On control variable transforms in the Met Office 3d and 4d Var., and a description of the proposed waveband summation transformation

Ross Bannister*

Document started: February 2003

A review of the need for control variable transforms, a description of the current scheme, and a discussion of a possible extension to the current scheme using wavebands.

1 Modification History from November 2003

28/11/03	1.3	Added note that non-correlation is maintained when rotating modes.
	2.3	Make clear that separate horizontal transforms are performed for each vertical mode.
	2.4	Add comment about emsemble method for estimating B. Make clear that separate horizontal length scales emerge for each vertical mode.
15/12/03	2.5	Rewrite section on vertical transforms and change some of the notation (hopefuly for improved clarity).
16/12/03	2.6	Highlight the shortcomings of the current transforms (section 2.5).
21/12/03	2.7	Update appendix 3 showing that current structure functions are not separable (with and without rotation of the vertical modes) except for special cases.
06/01/04	2.8	Add discussion about calibration of WS transform.
11/01/04	2.9	General tidy-up.
20/01/04	3.0	Add figs. from Ingleby. Whole text revise.
27/01/04	3.1	Finish draft discussion of calibration of WS transform.
09/03/04	3.2	Complete document revisions to completion.
21/05/04	3.3	Add 'v' (for vertical) and band indices on vertical transforms for WS transform. Modify comments about horizontal variances for band j being zero outside band j's wavenumber band (as this may not be true).
22/07/04	3.4	Include option to calibrate WS transform with triangular functions.
30/08/04	3.5	Discuss the gamma spectral correction function (appendix 6).

2 Contents

*Data Assimilation Research Centre, Dept. of Meteorology, Univ. of Reading, Reading, RG6 6BB, UK

- Acknowledgements

Variational data assimilation (Var.) for the atmosphere is the procedure of combining observations with dynamical knowledge of atmospheric behaviour. The objective is to produce an atmospheric state that is optimally consistent with all of the information. The cost function J, below, is a measure of the misfit between (i) a state vector, $\vec{x}$ (to be determined), and the prior knowledge in the form of the background, $\vec{x}_B$, and (ii) the real observations, $\vec{y}$, and the observations predicted by $\vec{x}$ through the forward model, $\vec{H}$. The cost function is [1],

$$\begin{aligned} J[\vec{x}'] &= \frac{1}{2}(\vec{x}' - \vec{x}'_B)^T \mathbf{B}^{-1} (\vec{x}' - \vec{x}'_B) + \frac{1}{2}(\vec{y} - \vec{H}[\vec{x}', \vec{x}_0])^T (\mathbf{E} + \mathbf{F})^{-1} (\vec{y} - \vec{H}[\vec{x}', \vec{x}'_0]), \\ &= J_B + J_O, \end{aligned} \tag{1}$$

which is minimized at $\vec{x} = \vec{x}_a$. This state is called the analysis. In Eq. 1, primed quantities are perturbation quanties, *e.g.* $\vec{x}' = \vec{x} - \vec{x}_0$, where x_0 is a linearization state (often, but not always taken to be the background state), and $\vec{H}$ is the forward operator that predicts the observations from the model state at the current iteration. Provided that $\vec{x}_B$ is unbiased with respect to the true state, that $\vec{H}[\vec{x}', \vec{x}'_0]$ is unbiased with respect to $\vec{y}$ and the errors of $\vec{x}_B$ and $\vec{y}$ have been characterized correctly in the respective error covariance matrices $\mathbf{B}$ and $\mathbf{E} + \mathbf{F}$, $\vec{x}_a$ will be the optimal combination of information.

The focus of this paper is to describe how, at the time of writing, the Met Office represents the background term of the cost function (J_B, the first term of Eq. 1). The background error covariance matrix, $\mathbf{B}$ represents how errors in each model variable at each grid point covaries - in a Gaussian sense - with every other. Represented explicitly $\mathbf{B}$ holds a huge amount of information. The dimensionality of $\vec{x}$ is typically $\sim 10^6$-10^7 elements, and so $\mathbf{B}$ will consist of $\sim 10^{12}$-10^{14} matrix elements, which is beyond the storage capability of present day computer resources (and note in Eq. 1 that the matrix $\mathbf{B}$ has to be inverted).

Instead of attempting to representing $\mathbf{B}$ explicitly, the Met Office and other meteorological agencies that employ Var. circumvent this problem by constructing a model of $\mathbf{B}$ out of physical, statistical and intuitive arguments. The first simplification (which we make clear at the outset) is that $\mathbf{B}$ is determined climatologically, *ie*, it contains no synoptic dependencies. (An attempt to account for synoptically dependent errors has been introduced as an add-on to the cost function - see work on the "errors of the day" project [2][3]. Other approaches considered include use of the geostrophic coordinate transform [4][5].)

3.1 Control variable transforms

In the minimization, components of $\vec{x}'$ are varied iteratively to reduce - and hence minimize - the value of J. A vector such as $\vec{x}'$ which is varied in this way is called a control vector. Minimizing J by varying $\vec{x}'$ directly is conceptually straightforward, but it leaves us with the problem of representing and inverting $\mathbf{B}$, and it would actually be a very inefficient approach to Var. Instead of solving the variational problem involving $\vec{x}'$, the problem is posed in terms of a different control vector, $\vec{v}'$ which is designed, as below, to have beneficial properties. As for $\vec{x}'$, $\vec{v}'$ is primed, indicating that it is a perturbation quantity.

The vector $\vec{v}'$ is related to $\vec{x}'$ via the so-called $\mathbf{U}$-transform, transforming from 'control' to 'model' perturbations,

$$\vec{x}' = \mathbf{U}\vec{v}', \tag{2}$$

and its inverse is the so-called $\mathbf{T}$-transform,

$$\vec{v}' = \mathbf{T}\vec{x}'. \tag{3}$$

In order to see how this change of variables can help the problem, substitute Eq. 2 into the cost function, Eq. 1,

$$\begin{aligned} J[\vec{v}'] &= \frac{1}{2}(\mathbf{U}\vec{v}' - \mathbf{U}\vec{v}'_B)^T \mathbf{B}^{-1} (\mathbf{U}\vec{v}' - \mathbf{U}\vec{v}'_B) + \frac{1}{2}(\vec{y} - \vec{H}[\mathbf{U}\vec{v}', \vec{x}_0])^T (\mathbf{E} + \mathbf{F})^{-1} (\vec{y} - \vec{H}[\mathbf{U}\vec{v}', \vec{x}'_0]), \\ &= \frac{1}{2}(\vec{v}' - \vec{v}'_B)^T \mathbf{U}^T \mathbf{B}^{-1} \mathbf{U} (\vec{v}' - \vec{v}'_B) + \frac{1}{2}(y - \vec{H}[\mathbf{U}\vec{v}', \vec{x}_0])^T (\mathbf{E} + \mathbf{F})^{-1} (y - \vec{H}[\mathbf{U}\vec{v}', \vec{x}'_0]), && (4) \\ &= \frac{1}{2}(\vec{v}' - \vec{v}'_B)^T (\vec{v}' - \vec{v}'_B) + \frac{1}{2}(y - \vec{H}[\mathbf{U}\vec{v}', \vec{x}_0])^T (\mathbf{E} + \mathbf{F})^{-1} (y - \vec{H}[\mathbf{U}\vec{v}', \vec{x}'_0]). && (5) \end{aligned}$$

