Generalized Frames on the Sphere, with Application to Background Error Covariance Modelling

M Fisher

ECMWF
Shinfield Park, Reading, UK

1.   Introduction

The advent of wavelet analysis has fuelled an explosion in harmonic and functional analysis, signal processing and time-frequency analysis. At the same time, wavelets and related constructs have proved themselves useful at the practical level in a wide range of applications, particularly in the field of image processing.

A characteristic of many wavelet applications is that they deal with linear or rectangular domains. This is because the extension of wavelet analysis to more complicated domains, such as the sphere, is non-trivial. In particular, any wavelet-like basis that retains the symmetry properties of the sphere must necessarily be non-orthogonal. This fact requires a considerable generalization of the theory, and demands that we consider non-orthogonal expansions. The mathematical tool for dealing with such expansions is the "frame".

The aim of this paper is to present an introduction to the mathematical theory of frames, and in particular to the notion of a generalized frame. To put some flesh onto the bones of what is otherwise a rather abstract concept, a simple class of generalized frames is discussed. A practical application is presented in the form of a covariance model for background error that allows control over both spatial and spectral aspects of the variation of background error.

The paper is laid out as follows. First, a brief introduction to the theory of discrete frames is presented, and the concept of a generalized frame (Kaiser, 1990) is discussed. Next, a simple class of generalized frames for the sphere is presented, and the decomposition and reconstruction of a function, using a member of this class, is demonstrated. After this, an application of frames to background error covariance modelling is discussed. This application was already presented by Fisher (2003). This paper clarifies the theoretical derivation of the covariance model, and illustrates that the ability of the model to generate good approximations to prescribed, spatially varying, Gaussian covariance structures. We conclude with a brief discussion section.

2.   Discrete and Generalized Frames

Frames were first introduced by Duffin and Schaeffer (1952). However, it is only in recent years, with the advent of wavelet analysis, that interest in them has taken off. The reader is referred to Daubechies (1992) for a very readable introduction to the subject, on which the remainder of this section is closely based.

A family of functions $\{\psi_m ; m \in M\}$ in a Hilbert space $\mathcal{H}$, where M is a countable set of discrete indices, is called a frame if there exist finite, positive bounds A and B such that, for any function f in the space:

$$A\|f\|^2 \le \sum_{m \in M} |\langle f, \psi_m \rangle|^2 \le B\|f\|^2 \tag{1}$$

It can be shown (see, e.g. Daubechies, *op. cit.*), that this condition is sufficient to guarantee the existence of another frame, $\{\tilde{\psi}_m ; m \in M\}$, called the dual frame, with the property:

$$\sum_{m \in M} \langle f, \tilde{\psi}_m \rangle \psi_m = \sum_{m \in M} \langle f, \psi_m \rangle \tilde{\psi}_m = f \quad \forall f \in \mathcal{H} \,. \tag{2}$$

Equation (2) states that we may decompose the function f into a linear combination of the elements of the frame. The coefficients for this decomposition are the projections of the function onto the elements of the dual frame. Alternatively, we may decompose f into a linear combination of elements of the dual frame, in which case the projections onto the original frame provide the coefficients.

The ability to decompose a function into a linear combination of "basis functions", and then reconstruct it, has clear parallels with orthogonal decomposition. The similarity is even more striking in the case of a "tight frame". That is, a frame for which the bounds A and B are equal. In this case, the frame and the dual frame are identical up to a multiplying factor, and we have:

$$f = \frac{1}{A} \sum_{m \in M} c_m \psi_m \quad \forall f \in \mathcal{H} \,,$$
$$\text{where} \quad c_m = \langle f, \psi_m \rangle \tag{3}$$

Compare this, for example, with the Fourier series representation of a function defined on an interval $[a,b]$:

$$f(x) = \sum_{m=-\infty}^{\infty} \hat{f}_m e^{2\pi imx/(b-a)}$$
$$\text{where} \quad \hat{f}_m = \langle f, e^{2\pi imx/(b-a)} \rangle = \frac{1}{b-a} \int_a^b f(x) e^{-2\pi imx/(b-a)} \, dx. \tag{4}$$

Indeed, orthogonal bases (such as the Fourier basis) are tight frames with the additional constraints that $A = 1$ and $\|\psi_m\| = 1 \, \forall m \in M$ (see Daubechies, *op. cit.*).

To illustrate that tight frames need not be orthogonal, consider the following example, due to Daubechies (*op. cit.*). The Hilbert space for this frame is the Euclidean plane, and the frame is defined by the vectors:

$$\psi_1 = (0,1), \quad \psi_2 = \left(-\frac{\sqrt{3}}{2}, -\frac{1}{2} \right), \quad \psi_3 = \left(\frac{\sqrt{3}}{2}, -\frac{1}{2} \right). \tag{5}$$

Given any vector in the plane, $f = (f_x, f_y)$, we have:

$$\sum_{m=1}^{3} |\langle f, \psi_j \rangle|^2 = |f_y|^2 + \left| -\frac{\sqrt{3}}{2} f_x - \frac{1}{2} f_y \right|^2 + \left| \frac{\sqrt{3}}{2} f_x - \frac{1}{2} f_y \right|^2$$
$$= \frac{3}{2} |f|^2. \tag{6}$$

Hence, $\{\psi_1, \psi_2, \psi_3\}$ constitute a tight frame for the Euclidean plane, with frame bound $A = 2/3$. However, the frame is clearly not an orthogonal basis – it consists of three vectors, 120° apart. Proof that the "transform property", i.e. equation (3), holds for this example is left as an exercise for the reader.

The discrete frames described above restrict the set of functions ψ_m to be countable. Recently, Kaiser (1990) (see also Kaiser, 1994) has generalized the notion of a frame by removing this restriction. Specifically, the

set M of indices is required only to be measurable by some measure μ[1]. The condition for a frame (equation (1)) becomes:

$$A\|f\|^2 \le \int_M \left|\left\langle f, \psi_m \right\rangle\right|^2 \, \mathrm{d}\mu(m) \le B\|f\|^2 \quad \forall f \in \mathcal{H} \,. \tag{7}$$

(Note that if M is countable and μ is the counting measure then we retrieve equation (1), so that generalized frames include discrete frames as a special case.)

Most of the properties of discrete frames carry over in an obvious way to the generalized case. In particular, we will make use of the generalization of equation (3) for tight frames, which becomes:

$$f = \frac{1}{A} \int_M c_m \psi_m \, \mathrm{d}\mu(m)$$
$$\text{where} \quad c_m = \left\langle f, \psi_m \right\rangle \tag{8}$$

There is an obvious similarity between this expression and the Fourier transform:

$$f(x) = \int_{-\infty}^{\infty} \hat{f}_m e^{2\pi i m x} \, \mathrm{d}m$$
$$\text{where} \quad \hat{f}_m = \left\langle f, e^{2\pi i m x} \right\rangle = \int_{-\infty}^{\infty} f(x) e^{-2\pi i m x} \, \mathrm{d}x. \tag{9}$$

However, it is worth stressing that, unlike the complex exponentials of the Fourier transform, the functions ψ_m in equation (8) are not generally orthogonal. It is nevertheless useful to think of the c_m as coefficients, albeit in a non-orthogonal expansion, of the corresponding functions ψ_m.

