course the length scale analysis is solely not sufficient, but it gives a preliminary view of the heterogeneity of each formulation.

When the ensemble is small, the above results are affected by the sampling noise. As stated in sections 2b,c, the wavelet formulation reduces the sampling noise. However, the correlation model with the diffusion is estimated from the computation of the local correlation, which is sensitive to the ensemble size. Under a local ergodicity assumption, the knowledge of the correlation sampling distribution, versus the ensemble size, is a possible way to improve the estimation and to reduce the sampling effects (Pannekoucke et al. 2008).

6. Conclusions and perspectives

In this paper, three formulations of correlations model have been studied. Two of them are based on the wavelet diagonal assumption while the third one is based on the diffusion equation.

The spectrally based wavelets (SBWs) has been recalled in detail, especially the filtering properties associated with local spatial averaging offered by the formulation. These properties have been described in a first experimental framework: a simple 1D context with heterogeneous correlation functions. In this framework,

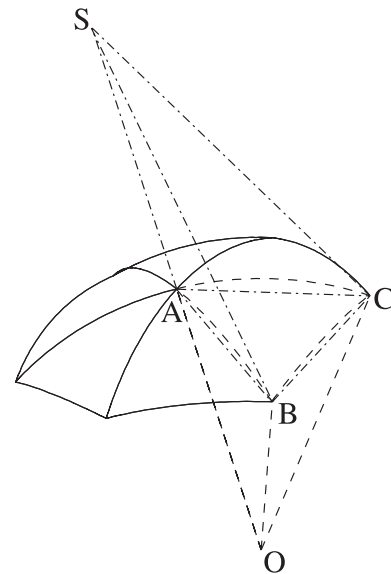


FIG. A1. The scheme to explain the computation of the integral $I_{j,k}$ at the finest scale J . The triangular mesh is in black except the spherical triangle ABC. Here O is the center of the unit sphere (see the text for details).

the filtering of sampling noise that affects the estimation of the local correlation function has been shown to be associated with the cone of influence related to a point. All wavelet variances, outside of the cone of influence, do not contribute to the local correlation function modeled with the diagonal assumption in wavelet space. A second generation wavelet (SGW) formulation has been introduced. As for the spectrally based wavelets, the SGW formulation must present some filtering properties of the sampling noise but this has not been studied here.

The ability of the wavelet diagonal assumption to represent the geographical variations of local correlation functions has been illustrated on a toy experiment on the sphere. It consists of generating a large ensemble with a toy barotropic model to construct complex correlation functions, taken as the reference. The diagnosis of length scale, modeled with the wavelet formulations, has shown the ability of both wavelet formulations to catch geographical variations. Length scale fields of the spectrally based wavelet formulation were too smooth compared with the references but main geographical variations were represented. A possible suspect is the failure of the SBW formulation to model filament structures, but this is purely speculative at this stage and requires more investigation. The length scale fields of the SGW formulation represent well the range of value of the reference but the correlation functions are dependent on the grid. However, this SGW was introduced as a

preliminary test case to more complex usage (e.g., with the introduction of boundary constraints).

It has been demonstrated that it is possible to estimate the local diffusion tensor for the diffusion model. This estimation has been illustrated in the toy experiment, where its representation of anisotropy has been shown to be more accurate than the wavelet formulations presented. Note that the issue of how to estimate the local tensor of diffusion in the generalized diffusion equation is still open. The correlation model based on the diffusion operator has been generalized into a hybrid version that includes a representation of the vertical correlations in wavelet space: there is one vertical correlation matrix per scale and geographical position. The resulting formulation is nonseparable. A simple two-dimensional framework has been used to illustrate this correlation model. Nevertheless the estimation of the vertical correlations in this hybrid formulation leads to an ill-posed problem. Until a satisfactory solution is found a parameterization can be used. This should be further investigated.

Acknowledgments. The author would like to thanks Kayo Ide, Pierre Gauthier, Chris Jones, and Keith Thompson for organizing the workshop on Mathematical Advancement in Geophysical Data Assimilation and for their invitation. He would also like to thank participants in this workshop for stimulating discussions, especially Mark Buehner. He is grateful to Andrew Tangborn for his critical review and fruitful suggestions on the manuscript. A part of this work (section 2) has been done during the Ph.D. of the author, and he would like to warmly thank Gérald Desroziers and Loïk Berre for their supervision and for their support. He would also like to thank Sébastien Massart for his collaboration on the diffusion part.