matrix in $\vec{v}$-space is $\mathbf{U}^T\mathbf{B}^{-1}\mathbf{U}$. The strategy is to design the $\mathbf{U}$-transform to make this matrix into the identity,

$$\mathbf{U}^T\mathbf{B}^{-1}\mathbf{U} = \mathbf{I}. \tag{6}$$

This means that in $\vec{v}$ co-ordinates, elements of the background state vector are uncorrelated with all other elements and each has a unit variance. This information has been imposed to give Eq. 5. The cost function in the form of Eq. 5 is the one that is to be minimized with $\vec{v}'$ as the new control vector. As long as Eq. 6 holds, this cost function is equivalent to the original cost function, Eq. 1. Due to the property of Eq. 6, the process of minimizing Eq. 5 is simpler than before as there is now no background error covariance matrix to invert. Once a $\vec{v}'$ has been found that minimizes J, the analysis in model space is,

$$\vec{x}_A = \vec{x}_0 + \mathbf{U}\vec{v}'. \tag{7}$$

Before proposing a general way in which the $\mathbf{U}$-transform can be designed to achieve the special property, we note how this change of variables helps us represent $\mathbf{B}$ and can also, at the same time, improve the conditioning of the variational problem.

3.1.1 How can this change of variables help represent B?

Rearranging the operators in Eq. 6 gives an expression for $\mathbf{B}$ in terms of the transformation operators,

$$\mathbf{B} = \mathbf{U}\mathbf{U}^T, \tag{8}$$

which, incidentally, has the required property of being symmetric. Thus by designing $\mathbf{U}$ to yield a unit background error covariance matrix in the transformed space, the operator $\mathbf{U}$ is found to be the square root of $\mathbf{B}$. The information contained in $\mathbf{B}$ as in the original cost function, Eq. 1, is transferred into the transformation operator, $\mathbf{U}$ in the new cost function, Eq. 5. $\mathbf{B}$ is now no longer explicit, but is, instead, implied by the transformation. The task of defining $\mathbf{B}$ is now one of designing the $\mathbf{U}$-operator.

3.1.2 How does changing the control vector improve the conditioning of the problem?

It is found that the cost function in the $\vec{x}'$-formulation, Eq. 1 is badly conditioned. To show how badly conditioned the original problem is requires knowledge of the properties of the operators in Eq. 1, especially $\mathbf{B}$. The shape of the cost function is specified by the Hessian matrix. This is the matrix of second derivatives of J with respect to elements of the control vector [6]. In $\vec{x}'$-space, the Hessian is $\mathbf{A}_\mathbf{x}$,

$$\mathbf{A}_\mathbf{x} = \mathbf{B}^{-1} + \mathbf{H}^T(\mathbf{E}+\mathbf{F})^{-1}\mathbf{H}, \tag{9}$$

which might describe the shape of the cost function to something like that in Fig. 1 (left panel). The skewed arrows denote the principal axes of J in this space, which are the eigenvectors of $\mathbf{A}_\mathbf{x}$, and the eigenvalues give the curvature of J along these directions. A small curvature, such as along the elongated axis in the Fig. represents a 'mode' of the data assimilation system about which there is little information (these modes have a high variance). Conversely, a large curvature mode, as perpendicular to the elongated axis, represents a mode of the data assimilation system about which there is much information (these have a small variance and so constrain the analysis strongly). In general, and as shown in Fig. 1, the principal axes will not be coincident with the model variables, which means that they are not 'pure' meteorological quantities, but a combination of such quantities.

The degree of skewness is determined by the range of Hessian eigenvalues, such that a broad range gives rise to contours that are highly skewed. As a consequence, the descent towards the minimum of J will be inefficient and prone to numerical error. A measure of the conditioning of J is the conditioning number defined as,

$$\text{conditioning number} = \frac{\text{maximum eigenvalue}}{\text{minimum eigenvalue}}, \tag{10}$$

with that note that all eigenvalues will be positive as the Hessian is positive definite. The conditioning number will be large for badly conditioned problems and unity for well conditioned problems.

It is believed that the term in Eq. 9 responsible for most of the bad conditioning is $\mathbf{B}^{-1}$. Typically $\mathbf{B}^{-1}$, which has the same conditioning as $\mathbf{B}$, is found to have a very large conditioning number of $\sim 10^{10}$ [7]. Since we are transforming the effective background error covariance matrix in $\vec{v}$-space to resemble the unit matrix, the conditioning in that space is expected to be better - see Fig. 1 (right panel). The

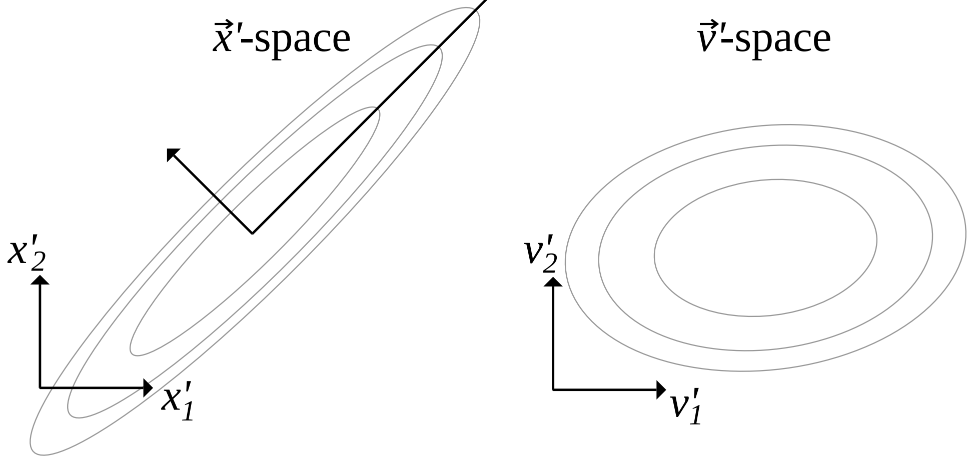

Figure 1: Schematic showing contours of the cost function J for model variables ($\vec{x}'$, left), for which the problem is badly conditioned, and for control variables ($\vec{v}'$, right), for which the conditioning is improved. The axes superimposed on the centre of the ellipses in the left panel denote the principle axes of the Hessian in that case (Eq. 9).

Hessian in $\vec{v}$-space is found, once more, by the matrix of second derivatives of J but now with respect to elements of the new control vector. The second derivative of Eq. 5 with respect to $\vec{v}$ gives $\mathbf{A}_v$,

$$\mathbf{A}_v = \mathbf{I} + \mathbf{U}^T \mathbf{H}^T (\mathbf{E} + \mathbf{F})^{-1} \mathbf{H}\mathbf{U}. \tag{11}$$

Here the shape of J is much less skewed than before (having a relatively small range of eigenvalues of $\mathbf{A}_v$), but not perfectly circular due to the presence of the observation contribution to the Hessian (second term of Eq. 11). The conditioning number of the background term on its own in this space has the ideal value of unity. The benefit of the transformation is that the descent algorithm now acts within the shape of $J[\vec{v}']$ which allows it to work efficiently, and in a manner where similar weight is given to each element of the control vector.