3.   A Class of Tight Generalized Frames on the Sphere

To make the abstract notion of a generalized frame concrete, we consider in this section the definition of a simple class of generalized frames for the sphere. The Hilbert space is the set of complex, square-integrable functions on the sphere, equipped with the natural inner product:

$$\left\langle f, g \right\rangle = \frac{1}{4\pi} \int_\Omega f(\lambda, \phi) \overline{g}(\lambda, \phi) \cos(\phi) \mathrm{d}\lambda \mathrm{d}\phi \,. \tag{10}$$

Here, Ω denotes the surface of the unit sphere, λ and ϕ denote longitude and latitude, and the complex conjugate is denoted by an overbar. The normalization by 4π is conventional in meteorology (see e.g. Courtier *et al.*, 1998).

For the set of indices, M, we choose the set of triples $\left\{ (\lambda, \phi, j) ; \lambda \in [0, 2\pi], \phi \in [-\pi/2, \pi/2], j \in \mathbb{N} \right\}$. Here, λ and ϕ are real numbers denoting longitude and latitude, whereas j is an integer. We associate with M the measure which is the product of the counting measure over $\mathbb{N}$ and the measure $\cos(\phi)\mathrm{d}\lambda\mathrm{d}\phi$ over Ω.

Note that there is one function $\psi_{(\lambda, \phi, j)}$ defined for each point on the sphere and for each positive integer j. For any given triple, the function $\psi_{(\lambda, \phi, j)}$ is itself a function of latitude and longitude, and we will choose it to be a function of great-circle distance from the point (λ, ϕ):

$$\psi_{(\lambda, \phi, j)}(\lambda', \phi') = \Psi_j \left(r(\lambda', \phi', \lambda, \phi) \right), \tag{11}$$

---

[1] More rigorously, $\{M, \mu\}$ must define a σ-finite measure space, and the map $\psi : m \mapsto \psi_m$ must be weakly measurable.

where $\Psi_j \in \mathcal{L}^2(\mathbb{R})$ and where $r(\lambda', \phi', \lambda, \phi)$ denotes the great-circle distance between the points (λ', ϕ') and (λ, ϕ).

Consider now the "coefficients" c_m of equation (8). There is one coefficient for each point on the sphere and for each positive integer j. We will emphasize this by writing $f_j(\lambda, \phi) \equiv c_{(\lambda, \phi, j)}$. That is:

$$f_j(\lambda, \phi) = \left\langle f, \psi_{(\lambda, \phi, j)} \right\rangle = \frac{1}{4\pi} \int_\Omega f(\lambda', \phi') \overline{\Psi}_j \big((r(\lambda', \phi', \lambda, \phi)) \big) \cos(\phi') \mathrm{d}\lambda' \mathrm{d}\phi'. \tag{12}$$

We may regard $f_j(\lambda, \phi)$ as a function f_j of location on the sphere, and equation (12) shows that f_j is the result of a convolution:

$$f_j = \overline{\Psi}_j \otimes f. \tag{13}$$

We seek a tight frame. The condition (equation (7) with $A = B$) is:

$$\sum_{j \in \mathbb{N}} \int_\Omega \left| \left\langle f, \psi_{(\lambda, \phi, j)} \right\rangle \right|^2 \cos(\phi) \mathrm{d}\lambda \mathrm{d}\phi = A \|f\|^2 \quad \forall f \in \mathcal{H}. \tag{14}$$

Writing the inner product as $f_j(\lambda, \phi)$, we see that equation (14) is simply:

$$\sum_{j \in \mathbb{N}} \|f_j\|^2 = A \|f\|^2. \tag{15}$$

That is:

$$\sum_{j \in \mathbb{N}} \left\| \overline{\Psi}_j \otimes f \right\|^2 = A \|f\|^2. \tag{16}$$

We now invoke Parseval's identity, and write equation (16) in terms of the coefficients, $\hat{f}(m, n)$ (etc.) of the spherical-harmonic expansions of f (etc.). Courtier *et al.* (1998) show that the spherical harmonic coefficients of the convolution $\overline{\Psi}_j \otimes f$ are:

$$\left(\widehat{\overline{\Psi}_j \otimes f} \right)(m, n) = \frac{1}{\sqrt{2n+1}} \overline{\hat{\Psi}}_j(0, n) \hat{f}(m, n). \tag{17}$$

Thus, equation (16) gives:

$$\sum_{j \in \mathbb{N}} \sum_{n=0}^\infty \sum_{m=-n}^n \frac{1}{2n+1} \left| \overline{\hat{\Psi}}_j(0, n) \hat{f}(m, n) \right|^2 = A \sum_{n=0}^\infty \sum_{m=-n}^n \left| \hat{f}(m, n) \right|^2. \tag{18}$$

Noting that equation (18) must hold for all functions $f \in \mathcal{L}^2(\Omega)$, we see at once that the condition for a tight frame reduces to:

$$\sum_{j \in \mathbb{N}} \frac{1}{2n+1} \left| \hat{\Psi}_j(0, n) \right|^2 = A \quad \forall n. \tag{19}$$

The "transform" (equation (8)) for the particular frame under consideration, is:

$$f(\lambda, \phi) = \frac{1}{A} \sum_{j \in \mathbb{N}} \int_\Omega f_j(\lambda', \phi') \psi_{(\lambda', \phi', j)}(\lambda, \phi) \cos(\phi') \mathrm{d}\lambda' \mathrm{d}\phi'. \tag{20}$$

Writing $\psi_{(\lambda, \phi, j)}$ in terms of Ψ_j, we see that the integral is another convolution. Thus:

$$f = \frac{1}{A}\sum_{j\in\mathbb{N}}\Psi_j \otimes f \tag{21}$$

where $f_j = \overline{\Psi}_j \otimes f$.