APPENDIX

Computation of $I_{j,k}$ Integrals

With notation in Fig. 8, the integrals $I_{j,k}$ can be computed recursively (Schröder and Sweldens 1995b) as follows:

$$\forall m \in M^j : \begin{cases} \forall k \in K^j : I_{j,k} = I_{j+1,k} \\ \forall k \in \mathcal{K}_m = \{v_1, v_2\}, I_{j,k} = I_{j,k} + \frac{1}{2} I_{j+1,k} \\ \forall k \in \mathcal{K}_m = \{f_1, f_2\}, I_{j,k} = I_{j,k} + \frac{1}{8} I_{j+1,k} \\ \forall k \in \mathcal{K}_m = \{e_1, e_2, e_3, e_4\}, I_{j,k} = I_{j,k} - \frac{1}{16} I_{j+1,k} \end{cases}$$

For $j = J$ on S^J , the scaling functions ϕ^J correspond to hat functions. The initial value of $I_{j,k}$ can be computed as described now, with the help of Fig. A1. In that figure, ϕ_A^J is the hat function of value one at A and null at the other points of the triangular mesh (e.g., at B or at C). For this particular case where the point A is of degree 6, $I_{J,A}$ appears as the sum of six elementary volumes such as SABC of spherical triangle base ABC. These elementary volumes are computed as for the last volume SABC: it corresponds to the volume of tetrahedron

SABC (dashed line) minus the spherical sector OABC (dashed-dotted line).

The volume of the tetrahedron SABC is also the volume of the tetrahedron OABC that is $||[\mathbf{A}, \mathbf{B}, \mathbf{C}]||/6$, where $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ is the determinant of vectors $\mathbf{A} = \mathbf{A} - \mathbf{O}$, $\mathbf{B} = \mathbf{B} - \mathbf{O}$, and $\mathbf{C} = \mathbf{C} - \mathbf{O}$.

The volume of the spherical sector OABC is $E/3$, where E is the area of the spherical triangle. This area is $E = \alpha + \beta + \gamma - \pi$, with α the angle \widehat{BAC} , $\beta = \widehat{ABC}$ and $\gamma = \widehat{BCA}$; and it can be computed by using l'Huilier's theorem:

$$E = 4 \arctan \left[\sqrt{\tan\left(\frac{p}{2}\right) \tan\left(\frac{p-\alpha}{2}\right) \tan\left(\frac{p-\beta}{2}\right) \tan\left(\frac{p-\gamma}{2}\right)} \right],$$

where $p = (\alpha + \beta + \gamma)/2$.

However, when J is large enough, the volume of the tetrahedron SABC can be a good approximation of the elementary volume.

REFERENCES

Antoine, J., L. Demanet, L. Jacques, and P. Vandergheynst, 2002: Wavelets on the sphere: Implementation and approximation. *Appl. Comput. Harmon. Anal.*, **13**, 177–200.