3.1.3 The transforms U, $\mathbf{U}^T$ and T

Apart from $\mathbf{U}$, there is also the adjoint transformation, $\mathbf{U}^T$, as used in Eq. 11, and the inverse $\mathbf{T}$ as defined in Eq. 3. The adjoint operator also appears in the gradient of the cost function with respect to $\vec{v}'$,

$$\left(\frac{\partial J[\vec{v}']}{\partial \vec{v}'} \right)^T = (\vec{v}' - \vec{v}'_B) + \mathbf{U}^T \mathbf{H}^T (\mathbf{E} + \mathbf{F})^{-1} (y - \vec{H}[\mathbf{U}\vec{v}', \vec{x}'_0]), \tag{12}$$

which is needed for the descent algorithm. The $\mathbf{T}$-transform has not yet appeared explicitly in any equations so far, but it is needed in two places in Var. The first is in the (off-line) calibration stage [8] in order to determine the background error covariance statistics (and hence determine the $\mathbf{U}$-transform - see section 4.4), and the second is in the determination of $\vec{v}'_B$ in equations 5 and 12 in those cases when the linearization state, $\vec{x}_0$ is not coincident with the background, $\vec{x}_B$. Then, $\vec{v}'_B = \mathbf{T}(\vec{x}_B - \vec{x}_0)$ (if the linearization state is the same as the background - as is the usual case - then the $\mathbf{T}$-transform is not needed there as $\vec{v}'_B$ is automatically zero).

3.2 A generic transform

Before discussing how the Met Office have designed their control variable transforms, we introduce a method for formulating $\mathbf{U}$ which could be used if unlimited computing resources were available. Although the approach to be described cannot be applied directly to a vector space of the size that is needed for an atmospheric forecast model, it is used, in a partial sense, in the practical solution discussed later.

The first step is to diagonalize $\mathbf{B}$ (this is what prohibits direct use of this approach). Before we do this, $\mathbf{B}$ must be scaled accordingly by choosing an appropriate 'inner product'. The simplest inner product has the form, $\vec{x}^T\vec{x}$, which is inappropriate in $\vec{x}$-space. Reasons for this [12] are that firstly some elements of $\vec{x}$ represent gridboxes larger than other elements and so will contribute disproportionally to terms, etc.). As a rule, whenever an inner product in $\vec{x}$-space is required, $\vec{x}^T\vec{x}$ should be replaced with $\vec{x}^T\mathbf{P}\vec{x}$ where $\mathbf{P}$ is an inner product matrix, which must be symmetric. It will be useful to decompose $\mathbf{P}$ into its square roots,

$$\mathbf{P} = \mathbf{P}^{T/2}\mathbf{P}^{1/2}. \tag{13}$$

Often $\mathbf{P}$ is taken to be a diagonal matrix of grid box volumes or areas [9]. It is not recommended that $\mathbf{B}$ be diagonalized without using a physically meaningful inner product matrix as otherwise its eigenfunctions will have (strictly speaking) no physical meaning.

The eigenvalue problem can now be meaningfully posed as,

$$\begin{aligned}
\mathbf{BPX} &= \mathbf{X}\Lambda, & (14)\\
\mathbf{P}^{1/2}\mathbf{BP}^{T/2}\mathbf{P}^{1/2}\mathbf{X} &= \mathbf{P}^{1/2}\mathbf{X}\Lambda, & (15)\\
\mathbf{B}'\mathbf{F} &= \mathbf{F}\Lambda, & (16)\\
\text{where}\quad \mathbf{F} &= \mathbf{P}^{1/2}\mathbf{X}, & (17)\\
\text{and}\quad \mathbf{B}' &= \mathbf{P}^{1/2}\mathbf{BP}^{T/2}. & (18)
\end{aligned}$$

The second two equations are a development of the first equation and the last two expressions are definitions. In the above, Λ is the diagonal matrix of eigenvalues and the columns of $\mathbf{X}$ are the eigenfunctions of $\mathbf{B}$ within the $\mathbf{P}$ inner-product framework. The matrix $\mathbf{P}$ has been included in Eq. 14 as the action of $\mathbf{B}$ on $\mathbf{X}$ is an inner product in $\vec{x}$-space. Equation 16 is a rewrite of Eq. 14 using alternative $\mathbf{F}$-vectors defined in Eq. 17, which are eigenvectors of the transformed matrix $\mathbf{B}'$ defined in Eq. 18. Transformed objects like $\mathbf{B}'$, which have physical meaning, are sometimes called filters. There is no need for an explicit inner product matrix in Eq. 16 as it has been subsumed into the other matrices.

The background error covariance matrix is symmetric, and so is the matrix $\mathbf{P}^{1/2}\mathbf{BP}^{T/2}$ (Eq. 18). This means that its eigenvectors are orthogonal [10][11],

$$\mathbf{FF}^T = \mathbf{I} \quad \text{and} \quad \mathbf{F}^T\mathbf{F} = \mathbf{I}, \tag{19}$$

and so $\mathbf{F}^T = \mathbf{F}^{-1}$.

By simple matrix manipulation, Eq. 16 can be developed further, and inverted,

$$\Lambda^{1/2}\mathbf{F}^T\mathbf{P}^{-T/2}\mathbf{B}^{-1}\mathbf{P}^{-1/2}\mathbf{F}\Lambda^{1/2} = \mathbf{I}, \tag{20}$$

and then compared to Eq. 6. This gives a general form for the $\mathbf{U}$ transform (Eq. 21), from which its adjoint (22) and inverse (23) follow,

$$\begin{aligned}
\mathbf{U} &= \mathbf{P}^{-1/2}\mathbf{F}\Lambda^{1/2}, & (21)\\
\mathbf{U}^T &= \Lambda^{1/2}\mathbf{F}^T\mathbf{P}^{-T/2}, & (22)\\
\mathbf{T} &= \Lambda^{-1/2}\mathbf{F}^T\mathbf{P}^{1/2}. & (23)
\end{aligned}$$

These results have a simple interpretation by looking at Eq. 23 for the $\mathbf{T}$-transform, which transforms model to control increments (Eq. 3). After applying the square root of the inner product matrix, it projects the result onto the eigenvectors of $\mathbf{P}^{1/2}\mathbf{BP}^{T/2}$. This is a rotation of the basis to the principal axes of $\mathbf{B}$ before scaling by the square root of the inverse eigenvalue for that direction.

This choice of transformation was first pointed out by Parrish and Derber in 1992 [13]. They suggested that if the control vector is chosen to represent the weights of the eigenvectors of $\mathbf{B}$ then the background term could be made diagonal (and hence simple to treat). The extra step of scaling with the square roots of the inverse eigenvalues then help to precondition the problem.