We may verify that equation (21) holds by considering an arbitrary $(m,n)^{\text{th}}$ coefficient of the spherical-harmonic expansion of the sum in equation (21). The convolution becomes a product, so that:

$$\widehat{\left(\sum_{j\in\mathbb{N}}\Psi_j \otimes f_j\right)}(m,n) = \sum_{j\in\mathbb{N}}\frac{1}{\sqrt{2n+1}}\hat{\Psi}_j(n)\hat{f}_j(m,n). \tag{22}$$

But, f_j is itself the result of a convolution with $\overline{\Psi}_j$, so that we may write equation (22) as:

$$\widehat{\left(\sum_{j\in\mathbb{N}}\Psi_j \otimes f_j\right)}(m,n) = \sum_{j\in\mathbb{N}}\frac{1}{2n+1}\left|\hat{\Psi}_j(n)\right|^2 \hat{f}(m,n). \tag{23}$$

Eliminating $\displaystyle\sum_{j\in\mathbb{N}}\frac{1}{2n+1}\left|\hat{\Psi}_j(n)\right|^2$ using equation (19), and dividing through by A gives:

$$\widehat{\left(\frac{1}{A}\sum_{j\in\mathbb{N}}\Psi_j \otimes f_j\right)}(m,n) = \hat{f}(m,n). \tag{24}$$

Since two square-integrable functions whose spherical-harmonic coefficients are all equal must themselves be equal (see e.g. Davis, 1963), it follows from equation (24) that $f = \dfrac{1}{A}\displaystyle\sum_{j\in\mathbb{N}}\Psi_j \otimes f$.

4. Example

In this section, we present an example of the decomposition and reconstruction of a scalar field using a tight frame from the class defined in the preceding section. We begin by defining a set of functions Ψ_j that satisfy the condition for a tight frame, equation (19).

Let us choose a sequence of wavenumbers $\left\{N_j; j\in\mathbb{Z}^*\right\}$ with $N_0 = 0$ and $N_j < N_{j+1}$. For each $j\in\mathbb{N}$, we regard the wavenumbers N_j as nodal points for a spline interpolation, and define $\hat{\Psi}_j(0,n)$ in terms of B-spline basis functions as:

$$\hat{\Psi}_j(0,n) = \sqrt{(2n+1)B_{j-p,p}(n)} \tag{25}$$

where the B-spline basis functions of order p are defined by:

$$B_{j,0}(n) = \begin{cases} 1 & N_j \le n < N_{j+1} \\ 0 & \text{otherwise} \end{cases}$$
$$B_{j,p}(n) = \frac{n-N_j}{N_{j+p}-N_j}B_{j,p-1}(n) + \frac{N_{j+p+1}-n}{N_{j+p+1}-N_{j+1}}B_{j+1,p-1}(n). \tag{26}$$

In particular, for $p = 1$ we have:

$$\hat{\Psi}_j(0,n) = \begin{cases} \sqrt{(2n+1)(n-N_{j-1})/(N_j-N_{j-1})} & N_{j-1} \leq n < N_j \\ \sqrt{(2n+1)(N_{j+1}-n)/(N_{j+1}-N_j)} & N_j \leq n < N_{j+1} \\ 0 & \text{otherwise.} \end{cases} \qquad (27)$$

Figure 1 shows the functions $\sqrt{2n+1}\hat{\Psi}_j(0,n)$ defined by equation (27) for the nodes $0, 2, 4, 8, 16, \ldots 512$. The spline function $B_{j,p}$ is identically zero outside the interval $[N_j, N_{j+p+1})$. It follows that the functions f_j defined in equation (13) are strictly band-limited. Their spherical harmonic coefficients are zero outside the range of wavenumbers $N_{j-p} \leq n < N_{j+1}$. As a consequence, each f_j is exactly determined by its values at a finite set of points on the sphere (for example, the points of a "Gaussian grid" with $2N_{j+1}$ longitudes and $(2N_{j+1}+1)/2$ latitudes).

The correspondence between total wavenumber and spatial scale implies that each of the functions f_j may be thought of as representing a particular spatial scale. At the same time, in the tight frame interpretation described above, we regard the value of f_j at a particular point (λ, ϕ) as representing the coefficient of $\psi_{(\lambda,\phi,j)}$. For this interpretation to be meaningful, we require that the function $\psi_{(\lambda,\phi,j)}$ should be localised around the point (λ, ϕ). Although strict compact support is not possible for band-limited functions, the plots of $\psi_{(\lambda,\phi,j)}$ shown in Figure 2 and Figure 3 demonstrate that for practical purposes the functions are well localised.
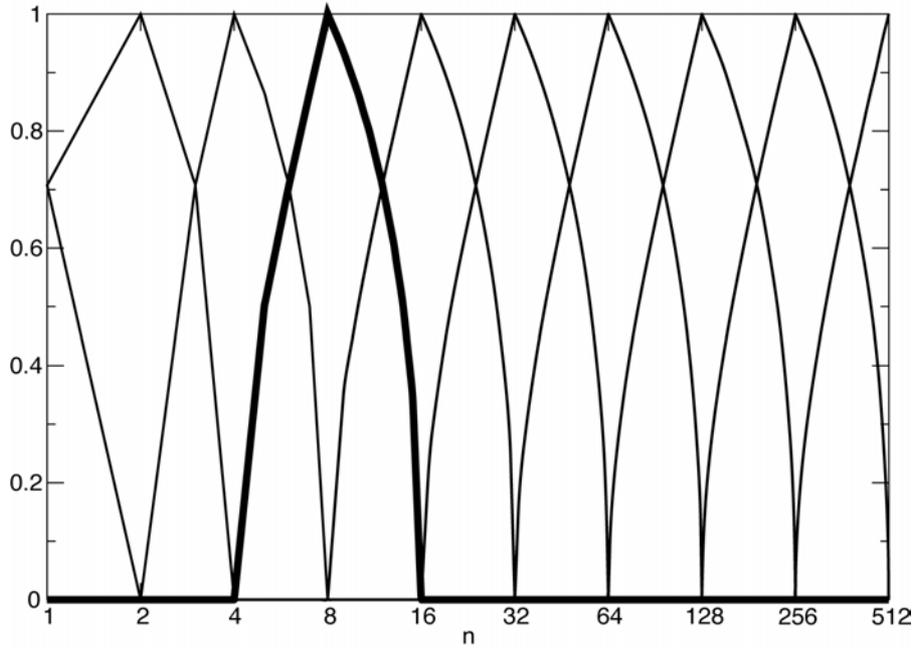