- Asselin, R., 1972: Frequency filter for time integrations. *Mon. Wea. Rev.*, **100**, 487–490.
- Auger, L., and A. Tangborn, 2004: A wavelet-based reduced rank Kalman filter for assimilation of stratospheric chemical tracer observations. *Mon. Wea. Rev.*, **132**, 1220–1237.
- Belo Pereira, M., and L. Berre, 2006: The use of an ensemble approach to study the background error covariances in a global NWP model. *Mon. Wea. Rev.*, **134**, 2466–2489.
- Berre, L., O. Pannekoucke, G. Desroziers, S. Stefanescu, B. Chapnik, and L. Raynaud, 2007: A variational assimilation ensemble and the spatial filtering of its error covariances: Increase of sample size by local spatial averaging. *Proc. Workshop on Flow-Dependent Aspects of Data Assimilation*, Reading, United Kingdom, ECMWF, 151–168.
- Boer, G., 1983: Homogeneous and isotropic turbulence on the sphere. *J. Atmos. Sci.*, **40**, 154–163.
- , and T. Shepherd, 1983: Large-scale two-dimensional turbulence in the atmosphere. *J. Atmos. Sci.*, **40**, 164–184.
- Bouttier, F., 1993: The dynamics of error covariances in a barotropic model. *Tellus*, **45A**, 408–423.
- , 1994: A dynamical estimation of error covariances in an assimilation system. *Mon. Wea. Rev.*, **122**, 2376–2390.
- Boyd, J. P., 2001: *Chebyshev and Fourier Spectral Methods*. 2nd ed. Dover Publications, 688 pp.
- Buehner, M., and M. Charron, 2007: Spectral and spatial localization of background error correlations for data assimilation. *Quart. J. Roy. Meteor. Soc.*, **133**, 615–630.
- Coifman, R., and W. Wickerhauser, 1992: Entropy-based algorithms for best-basis selection. *IEEE Trans. Info. Theory*, **38**, 713–718.
- Courtier, P., and J. Geleyn, 1988: A global numerical weather prediction model with variable resolution: Application to the shallow-water equations. *Quart. J. Roy. Meteor. Soc.*, **114**, 1321–1346.
- , and Coauthors, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-VAR). I: Formulation. *Quart. J. Roy. Meteor. Soc.*, **124**, 1783–1807.
- Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, 472 pp.
- Daubechies, I., 1992: *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, 377 pp.
- Deckmyn, A., and L. Berre, 2005: A wavelet approach to representing background error covariances in a LAM. *Mon. Wea. Rev.*, **133**, 1279–1294.
- Derber, J., and A. Rosati, 1989: A global oceanic data assimilation system. *J. Phys. Oceanogr.*, **19**, 1333–1347.
- , and F. Bouttier, 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, **51A**, 195–221.
- Desroziers, G., 1997: A coordinate change for data assimilation in spherical geometry of frontal structures. *Mon. Wea. Rev.*, **125**, 3030–3038.
- Dyn, N., D. Levin, and J. Andregory, 1990: Butterfly subdivision scheme for surface interpolation with tension control. *Trans. Graph.*, **2**, 160–169.
- Egbert, G. D., M. G. Foreman, and A. F. Bennett, 1994: TOPEX/Poseidon tides estimated using a global inverse model. *J. Geophys. Res.*, **99**, 24 821–24 852.
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error. *J. Geophys. Res.*, **99**, 10 143–10 162.
- Fisher, M., 2003: Background error covariance modelling. *Proc. ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean*, Reading, United Kingdom, ECMWF, 45–63.
- , 2004: Generalized frames on the sphere, with application to background error covariance modelling. *Proc. ECMWF Seminar on Recent Developments in Numerical Methods for Atmospheric and Ocean Modelling*, Reading, United Kingdom, ECMWF, 87–102.
- , and P. Courtier, 1995: Estimating the covariance matrices of analysis and forecast error in variational data assimilation. Tech. Rep. 220, ECMWF Tech. Memo., ECMWF, 28 pp.
- , and E. Andersson, 2001: Developments in 4D-VAR and Kalman filtering. Tech. Rep. 347, ECMWF Tech. Memo., ECMWF, 36 pp.
- Freeden, W., and U. Windheuser, 1996: Spherical wavelet transform and its discretization. *Adv. Comput. Math.*, **5**, 51–94.
- , and M. Schreiner, 1998: Orthogonal and non-orthogonal multiresolution analysis, scale discrete and exact fully discrete wavelet transform on the sphere. *Constr. Approx.*, **14**, 493–515.
- Gaspari, G., and S. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757.
- Holshneider, M., 1990: Wavelet analysis on the circle. *J. Math. Phys.*, **31**, 39–44.
- , 1996: Continuous wavelet transforms on the sphere. *J. Math. Phys.*, **37**, 4156–4165.
- Houtekamer, P., and H. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811.
- , L. Lefaiivre, J. Derome, H. Ritchie, and H. Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.*, **124**, 1225–1242.
- Ingleby, B., 2001: The statistical structure of forecast errors and its representation in the Met Office global 3-D variational data assimilation scheme. *Quart. J. Roy. Meteor. Soc.*, **127**, 209–231.
- Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 364 pp.
- Kevlahan, N., and M. Farge, 1997: Vorticity filaments in two-dimensional turbulence: Creation, stability and effect. *J. Fluid Mech.*, **346**, 49–76.
- Lönnberg, P., 1988: Developments in the ECMWF analysis scheme. *Proc. ECMWF Seminar on Data Assimilation and the Use of Satellite Data*, Reading, United Kingdom, ECMWF, 75–120.
- Lorenc, A., 2003: The potential of ensemble Kalman filter for NWP—A comparison with 4D-VAR. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183–3203.
- Mallat, S., 1999: *A Wavelet Tour of Signal Processing*. 2nd ed. Academic Press, 637 pp.
- , G. Papanicolaou, and Z. Zhang, 1998: Adaptive covariance estimation of locally stationary processes. *Ann. Stat.*, **26**, 1–47.
- McWilliams, J., 1984: The emergence of isolated coherent vortices in turbulent flow. *J. Fluid Mech.*, **146**, 21–43.
- Mehra, M., and N. Kevlahan, 2008: An adaptive wavelet collocation method for the solution of partial differential equations on the sphere. *J. Comput. Phys.*, **227**, 5610–5632.
- Müller, C., 1966: *Spherical Harmonics*. Springer-Verlag, 45 pp.
- Pannekoucke, O., 2008: Modelisation des structures locales de covariance des erreurs de prévision l'aide des ondelettes. Ph.D. thesis, Université de Toulouse, France, 208 pp.
- , and S. Massart, 2008: Estimation of the local diffusion tensor and normalization for heterogeneous correlation modelling using a diffusion equation. *Quart. J. Roy. Meteor. Soc.*, **134**, 1425–1438.