3.3 Other transform options - rotation

The $\mathbf{U}$-transformation, representing the square root of $\mathbf{B}$ as in Eq. 21 is not unique. Alternatives for the square roots have the general form,

$$\mathbf{U} = \mathbf{P}^{-1/2}\mathbf{F}\Lambda^{1/2}\mathbf{R}, \tag{24}$$

where $\mathbf{R}$ is an orthogonal rotation matrix ($\mathbf{RR}^T = \mathbf{I}$). There are an infinite number of possibile alternatives of $\mathbf{R}$ (Eq. 21 has $\mathbf{R} = \mathbf{I}$). Although the particular choice of $\mathbf{R}$ will influence the transform (as in Eq. 24), the background error covariance matrix - which the transformation is meant to represent inserted into Eq. 8). Rotation of the modes in the vertical transform (see section 4.2) has been used successfully to improve the representation of background error covariances in the stratosphere (see [14]). The adjoint and inverse of Eq. 24 are,

$$\mathbf{U}^T = \mathbf{R}^T \Lambda^{1/2} \mathbf{F}^T \mathbf{P}^{-T/2}, \tag{25}$$
$$\mathbf{T} = \mathbf{R}^T \Lambda^{-1/2} \mathbf{F}^T \mathbf{P}^{1/2}. \tag{26}$$

If we choose the rotation to be $\mathbf{R} = \mathbf{F}^T$, where $\mathbf{F}$ is the matrix of eigenvectors found in section 3.2, then this will satisfy the required orthogonality condition $\mathbf{R}\mathbf{R}^T = \mathbf{I}$. The $\mathbf{U}$-transform in Eq. 24 will then become,

$$\mathbf{U} = \mathbf{P}^{-1/2} \mathbf{F} \Lambda^{1/2} \mathbf{F}^T, \tag{27}$$

which is known as the 'symmetric' square root of $\mathbf{B}$ (Eq. 27 is symmetric however only if $\mathbf{P}$ is absent).

Rotation has the consequence that the components of the control vector $\vec{v}'$ no longer represent the eigenvectors contained in the columns of $\mathbf{F}$, but instead by the space represented by the additional transform $\mathbf{R}^T$ as in Eq. 26. In particular, if the choice, $\mathbf{R} = \mathbf{F}^T$ is made (as above) then the control vector will be in the same representation as the model vector, $\vec{x}'$.

Presence of rotation does not compromise the property that the background errors in $\vec{v}'$ should be uncorrelated. In the example of rotation discussed above (that $\mathbf{R} = \mathbf{F}^T$) e.g., it might be thought that, since the control vector, $\vec{v}'$, is expressed in the same basis as the model vector, $\vec{x}'$, and the model's basis members are correlated then components of $\vec{v}'$ should also be correlated. It is easy to show that a rotation satisfying $\mathbf{R}\mathbf{R}^T = \mathbf{I}$ will not disrupt the uncorrelated nature of members of $\vec{v}'$. Given that $\mathbf{B}$ can be defined by the expectation, $\mathbf{B} = \langle \vec{x}' \vec{x}'^T \rangle$, that $\vec{v}' = \mathbf{T}\vec{x}'$, and Eq. 26 holds, then we can ask: does $\langle \vec{v}' \vec{v}'^T \rangle = \mathbf{I}$? (Note that expectation over realization is denoted by the angular brackets, $\langle \rangle$.)

$$\begin{aligned}
\langle \vec{v}' \vec{v}'^T \rangle &= \langle \mathbf{T} \vec{x}' \vec{x}'^T \mathbf{T}^T \rangle, \\
&\approx \mathbf{T} \langle \vec{x}' \vec{x}'^T \rangle \mathbf{T}^T, \\
&= \mathbf{T}\mathbf{B}\mathbf{T}^T, \\
&= \mathbf{R}^T \Lambda^{-1/2} \mathbf{F}^T \mathbf{P}^{1/2} \mathbf{B} \mathbf{P}^{T/2} \mathbf{F} \Lambda^{-1/2} \mathbf{R},
\end{aligned}$$

Use Eq. 18 and 16,

$$\begin{aligned}
\langle \vec{v}' \vec{v}'^T \rangle &= \mathbf{R}^T \Lambda^{-1/2} \Lambda \Lambda^{-1/2} \mathbf{R}, \\
&= \mathbf{I}.
\end{aligned} \tag{28}$$

Although Eq. 24, is a generic and correct form of $\mathbf{U}$, diagonalization of the full $\mathbf{B}$ is not possible, and so an alternative approach is needed (that said, the result 24 remains useful and will be used to diagonalize submatrices of $\mathbf{B}$ - see below). At the time of writing the current transform between model and control vectors is comprised of three stages. These are the parameter transform, the vertical transform and the horizontal transform (given subscripts "p", "v" and "h" respectively). Within this decomposition, $\mathbf{T}$ and $\mathbf{U}$ are,

$$\mathbf{T} = \mathbf{T}_h \mathbf{T}_v \mathbf{T}_p, \tag{29}$$
$$\mathbf{U} = \mathbf{U}_p \mathbf{U}_v \mathbf{U}_h, \tag{30}$$

where the $\mathbf{T}$ and $\mathbf{U}$ components have an inverse relationship (sometimes exact, sometimes approximate),

$$\mathbf{U}_p^{-1} = \mathbf{T}_p, \ \mathbf{U}_v^{-1} = \mathbf{T}_v \text{ and } \mathbf{U}_h^{-1} = \mathbf{T}_h, \tag{31}$$

We will proceed by explaining the role of each transform in the order $\mathbf{T}_p$, $\mathbf{T}_v$ and $\mathbf{T}_h$. Sometimes it is easier to explain the workings of each transform by its $\mathbf{T}$ component, other times with its $\mathbf{U}$ counterpart.

4.1 The parameter transform

The first stage of the $\mathbf{T}$-transform is to convert from model variables to new parameters which have mutually uncorrelated errors. As long as parameters are derived with truly uncorrelated errors, this step is essentially a block diagonalization of the problem, leaving only spatial covariances within each parameter's field.

The 'uncorrelated' parameters are taken to be streamfunction, ψ, velocity potential, χ, unbalanced pressure, ${}^{A}P$ and relative humidity, μ, as detailed in [15][16][6]. Figure 2 is a schematic of $\mathbf{B}_p$, the error covariance matrix imposed between the parameters. The role of $\mathbf{T}_p$ then is to break the big multivariate error covariance problem down into many simpler univariate problems, each of which can be treated further by the vertical and horizontal steps.

The basis for this choice of parameters is the assumption that their errors are uncorrelated between parameters. The implied error covariances of the model fields will be $\mathbf{U}_p \mathbf{B}_p \mathbf{U}_p^T$, but since the non-correlation assumption is poor, the impled covariances will not necessarily be a good representation of the true error covariances between model field errors. On practical grounds though the parameter transform is relatively efficient. There is currently much discussion on alternative control parameters that are better approximations to the problem (see [6][17][18][19]).

	ψ	χ	${}^{A}p$	μ
ψ	$\mathbf{B}_{\psi\psi}$	0	0	0
χ	0	$\mathbf{B}_{\chi\chi}$	0	0
${}^{A}p$	0	0	$\mathbf{B}_{{}^{A}p{}^{A}p}$	0
μ	0	0	0	$\mathbf{B}_{\mu\mu}$

Figure 2: The error covariance matrix between parameters, $\mathbf{B}_p$. Parameters, as output by $\mathbf{T}_p$ are designed to be approximately uncorrelated. Each block is a submatrix of spatial covariances, and the off diagonal matrices are set to zero as part of the covariance model.