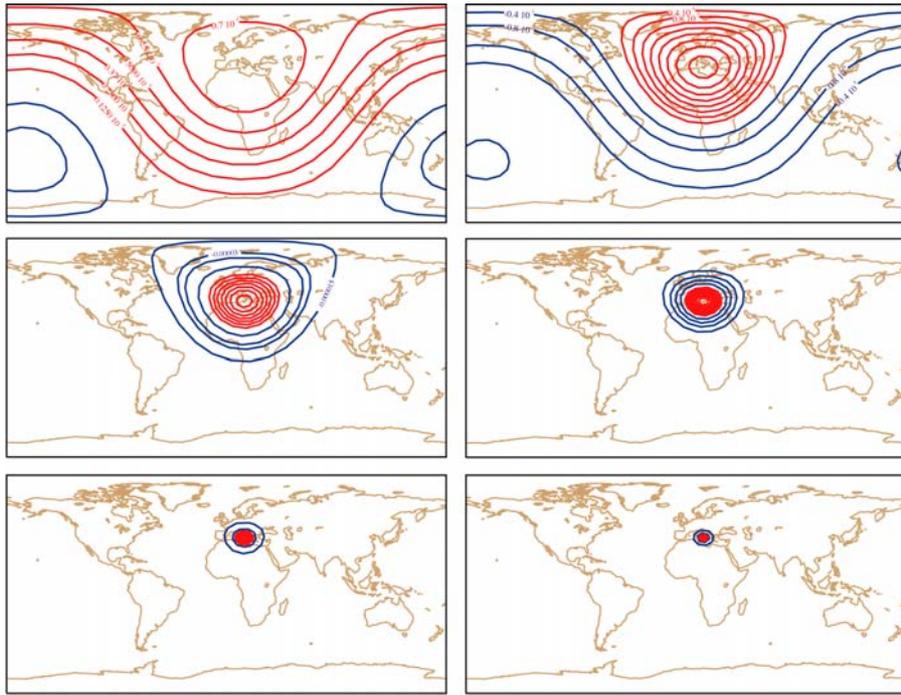Figure 1: The functions $\sqrt{2n+1}\hat{\Psi}_j(0,n)$ defined by equation (27) for one choice of nodes N_j.

Figure 2: The functions $\psi_{(\lambda,\phi,j)}$ for a point (λ,ϕ) over Sicily, and for $j = 1\ldots 6$, corresponding to the spectral coefficients shown in Figure 1.
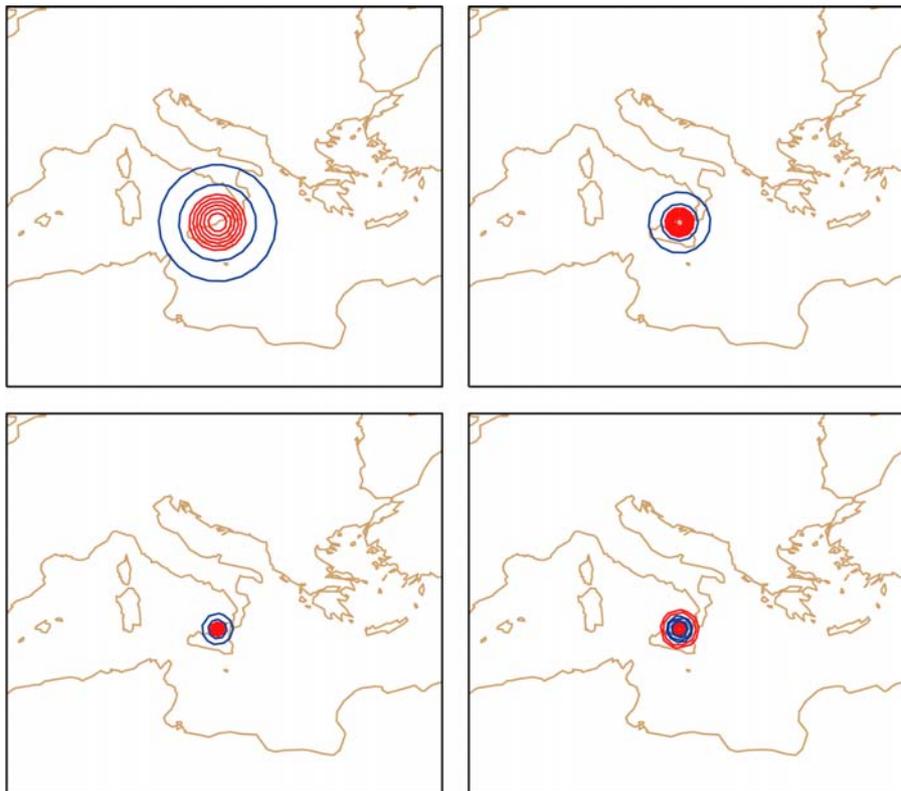

Figure 3: As in Figure 2, but for $j = 7\ldots 10$.

To illustrate the decomposition and reconstruction of a function on the sphere, we will consider the orography of the ECMWF $T_L 511$ model, shown in Figure 4. The functions f_j, corresponding to the convolutions of the orography with the functions shown in Figure 1, are shown in Figure 5. It is clear that the convolutions achieve a separation of spatial scale, while retaining the local nature of the features they

describe. This is consistent with the interpretation of each point in each of the plots shown in Figure 5 as the coefficient of a function that is simultaneously spatially and spectrally localised.
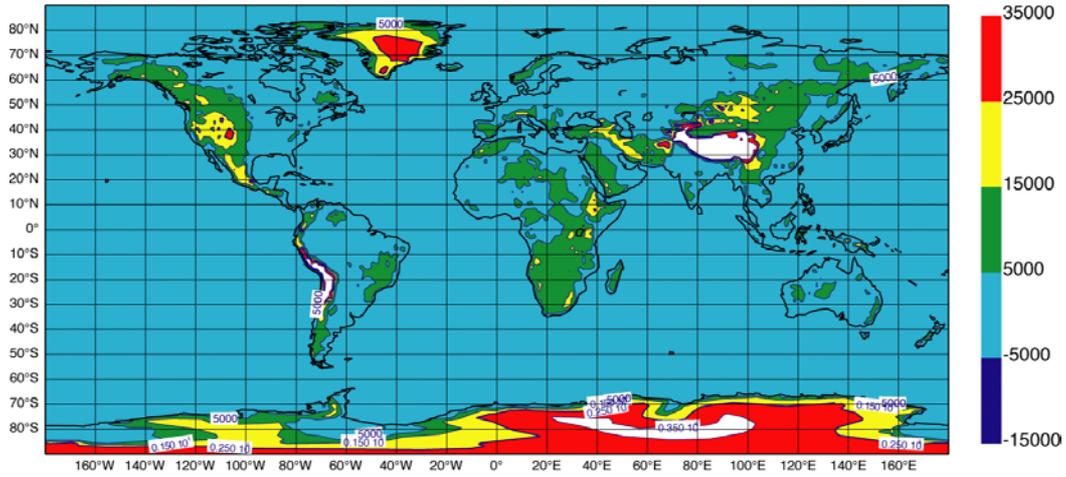