- , L. Berre, and G. Desroziers, 2007: Filtering properties of wavelets for local background-error correlations. *Quart. J. Roy. Meteor. Soc.*, **133**, 363–379.
- , —, and —, 2008: Background error correlation length-scale estimates and their sampling statistics. *Quart. J. Roy. Meteor. Soc.*, **134**, 497–508.
- Purser, R., W.-S. Wu, D. Parrish, and N. Roberts, 2003a: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. *Mon. Wea. Rev.*, **131**, 1524–1535.
- , —, —, and —, 2003b: Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.*, **131**, 1536–1548.
- Rabier, F., A. McNally, E. Andersson, P. Courtier, P. Undén, J. Eyre, A. Hollingsworth, and F. Bouttier, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-VAR). II: Structure functions. *Quart. J. Roy. Meteor. Soc.*, **124**, 1809–1829.
- Rosca, D., 2005: Locally supported rational spline wavelets on the sphere. *Math. Comput.*, **74**, 1803–1829.
- Sadourny, R., A. Arakawa, and Y. Mintz, 1968: Integration of the nondivergent barotropic vorticity equation with an icosahedral-hexagonal grid for the sphere. *Mon. Wea. Rev.*, **96**, 351–356.
- Schröder, P., and W. Sweldens, 1995a: Spherical wavelets: Efficiently representing functions on the sphere. *Proc. 22nd Conf. on Computer Graphics and Interactive Techniques*, Los Angeles, CA, SIGGRAPH, 161–172.
- , and —, 1995b: Spherical wavelets: Texture processing. Tech. Rep. 4, Industrial Mathematics Initiative, Department of Mathematics, University of South California, 13 pp.
- Sweldens, W., 1995: The lifting scheme: A new philosophy in bi-orthogonal wavelet constructions. *Wavelet Appl. Signal Image Process.*, **III**, 68–79.
- , 1998: The lifting scheme: A construction of second generation wavelets. *SIAM J. Math. Anal.*, **29**, 511–546.
- Tangborn, A., 2004: Wavelet approximation of error covariance propagation in data assimilation. *Tellus*, **56A**, 16–28.
- Weaver, A., and P. Courtier, 2001: Correlation modelling on the sphere using a generalized diffusion equation. *Quart. J. Roy. Meteor. Soc.*, **127**, 1815–1846.
- , and S. Ricci, 2003: Constructing a background-error correlation model using generalized diffusion operators. *Proc. ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean*, Reading, United Kingdom, ECMWF, 327–340.
- Williamson, D., 1968: Integration of the barotropic vorticity equation on a spherical geodesic grid. *Tellus*, **20**, 642–653.
- Yoden, S., and M. Yamada, 1993: A numerical experiment on two-dimensional decaying turbulence on a rotating sphere. *J. Atmos. Sci.*, **50**, 631–643.