The remaining vertical and horizontal transforms together are designed to project each parameter's field onto sets of orthogonal and uncorrelated modes, with variance scaling, e.g. as in Eq. 26. In the decomposition of Eqs. 29 and 30 - where the vertical transfom is meant to represent the vertical error covariances - the projection is rather like doing a fourier decomposition in the vertical direction, but instead of projecting onto plane waves, the projection is onto empirical orthogonal functions (EOFs), found from vertical error covariances matrices. Then, instead of thinking of error variances at different vertical levels and covariances between levels, the vertical covariance model is built of variances of these (uncorrelated) EOFs. Since there are no covariances in the vertical between these modes (or only weak covariances where approximations), this step is a simplification of the representation of $\mathbf{B}$.

Let the vertical background covariance matrix at position (λ, ϕ) for errors in one of the parameters discussed in section 4.1 be $\mathbf{B}_v(\lambda, \phi)$. We explain the vertical transform with reference to $\mathbf{T}_v$. This is performed at each horizontal position and for each parameter. For a given parameter, e.g. ψ, $\mathbf{T}_v$ yields (at horizontal position (λ, ϕ)),

$$\vec{\psi}(\lambda, \phi, \nu) = \mathbf{T}_v(\lambda, \phi)\vec{\psi}(\lambda, \phi, z), \tag{32}$$

where the arguments (λ, ϕ, z) and (λ, ϕ, ν) have been added to the state vectors to show explicitly the three dimensional space in which the components of the vectors exist. $\mathbf{T}_v$ will be different for different parameters, but we drop an explicit parameter label to simplify notation. z is the original vertical coordinate and ν is a new 'vertical' coordinate. Actually, the nature of ν depends upon the choice of the rotation (see below), but two possibilities are discussed here, these being either $\nu = z$, the original vertical co-ordinate or a new 'EOF index' (the EOFs being the eigenvectors of $\mathbf{B}_v(\lambda, \phi)$, weighted as in Eq. 16). The latter representation denotes the projection onto the νth EOF - see Fig. 3, left panel). Whatever the choice of rotation, projection onto vertical EOFs will be part of the vertical transform, and a selection of such vertical modes is shown in Fig. 4.

Since we are aiming towards new variables that are uncorrelated and have unit variance in each vertical column, we *could* choose $\mathbf{T}_v$ to have the form that is based on Eq. 26,

$$\mathbf{T}_v(\lambda, \phi) = \mathbf{R}_v^T(\lambda, \phi)\ \Lambda_v(\lambda, \phi)^{-1/2}\ \mathbf{F}_v(\lambda, \phi)^T\ \mathbf{P}_v^{1/2}, \tag{33}$$

where $\mathbf{R}_v$ is any vertical rotation (satisfying $\mathbf{R}_v\mathbf{R}_v^T = \mathbf{I}$), $\mathbf{F}_v$ are the vertical EOFs (eigenvectors of $\mathbf{B}_v$) and Λ_v are their variances. The transform in Eq. 33 acts on one parameter within one column, positioned at (λ, ϕ) and similar transforms need to be made for all other positions and parameters.

4.2.1 Difficulties with the vertical modes

The EOFs, $\mathbf{F}_v(\lambda, \phi)$ and their variances, $\Lambda_v(\lambda, \phi)$, will vary with horizontal position. For some choices of $\mathbf{R}_v$ (e.g. $\mathbf{R}_v = \mathbf{I}$ the $\mathbf{T}_v$-transform will project columns of parameters onto vertical modes that are a function of latitude and longitude. As a consequence of this variation with horizontal position, there is potential for loss of continuity of the fields on the resulting EOF surfaces (see Fig. 3, right panel). This is a difficulty especially if the order of the EOFs changes from one position to the next, so that if one were to plot the surface $\vec{\psi}(\lambda, \phi, \nu = n)$, for n constant, then one might be plotting the scaled weight of one mode at one position, but of another mode at the next.

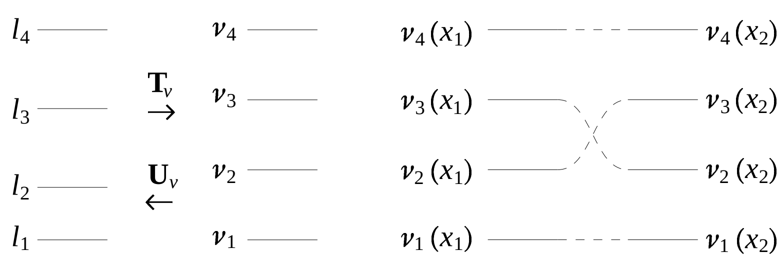

Figure 3: Left panel: in the $\mathbf{T}$-transform, the vertical transform acts on a column and projects onto a set of EOF modes. The $\mathbf{U}$-transform, is designed to be the inverse. Right panel: if the EOF modes are allowed to be a function of position $x_n = (\lambda_n, \phi_n)$ the ordering of the modes - ordered by their eigenvalues (variances) - may change. The fields in EOF space would then suffer a loss of continuity which is why global modes are used in the current transform (see text).

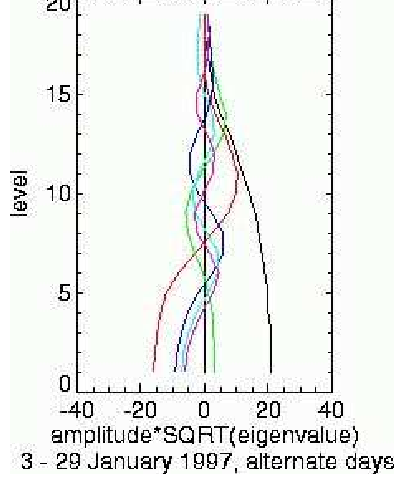

Figure 4: Example six leading vertical EOFs for the global mean vertical background error covariance matrix for unbalanced pressure, ${}^{A}p$. Figure taken from [7].

4.2.2 One solution to this problem ($\nu = z$)

A way round this problem is to use the freedom of rotation to project back onto the original vertical co-ordinate, $\nu = z$. Choosing $\mathbf{R}_v^T = \mathbf{F}_v(\lambda, \phi)$ in Eq. 33 will result in the vertical transform from every vertical column to yield an output in the model's height space. This avoids loss of continuity. This solution leads to the 'symmetric square-root' transformation for the vertical error covariances and results in the following transforms based on Eqs. 24 to 26,

$$\mathbf{U}_v = \mathbf{P}_v^{-1/2}\mathbf{F}_v\Lambda_v^{1/2}\mathbf{F}_v^T, \tag{34}$$
$$\mathbf{U}_v^T = \mathbf{F}_v\Lambda_v^{1/2}\mathbf{F}_v^T\mathbf{P}_v^{-T/2}, \tag{35}$$
$$\mathbf{T}_v = \mathbf{F}_v\Lambda_v^{-1/2}\mathbf{F}_v^T\mathbf{P}_v^{1/2}. \tag{36}$$

Although this rotation solves the continuity problem, it lacks certain features required (see section 5 for a discussion). These may be added later however by using a wavelet or waveband solution, introduced in section 6.