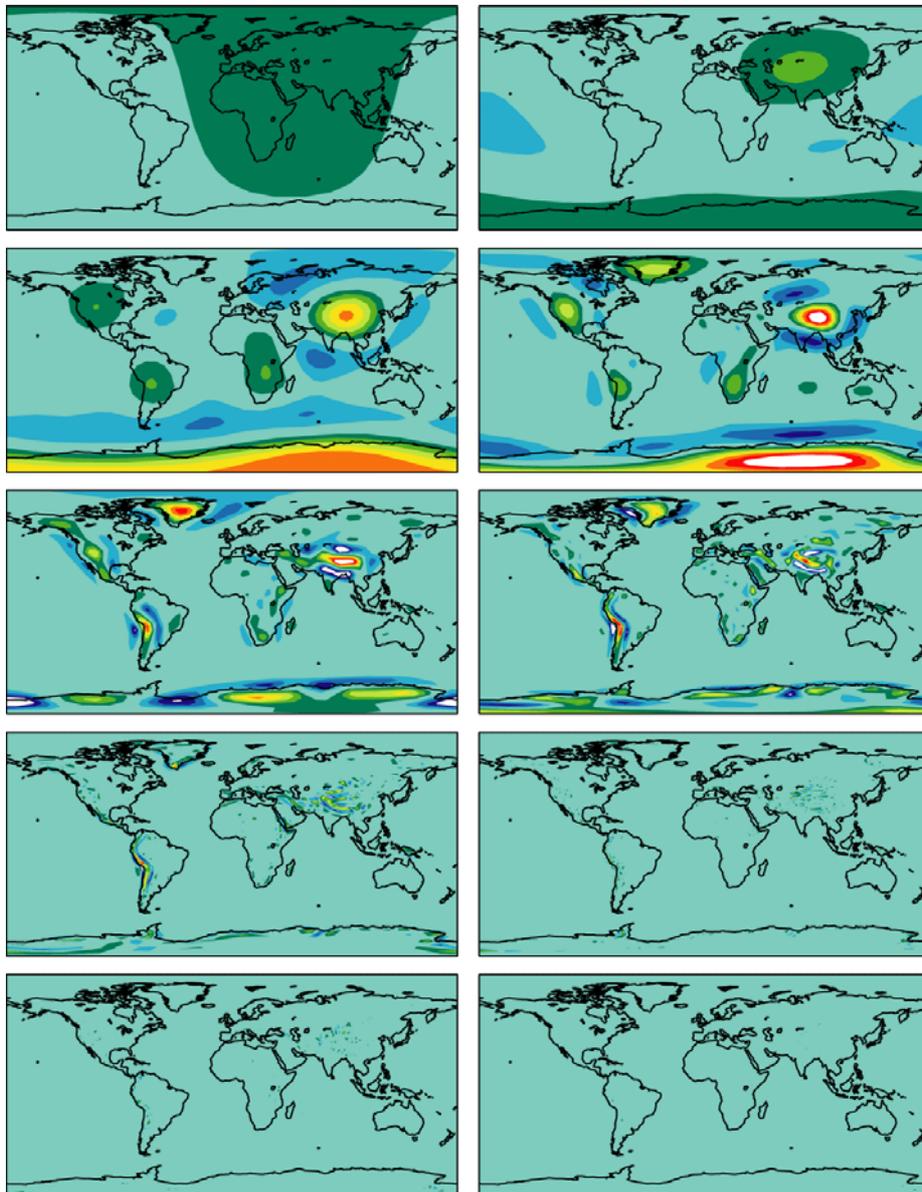*Figure 4: The orography of the ECMWF T_L511 model.*

Figure 5: The functions f_j for the orography field.

The result of convolving each function f_j with the corresponding function Ψ_j is shown in Figure 6. The sum of these fields (not shown) is identical to the original orography. Importantly, the reconstruction of f from the functions f_j retains the local nature of the features shown in Figure 5.
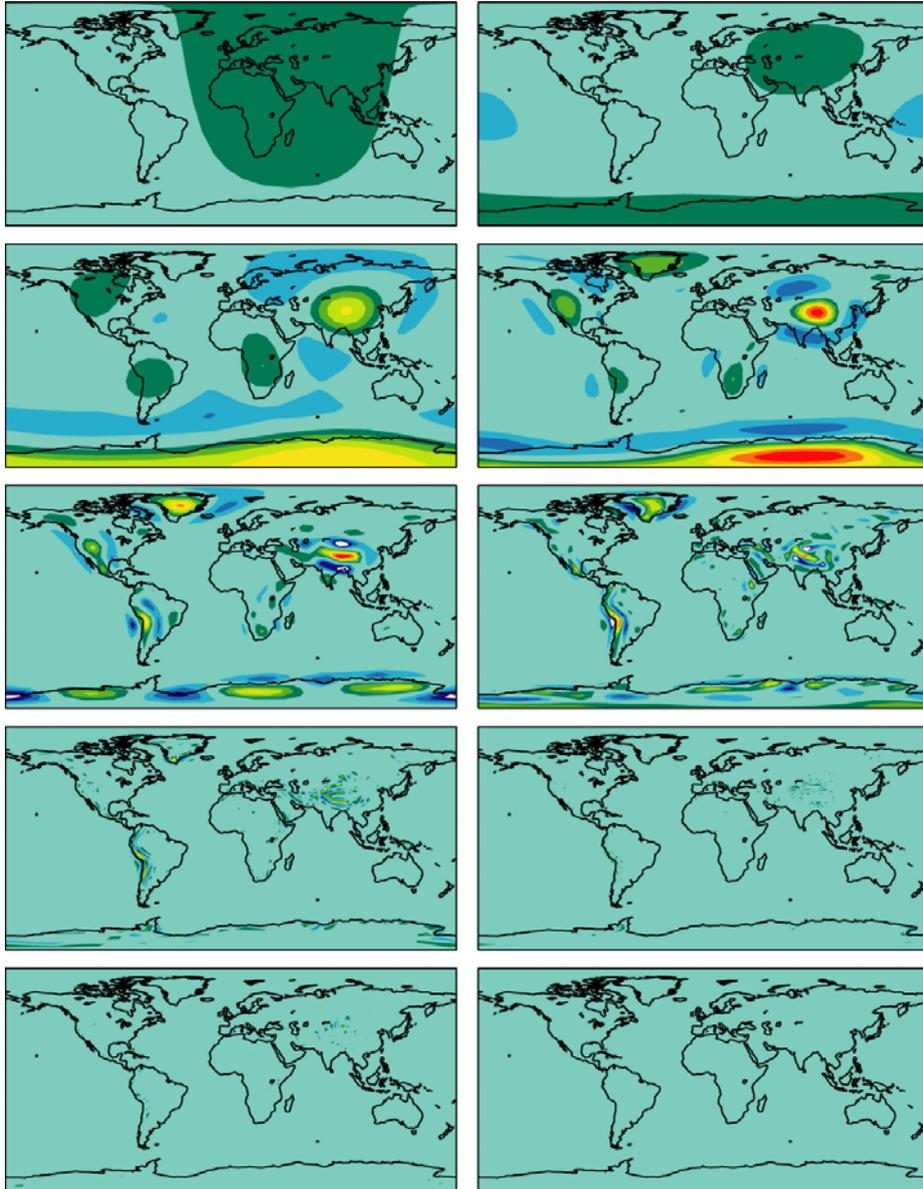

Figure 6: The functions $\Psi_j \otimes f_j$. Summing the fields shown in this figure reconstructs the original orography.