4.2.3 Another solution ($\nu =$ global EOF index)

Another solution, which is the one currently adopted by the Met Office, is to scrap the position dependence of the EOFs and replace them with a single set of EOFs used at every horizontal position. In this case the vertical coordinate, ν, output by the $\mathbf{T}_v$ transform is EOF index. This set of EOFs is taken to be the eigenvectors of the globally averaged vertical covariance matrix. This step is made with two costs: (i) there is no variation of vertical mode with position (the eigenvalues can change though - see Eq. 37 below) and (ii) the global modes are no longer exactly uncorrelated locally, having a covariance matrix, $\hat{\Lambda}_v(\lambda, \phi)$ at position (λ, ϕ). Let the global EOFs be columns of the matrix $\mathbf{F}_v$; $\hat{\Lambda}_v(\lambda, \phi)$ is then found by projecting $\mathbf{B}_v(\lambda, \phi)$ onto $\mathbf{F}_v$,

$$\mathbf{F}_v^T\mathbf{P}_v^{1/2}\mathbf{B}_v(\lambda, \phi)\mathbf{P}_v^{T/2}\mathbf{F}_v = \hat{\Lambda}_v(\lambda, \phi), \tag{37}$$

where the matrix $\hat{\Lambda}_v(\lambda, \phi)$ is expected to be nearly diagonal.

Instead of basing their vertical transform directly on Eq. 33 for this solution to the continuity problem, the Met Office use a slightly different form which we now try to build by arguments that start with Eq. 37. With the right hand side made exactly diagonal, Eq. 37 becomes,

$$\mathbf{F}_v^T\mathbf{P}_v^{1/2}\mathbf{B}_v(\lambda, \phi)\mathbf{P}_v^{T/2}\mathbf{F}_v \approx \text{diag}\left(\hat{\Lambda}_v(\lambda, \phi)\right), \tag{38}$$

where the 'diag' operator strips the off diagonal elements, leaving the diagonal elements unaffected.

On the right hand side, multiply by $\Lambda^{1/2}\mathbf{R}_v\mathbf{R}_v^T\Lambda^{-1/2}$ to the left, and by its adjoint, $\Lambda^{-1/2}\mathbf{R}_v\mathbf{R}_v^T\Lambda^{1/2}$ to the right. Here Λ is the diagonal matrix of eigenvalues of the globally averaged vertical covariance these strings of matrices is equal to the identity matrix and so neither will alter the expression in any way (recall that the rotations satisfy $\mathbf{R}_v \mathbf{R}_v^T = \mathrm{I}$). In addition, take the matrices of global EOFs (F_v terms in Eq. 37) to the right hand side,

$$\mathbf{P}_v^{1/2} \mathbf{B}_v(\lambda, \phi) \mathbf{P}_v^{T/2} \approx \mathbf{F}_v \Lambda^{1/2} \mathbf{R}_v \mathbf{R}_v^T \Lambda^{-1/2} \mathrm{diag}\left(\hat{\Lambda}_v(\lambda, \phi)\right) \Lambda^{-1/2} \mathbf{R}_v \mathbf{R}_v^T \Lambda^{1/2} \mathbf{F}_v^T. \tag{39}$$

Since $\Lambda^{\pm 1/2}$ is diagonal, the matrices of eigenvalues adjacent to the 'diag' operator can be taken inside. The adjacent $\mathbf{R}_v^T$ and $\mathbf{R}_v$ matrices are also taken inside, but this move is an approximation,

$$\mathbf{P}_v^{1/2} \mathbf{B}_v(\lambda, \phi) \mathbf{P}_v^{T/2} \approx \mathbf{F}_v \Lambda^{1/2} \mathbf{R}_v \ \mathrm{diag}\left(\mathbf{R}_v^T \Lambda^{-1/2} \hat{\Lambda}_v(\lambda, \phi) \Lambda^{-1/2} \mathbf{R}_v\right) \ \mathbf{R}_v^T \Lambda^{1/2} \mathbf{F}_v^T. \tag{40}$$

We have now made all the steps and approximations necessary to give Eq. 1.21 of the official documentation [20] (with some sign errors presumably corrected). Equation 40 can be rewritten by parametrizing the right hand side,

$$\mathbf{P}_v^{1/2} \mathbf{B}_v(\lambda, \phi) \mathbf{P}_v^{T/2} \approx \mathbf{F}_v \Lambda^{1/2} \mathbf{R}_v \mathbf{M}(\lambda, \phi) \mathbf{R}_v^T \Lambda^{1/2} \mathbf{F}_v^T, \tag{41}$$

where $\mathbf{M}(\lambda, \phi)$, as in [20] is,

$$\mathbf{M}(\lambda, \phi) = \mathrm{diag}(\mathbf{Z} \mathbf{B}_v(\lambda, \phi) \mathbf{Z}^T), \tag{42}$$

and, making use of Eq. 37,

$$\mathbf{Z} = \mathbf{R}_v^T \Lambda^{-1/2} \mathbf{F}_v^T \mathbf{P}_v^{1/2}. \tag{43}$$

Taking the inner products of Eq. 41 to the right hand side, allows $\mathbf{B}_v(\lambda, \phi)$ to be written,

$$\begin{aligned} \mathbf{B}_v(\lambda, \phi) &\approx \mathbf{P}_v^{-1/2} \mathbf{F}_v \Lambda^{1/2} \mathbf{R}_v \mathbf{M}(\lambda, \phi) \mathbf{R}_v^T \Lambda^{1/2} \mathbf{F}_v^T \mathbf{P}_v^{-T/2}, \\ &\approx \mathbf{Z}^{-1} \mathbf{M}(\lambda, \phi) \mathbf{Z}^{-T}. \end{aligned} \tag{44}$$

By using analogous arguments to those used in section 3 (that the that the effective covariance matrix in the vertical is to become the identity matrix), the vertical transforms can be formed by taking the square root of $\mathbf{B}_v(\lambda, \phi)$ as $\mathbf{U}_v \mathbf{U}_v^T$. Comparing Eq. 44 with $\mathbf{B}_v = \mathbf{U}_v(\lambda, \phi) \mathbf{U}_v^T(\lambda, \phi)$ (c.f. Eq. 8) yields the $\mathbf{U}_v$, $\mathbf{U}_v^T$ and $\mathbf{T}_v$ transforms for this vertical transform,

$$\mathbf{U}_v(\lambda, \phi) = \mathbf{Z}^{-1} \mathbf{M}(\lambda, \phi)^{1/2}, \tag{45}$$

$$\mathbf{U}_v(\lambda, \phi)^T = \mathbf{M}(\lambda, \phi)^{T/2} \mathbf{Z}^{-T}, \tag{46}$$

$$\mathbf{T}_v(\lambda, \phi) = \mathbf{M}(\lambda, \phi)^{-1/2} \mathbf{Z}. \tag{47}$$

The result of Eq. 47 is different to that proposed before as Eq. 33, which why we had to take the above route to demonstrate the legitimacy of the transforms used in the current scheme.

4.2.4 Note about local and global vertical EOFs

As it stands, we have two forms of the vertical transforms:

- Equations 34 to 36 have been designed with a rotation in mind that projects the vertical EOFs back onto the height levels. This is desirable as the vertical EOFs can be a function of horizontal position. The use of height levels is undesirable from another point of view however (see section 5),

- Eqs. 45 to 47 (the standard forms of the transforms) have been designed to input or output on global EOF levels. There is thus no dependence on horizontal position of the vertical modes (only of their variances). If a rotation matrix is used here, it must transform from the EOF modes common to every horizontal position. We have not considered any such rotation matrix in this report. It is difficult to see how a rotation matrix in these transforms can introduce horizontal position dependence, which is not very useful. This is why the (simpler) first set of transforms has been added to the discussion.