5.   Application to Background Error Covariance modelling

A background error covariance model based on the generalised frame decomposition described above was discussed by Fisher (2003), and dubbed "Wavelet J_b". In this section we examine this covariance model in further detail, and provide evidence of its ability to capture important aspects of the spatial and spectral variation of the covariance of background error. We begin by briefly reviewing the variational form of the data assimilation problem.

Variational data assimilation estimates a discretization of the true state of a system (e.g. the atmosphere) as the vector **x** that minimizes the cost function:

$$J_x(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + (\mathbf{y} - H(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y} - H(\mathbf{x})). \tag{28}$$

Here, $\mathbf{x}_b$ is a prior (background) estimate for $\mathbf{x}$. The vector $\mathbf{y}$ is a vector of observations, and H is an operator that attempts to map the state vector to the space of the observations. The matrices $\mathbf{R}$ and $\mathbf{B}$ are covariance matrices of observation error and background error.

Typically, direct minimization of the cost function defined in equation (28) is numerically very poorly conditioned. For this reason, it is usual to define a control vector, $\boldsymbol{\chi}$, which diagonalizes the background error covariance matrix. The state vector, $\mathbf{x}$, is determined as an affine transformation of $\boldsymbol{\chi}$:

$$\mathbf{x} = \mathbf{x}_b + \mathbf{L}\boldsymbol{\chi}. \tag{29}$$

Written in terms of $\boldsymbol{\chi}$, the cost function becomes:

$$J_\chi = \boldsymbol{\chi}^T\boldsymbol{\chi} + (\mathbf{y} - H(\mathbf{x}_b + \mathbf{L}\boldsymbol{\chi}))^T \mathbf{R}^{-1}(\mathbf{y} - H(\mathbf{x}_b + \mathbf{L}\boldsymbol{\chi})). \tag{30}$$

Note that the background error covariance matrix does not appear explicitly in equation (30). Rather, it is defined implicitly by the choice of the matrix $\mathbf{L}$. It is straightforward to show that minimization of J_χ and J_x give identical state estimates if $\mathbf{B}=\mathbf{L}\mathbf{L}^T$. Thus, modelling of the background covariances in variational assimilation is achieved in practice by choosing a suitable control variable $\boldsymbol{\chi}$ and transformation matrix $\mathbf{L}$.

There is no requirement for the matrix $\mathbf{L}$ to be invertible, or even square. Fisher (2003) took advantage of this fact, and chose:

$$\boldsymbol{\chi} = \begin{pmatrix} \boldsymbol{\chi}_1 \\ \boldsymbol{\chi}_2 \\ \vdots \\ \boldsymbol{\chi}_K \end{pmatrix}. \tag{31}$$

Here, $\boldsymbol{\chi}_1 \ldots \boldsymbol{\chi}_K$ are exact discrete representations, on appropriate grids, of the band-limited functions f_j defined in the preceding section.

Whereas Fisher (*op. cit.*) demonstrated how a three-dimensional covariance model could be constructed, in this paper we describe a two-dimensional covariance model. To cast things into discrete form, let us define the diagonal matrix $\boldsymbol{\Psi}_j$ with diagonal elements $\hat{\Psi}_j(0,n)/\sqrt{2n+1}$. Then, letting $\mathbf{S}_j$ represent the discrete spherical transform from the grid associated with $\boldsymbol{\chi}_j$, the spectral representation of the convolution $\Psi_j \otimes f_j$ may be written as $\boldsymbol{\Psi}_j\mathbf{S}_j\boldsymbol{\chi}_j$.

The transformation matrix $\mathbf{L}$ defining the covariance model will be defined as:

$$\mathbf{L} = (\boldsymbol{\Psi}_1\mathbf{S}_1\boldsymbol{\Sigma}_1 \mid \boldsymbol{\Psi}_2\mathbf{S}_2\boldsymbol{\Sigma}_2 \mid \ldots \mid \boldsymbol{\Psi}_K\mathbf{S}_K\boldsymbol{\Sigma}_K). \tag{32}$$

Here, the matrices $\boldsymbol{\Sigma}_j$ are diagonal in grid space. Their elements determine both the spatial and the spectral variation of the standard deviation of background error.

The background error covariance matrix (in spectral space) associated with L is:

$$\mathbf{B} = \mathbf{L}\mathbf{L}^T = \sum_{j=1}^{K} \boldsymbol{\Psi}_j\mathbf{S}_j\boldsymbol{\Sigma}_j^2\mathbf{S}_j^{-1}\boldsymbol{\Psi}_j \tag{33}$$

Consider the special case $\Sigma_j = \sigma_j \mathbf{I}$. In this case, the matrices Σ_j and $\mathbf{S}_j$ commute, and equation (33) reduces to $\mathbf{B} = \sum_{j=1}^{K} \sigma_j^2 \mathbf{\Psi}_j^2$. This is a diagonal matrix whose elements vary only with wavenumber n. The diagonal elements of $\mathbf{B}$ are therefore modal variances (see Courtier *et al.* 1998), and are given by:

$$b_n = \sum_{j=1}^{K} \frac{\sigma_j^2}{\sqrt{2n+1}} \hat{\Psi}_j^2(0,n) . \tag{34}$$

If the functions Ψ_j are defined using splines, as in equation (26), then equation (34) is:

$$b_n = \sum_{j=1}^{K} \sigma_j^2 B_{j-p,p}(0,n) . \tag{35}$$

Thus, for $\Sigma_j = \sigma_j \mathbf{I}$, the Wavelet J_b covariance model is identical with the two-dimensional covariance model defined by Courtier *et al.* (1998), with the restriction that the modal variances, b_n, are defined by a spline interpolation between values $b_{N_j} = \sigma_j^2$ specified at the nodal wavenumbers N_j. At intermediate wavenumbers, the modal variances are interpolated. Since the spline functions $B_{j,p}$ are positive, the interpolated variances are also positive. Hence, $\mathbf{B}$ represents a valid covariance model.

Analysis of the case where the elements of Σ_j vary is difficult. However, the local nature of the functions $\psi_{(\lambda,\phi,j)}$ implies that at any given point (λ,ϕ) on the sphere, the covariance with neighbouring points will be determined by the values of Σ_j corresponding to points near (λ,ϕ). That is, we should regard the elements of Σ_j as a local equivalent of the modal variance. To demonstrate that this interpretation is valid, we construct a covariance model whose structures are isotropic and approximately Gaussian, but whose length scales vary with location.