In the $\mathbf{T}$-transform sense, after acting with $\mathbf{T}_p$ and then with $\mathbf{T}_v(\lambda, \phi)$ at each horizontal location results in fields of parameter perturbations, each a function of λ, ϕ and ν, where the nature of the vertical coordinate, ν, depends on the choice of rotation used in the vertical transform (section 4.2). These fields are meant to have no correlations between parameters and no correlations between ν levels. The remaining transform is the horizontal transform which is designed to produce variables that are uncorrelated in the horizontal.

The methodology of the horizontal transform [9], in the details, is different from the general one given in section 3.2, and from the one used for the vertical transform. Instead of basing our approach directly on representing an error covariance matrix - which has the property of being symmetric - the transform first represents a different kind of object that is self adjoint (the difference between the adjoint and the transpose is outlined in appendix 1). The new object that will be modelled is the horizontal filter $\mathbf{S}_h$, which is explained in the next paragraph,

$$\begin{aligned} \mathbf{S}_h &= \mathbf{P}_h^{-1/2}\mathbf{B'}_h\mathbf{P}_h^{1/2}, && (48) \\ \text{where} \quad \mathbf{B'}_h &= \mathbf{P}_h^{1/2}\mathbf{B}_h\mathbf{P}_h^{T/2}, && (49) \end{aligned}$$

and where $\mathbf{P}_h$ is the inner product used in the horizontal transform - it is actually taken as a diagonal matrix of horizontal grid box areas (smaller values for elements that are near the poles than those near the equator). Although it is written as one operator in Eq. 30, the horizontal tranform is made up of separate transforms for each ν output by the vertical step. For the purpose of describing the horizontal transform, we think of each ν in isolation.

In Eq. 49, $\mathbf{B'}_h$ is the horizontal error covariance matrix transformed to have meaningful eigenvalues (see section 3.2, in particular Eq. 18). This is a representation of the usual horizontal background error covariance matrix, $\mathbf{B}_h$, and exists in the space transformed by multiplying by $\mathbf{P}_h^{1/2}$ (see Eq. 17). $\mathbf{B'}_h$ is symmetric. The corresponding filter in the model's horizontal space is Eq. 48 and is found by first transforming to the 'physical meaningful space' with $\mathbf{P}_h^{1/2}$, acting with $\mathbf{B'}_h$ and then transforming back with $\mathbf{P}_h^{-1/2}$ [21].

The filter $\mathbf{S}_h$ is not symmetric in the model's space, but is self adjoint. This can be shown by taking the adjoint (*) of $\mathbf{S}_h$ and writing this using the result of appendix 1 (Eq. 130 with $\mathbf{U} \to \mathbf{S}_h$, and $\mathbf{P}_v, \mathbf{P}_\mathbf{x} \to \mathbf{P}_h$),

$$\begin{aligned} \mathbf{S}_h^* &= \mathbf{P}_h^{-1}\mathbf{S}_h^T\mathbf{P}_h, \\ &= \mathbf{P}_h^{-1}\mathbf{P}_h^{T/2}\mathbf{B'}_h\mathbf{P}_h^{-T/2}\mathbf{P}_h, \\ &= \mathbf{P}_h^{-1/2}\mathbf{P}_h^{-T/2}\mathbf{P}_h^{T/2}\mathbf{B'}_h\mathbf{P}_h^{-T/2}\mathbf{P}_h^{T/2}\mathbf{P}_h^{1/2}, \\ &= \mathbf{P}_h^{-1/2}\mathbf{B'}_h\mathbf{P}_h^{1/2}, \\ &= \mathbf{S}_h. && (50) \end{aligned}$$

In the second line, we have substituted $\mathbf{S}_h$ from Eq. 48, and in the third line we have used $\mathbf{P}_h = \mathbf{P}_h^{T/2}\mathbf{P}_h^{1/2}$.

We proceed by constructing a model of $\mathbf{S}_h$ which has this self adjoint property. If the horizontal errors have a structure that is isotropic and homogeneous, then the filter is diagonal in spectral space (see appendix 2),

$$\begin{aligned} \mathbf{F}_h^{-1}\mathbf{S}_h\mathbf{F}_h &= \Lambda_h, \\ \text{ie} \quad \mathbf{S}_h &= \mathbf{F}_h\Lambda_h\mathbf{F}_h^{-1}, && (51) \end{aligned}$$

where the spectral transform, $\mathbf{F}_h$ takes a field from its spectral to its horizontal grid representation and Λ_h is a diagonal matrix, the diagonal elements being the damping required for each horizontal mode. Using the language of section 3.2 of this paper, and thinking of these objects as matrices, the columns of $\mathbf{F}_h$ are akin to the independent modes (e.g. spherical harmonics on the globe [22] or Fourier modes for a limited area model) and the Λ_h (making up a diagonal matrix) are akin to the variances of these modes. (Reading along the diagonal of Λ_h, we obtain a wavenumber dependent quantity. Later we will refer to the diagonal elements of Λ_h as correlation spectra.)

Evidently, the spectral transform has the property that the inverse is the adjoint, $\mathbf{F}_h^{-1} = \mathbf{F}_h^*$ [9]. This, and the assumption that the inner product matrix in wave space is the identity (making $\Lambda_h^* = \Lambda_h^T$),

$$
\begin{aligned}
\mathbf{S}_h &= (\mathbf{F}_h\Lambda_h^{1/2})(\mathbf{F}_h\Lambda_h^{1/2})^*, &(52)\\
&= (\mathbf{F}_h\Lambda_h^{1/2})(\mathbf{F}_h\Lambda_h^{1/2})^T\mathbf{P}_h. &(53)
\end{aligned}
$$

In the last line, we have applied the result relating the adjoint to the transpose (Eq. 130 with $\mathbf{U} \to \mathbf{F}_h\Lambda_h^{1/2}$, $\mathbf{P}_{\mathbf{x}} \to \mathbf{P}_h$ and $\mathbf{P}_v \to \mathbf{I}$). The horizonal background error covariance matrix then follows from Eqs. 48 and 49. This gives the first line of the follwing, and with Eq. 53 to give the second line,

$$
\begin{aligned}
\mathbf{B}_h &= \mathbf{S}_h\mathbf{P}_h^{-1},\\
&= (\mathbf{F}_h\Lambda_h^{1/2})(\mathbf{F}_h\Lambda_h^{1/2})^T. &(54)
\end{aligned}
$$

The horizontal transforms required to decorrelate the horizontal modes, $\mathbf{U}_h$, $\mathbf{U}_h^T$ and $\mathbf{T}_h$ then follow using the horizontal counterpart of Eq. 8,

$$
\begin{aligned}
\mathbf{U}_h &= \mathbf{F}_h\Lambda_h^{1/2}, &(55)\\
\mathbf{U}_h^T &= \Lambda_h^{1/2}\mathbf{F}_h^T,\\
&= \Lambda_h^{1/2}\mathbf{F}_h^{-1}\mathbf{P}_h^{-1}, &(56)\\
\text{and}\quad \mathbf{T}_h &= \Lambda_h^{-1/2}\mathbf{F}_h^{-1}, &(57)
\end{aligned}
$$

where Eq. 56 follows, as before, from the application of the relationship between the adjoint and transpose operators, Eq. 130, and $\mathbf{F}_h^* = \mathbf{F}_h^{-1}$.