Weaver and Courtier (2001) give the following expression for the modal variances of an approximately Gaussian isotropic correlation model on the sphere:

$$b_n = \frac{\exp\left(-n(n+1)L^2/2a^2\right)}{\sum_{n}^{N_{max}} (2n+1)\exp\left(-n(n+1)L^2/2a^2\right)} \tag{36}$$

Here, L is the length scale of the correlation function, and a is the radius of the Earth.

With this covariance model in mind, let us define the elements of Σ_j as:

$$\sigma_j^2(\lambda,\phi) = \frac{\left(N_{max}+1\right)^2 \exp\left(-N_j(N_j+1)L^2(\lambda,\phi)/2a^2\right)}{\sum_{n}^{N_{max}} (2n+1)\exp\left(-n(n+1)L^2(\lambda,\phi)/2a^2\right)} , \tag{37}$$

where the length scale L is a function of location on the sphere and $N_{max} = 511$ is the upper truncation limit. The factor $\left(N_{max}+1\right)^2$ in the numerator of equation (37) is included to give a covariance matrix whose variance is, to within the accuracy of the spline interpolation, equal to that of the identity matrix.

Figure 7 shows the spatial variation of length scale L used for this example. Note that L is a simple combination of trigonometric functions of latitude and longitude, and is not intended to be a realistic representation of the actual spatial variation of length scale for background error at any atmospheric level.
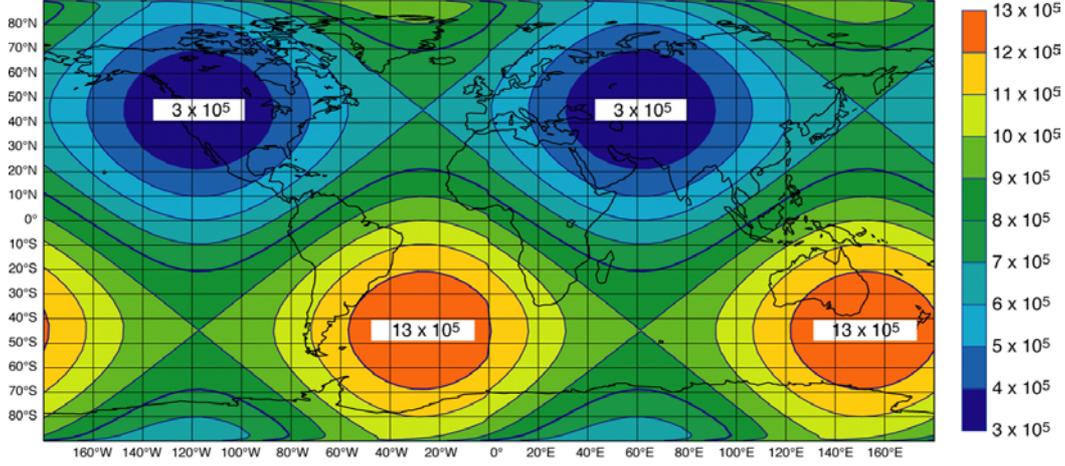

Figure 7: Length scale (metres) for the covariance model.

To reveal the covariance structures defined by the model, we applied the matrix **B**, defined in equation (33) to a field consisting of twelve delta functions. The result, shown in Figure 8, is the sum of twelve rows of the covariance matrix. However, since the delta functions are well separated, we expect the covariance structure near each delta function to be dominated by the corresponding row of **B**. It is clear from Figure 8 that the covariance model is successful in producing covariance structures with a wide range of length scales.

Cross-sections along the three lines of longitude containing the delta functions are shown by the solid lines in Figure 9. In general, the desired Gaussians (shown in blue) are well approximated. However, the peak amplitude is not correct, particularly for longer length scales. This is probably a consequence of inaccuracies in the spline interpolation. (Note that for constant L, the total variance implied by equation (37) is $\sum_{n=0}^{N_{max}} (2n+1)b_n$. Because b_n is interpolated between values specified at the nodal wavenumbers, the total variance is only approximately equal to $(N_{max} + 1)^2$).

The covariance structures corresponding to the shortest length scales are more sharply peaked than the corresponding Gaussian. This is again likely to be a consequence of the spline interpolation. This supposition is supported by the dotted black lines in Figure 9, which show cross sections for a covariance model constructed using equation (37), but with a larger set of nodal wavenumbers.
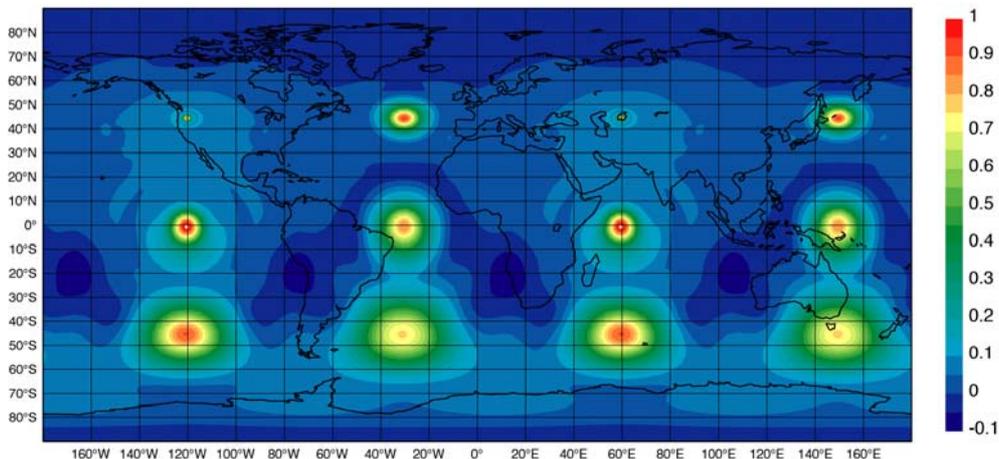

Figure 8: Covariance structures at twelve locations generated by the "Wavelet J_b" covariance model.
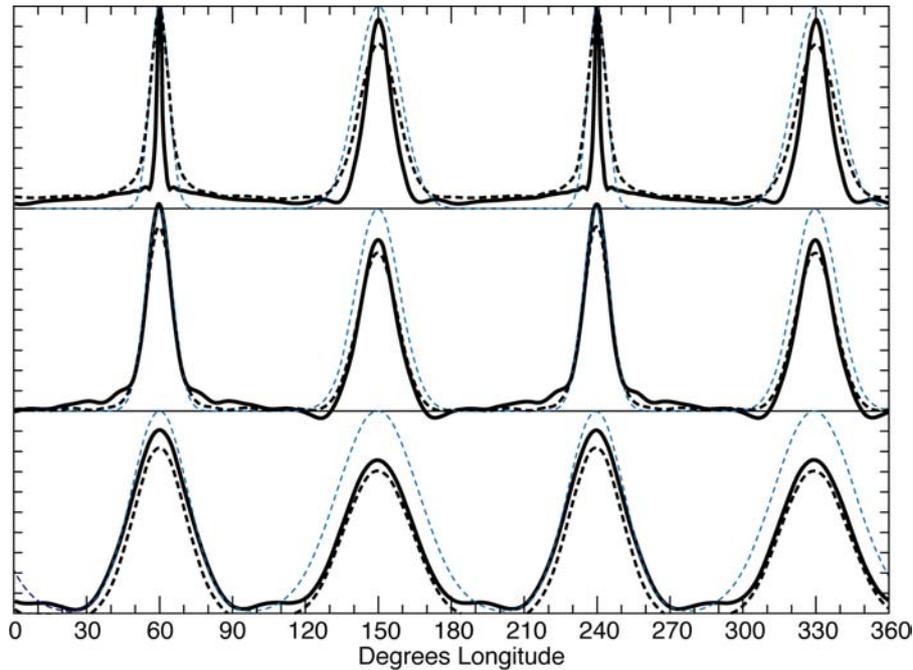
Figure 9: Cross-sections of the covariance structures along lines of longitude passing through the maxima. The dashed blue lines show Gaussians with the desired length scales. The solid black lines show cross sections of the structures shown in Figure 8. The dotted black lines show cross sections for a covariance model with a different set of nodal wavenumbers.

6.   Discussion

The theory of generalized frames provides a mathematically rigorous tool for constructing non-orthogonal decompositions which share many of the properties of their orthogonal cousins. In particular, they allow functions to be transformed into linear combinations of basis functions, and then reconstructed. The range of functions available to generalized frames is far wider than the range of orthogonal functions. In particular, it is relatively straightforward to construct basis functions that are localised in both space and scale, so that frames with many of the characteristics of wavelet bases can be constructed, even for such awkward domains as the surface of the sphere.

The particular generalized frames considered in this paper are rather simple. Indeed, they could easily have been discussed without recourse to generalized frame theory, since they are extensions to the spherical domain of the perfect reconstruction filter banks (there is one in your MP3 player!) of signal processing theory (see e.g. Nguyen, 1995). On the other hand, frame theory produces a particularly elegant derivation, and allows an interpretation of the coefficient functions f_j that is not obvious if, for example, the functions

$\hat{\Psi}_j$ are regarded simply as filters.

There are many different formulations of wavelet, and wavelet-like bases for the sphere. Some formulations (e.g. Göttelmann, 1997; Schröder and Sweldens, 1995) define orthogonal bases using refinable griddings of the sphere. Although these approaches are capable of producing complete bases for $\mathcal{L}^2(\Omega)$, they necessarily suffer from the problem that any finite truncation of the basis contains special points (e.g. the poles). Avoiding this problem requires giving up on orthogonality, so that we are forced to consider frames rather than orthogonal bases. There are several approaches (see, for example, Freeden and Windheuser, 1996; Antoine and Vandergheynst, 1999; and Mhaskar *et al.*, 2000). A very broad family of "spherical wavelets" is described by Freeden et al. (1998). This family includes the wavelet-like frames described in this paper.

The generalized frames considered in this paper are tight. By removing the tightness constraint, a broader class of decompositions may be generated. Such frames use different functions during decomposition and reconstruction, making interpretation of the functions f_j more difficult. For example, it is straightforward to define non-tight convolution-based frames that satisfy equations of the form:

$$f = \sum_{j \in \mathbb{N}} \tilde{\Psi}_j \otimes f$$
$$\text{where } f_j = \Psi_j \otimes f.$$

(38)

It is not obvious if such frames have any significant advantages over the corresponding tight frames.

We have presented a demonstration of the "Wavelet J_b" covariance model. The model was successful in producing approximately Gaussian covariance structures with spatially-varying length scales. Some problems were noted in reproducing the correct variance, and in generating Gaussian-shaped correlation structures for small length scales. It is likely that these problems could be significantly reduced by a more careful choice of nodal wavenumbers, and by the use of higher order (e.g. quadratic or cubic) spline interpolation. Ultimately, the degree to which both spectral and spatial variation of covariance structure can be modelled is limited by the spherical equivalent of the uncertainty principle (see e.g. Freeden *et al.* 1998).

7.   References

Courtier P., Andersson E., Heckley W., Pailleux J., Vasiljević D., Hamrud M., Hollingsworth A., Rabier F. and M. Fisher, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. Quar. J. Roy. Meteor. Soc., **124**, 1783-807.

Daubechies I., 1992: Ten Lectures on Wavelets. CBMS-NSF regional conference series in applied mathematics, SIAM.

Davis P.J., 1963: Interpolation and Approximation. Blaisdell Publishing Company.

Duffin R.J. and A.C. Schaeffer (1952): A class of nonharmonic Fourier series. Trans. Amer. Math. Soc., 72, 341-366.

Fisher M. (2003): Background Error Covariance Modelling. Proc. ECMWF Seminar on Recent developments in data assimilation for atmosphere and ocean, 8-12 September. 45-64.

Freeden W. and U. Windheuser, 1996: Spherical wavelet transform and its discretization. Adv. Comput. Math. 5, 51-94.

Freeden W., Gervens T. and M. Schreiner, 1998: Constructive approximation on the sphere. Oxford University Press series on Numerical Mathematics and Scientific Computation. ISBN 0-19-853682-8.

Göttelmann, J, 1996: Locally Supported Wavelets on the Sphere. Technical report, Department of Mathematics, University of Mainz, Germany.

Kaiser G., 1990: Quantum physics, relativity and complex spacetime, North-Holland Mathematics Studies, vol. 163, North-Holland, Amsterdam, ISBN 0-444-88465-3.

Kaiser G., 1994: A Friendly Guide to Wavelets. Birkhäuser Boston.

Mhaskar H.N., Narcowich F.J., Ward J.D. and J. Prestin, 2000: Polynomial frames on the sphere. Advances in Computational Mathematics.**13**, 387-403.

Nguyen T.Q., 1995: A Tutorial On Filter Banks And Wavelets. citeseer.ist.psu.edu/nguyen95tutorial.html

Schröder P. and W. Sweldens, 1995: Spherical wavelets: efficiently representing functions on the sphere. Computer Graphics, Annual Conference Series (Siggraph'95 Proceedings), pp. 161-172, http://citeseer.nj.nec.com/schroder95spherical.html

Weaver A. and P. Courtier, 2001: Correlation modelling on the sphere using a generalized diffusion equation. Quart. J. Roy. Meteor. Soc. **127**, 1815-1846.