There are a number of options for modelling the variances of the spectral modes, $\Lambda_h^{1/2}$ [9]. One choice is the second order auto regressive (SOAR) function. In real and spectral spaces, this has the form,

$$
\begin{aligned}
\mu(r) &= \left(1+\frac{r}{s}\right)\exp\left(-\frac{r}{s}\right), &(58)\\
\Lambda_h^{1/2}(\vec{k}) &= \left(\frac{16\pi s^2}{N_iN_j\delta_i\delta_j}\right)^{1/2}\left(1+\frac{s^2}{4}\left[\left(\frac{\pi k}{N_i\delta_i}\right)^2+\left(\frac{\pi l}{N_j\delta_j}\right)^2\right]\right)^{-5/2}, &(59)
\end{aligned}
$$

where μ is the correlation between two points separated by a distance r, s is a correlation length, $\vec{k}$ is the horizontal wavenumber vector, $\vec{k} = (\pi k/N_i\delta_i, \pi l/N_j\delta_j)$ for k and l integers. N_i and N_j are the number of gridpoints in the longitude and latitude directions respectively and δ_i and δ_j are the respective grid sizes. Another choice is the Gaussian form,

$$
\begin{aligned}
\mu(r) &= \exp\left(-\frac{1}{2}\left(\frac{r}{2s}\right)^2\right), &(60)\\
\Lambda_h^{1/2}(\vec{k}) &= \left(\frac{32\pi s^2}{N_iN_j\delta_i\delta_j}\right)^{1/2}\exp\left(-s^2\left[\left(\frac{\pi k}{N_i\delta_i}\right)^2+\left(\frac{\pi l}{N_j\delta_j}\right)^2\right]\right). &(61)
\end{aligned}
$$

Note that the two-dimensional fields for each vertical coordinate can have a different correlation length.

A plausibility argument as to why the above approach should model error structures that are isotropic and homogeneous, and why the fourier transform of the correlation function should give the wave damping is discussed in appendix 2.

4.4 Summary of current scheme

The parameter, vertical and horizontal transforms are often most easily explained in the **T**-transform sence, which takes perturbations in 'model' space (defined by the usual meterological variables, u, v, θ, etc., each as a function of λ, ϕ and z) to produce perturbations in 'control'-space, which is the state vector manipulated inside Var. The variable changes are designed to precondition the background error covariance matrix, allowing Var. to be a practicable and efficient procedure. (The parameter transform is also designed to control the degree of 'balance' in the analysis, but these related issues are raised elsewhere, e.g. [1], [6].)

Errors in control variables are uncorrelated and have unit variance (at least approximately, given the approximate nature of some of the transforms), allowing the effective background error covariance matrix for the control variables to be the identity matrix.

Errors in the parameters are assumed to be uncorrelated, but there remain spatial correlations (autocorrelations) within the parameters.

- Vertical error covariances within each parameter are dealt with by the vertical **T**-transform. The output of this transform is the set of rescaled parameters, each a function of (λ,ϕ,ν), ν being the resulting vertical coordinate, whose nature depends on the choice of rotation in the vertical transform. In $\vec{v}$-space, parameters are uncorrelated in the vertical.
- Horizontal error covariances within each parameter and for each ν are dealt with by the horizontal **T**-transform. This assumes that errors in spectral (or Fourier) modes are uncorrelated and the variances are a function of total wavenumber only. This implies that such horizontal errors are homogeneous and isotropic.
- The **T**-transform helps explain how the uncorrelated modes are found, but is used only:
  - in the **B**-calibration stage, performed off-line before Var. is run (see section 4.5 below).
  - to construct the background state perturbation in control space at the beginning of Var. in the case that the reference state - from which perturbations are measured - is not the same as the background state itself. In practice this is relevent only when Var. is run with many outer loops, as is coded in the current Met Office Var. algorithm, but actually can be avoided althogether (see e.g. [24]).
- The parameter, vertical and horizontal **U**-transforms (inverses to the **T**-transforms) and their transposes, $\mathbf{U}^T$ are used many times during the Var. iterations to compute gradients of the cost function with respect to the control variables.

4.5 Calibrating the transforms by calculating covariance statistics

The design of the parameter, vertical and horizontal transforms has now been completed. The remaining part is to determine the specific matrices and variances to be used, i.e. $\mathbf{B}_v$, Λ and $\mathbf{F}_v$ in the current vertical transform with no rotation, $\mathbf{F}_v$ and Λ_v in the simpler vertical transform with rotation, and Λ_h in the horizontal transform (the parameter transform is already completely determined). We now derive a means of calibrating these in a systematic way.

The error covariance matrix is found from the following outer product averaged over realizations,

$$\begin{aligned} \mathbf{B} &= \langle \vec{x}' \vec{x}'^T \rangle, & (62) \\ \text{where } \vec{x}' &= \vec{x} - \vec{x}^t, & (63) \end{aligned}$$

and where $\vec{x}^t$ is the 'truth'. We never know the truth, nor can we average over all realizations, and so we settle for the following approximation to replace Eq. 63,

$$\vec{x}' \approx \vec{x}^{48} - \vec{x}^{24}, \tag{64}$$

where $\vec{x}^{48}$ and $\vec{x}^{24}$ are, respectively, 48 and 24 hour forecasts valid at the same time. This is known as the "NMC method" [25] in which the average over realizations is simulated by an average over situations (actually giving twice the climalological covariances). Thus the desired (and synoptically dependent) error covariance matrix in Eq. 62 using Eq. 63 is replaced with an approximate and climatological version using Eq. 64. (Actually many different NMC runs are made at different times of the year and Var. will interpolate between them for the current time of the year. Also a scaling factor is required in Eq. 62 to take into account the use of 48-24 hour differences applied to the uncertainty in a background state found from a short (6 or 12 hour) forecast.) The NMC method is not considered to be the best way of estimating the background error statistics. Other ways are used by other agencies, including averaging over ensembles of assimilation-plus-forecast runs, where each member uses similar, but perturbed observations [26].

Whichever method is used, covariance calibration is done in stages, starting with the vertical transform, followed by the horizontal transform, and is performed for each parameter output by the paramater **T**-transform (ψ, χ, ${}^{A}p$, etc.).

We introduce two new operators: let the operator $\mathbf{T}_p^{\psi}$ be parameter $\mathbf{T}$ transform that yields only one parameter (e.g. ψ in this case) instead of all of them, and let $\mathbf{C}(\vec{r})$ act on this parameter's field to yield a single vertical column at position $\vec{r}$. The vertical error covariance matrix for this parameter at this horizontal position is then,

$$\mathbf{B}_v(\lambda, \phi) = \langle \mathbf{C}(\vec{r})\mathbf{T}_p^{\psi}\vec{x}'(\mathbf{C}(\vec{r})\mathbf{T}_p^{\psi}\vec{x}')^T \rangle, \tag{65}$$

where $\vec{x}'$ is as Eq. 64. The matrix $\mathbf{B}_v(\lambda, \phi)$ is the vertical covariance matrix (No. of levels $\times$ No. of levels) at position (λ, ϕ).

The information needed for the vertical transform depends on the choice of vertical rotation, $\mathbf{R}_v$, used. The two standard cases discussed in section 4.2 are considered.

- If we choose to rotate from the vertical EOFs back to the model's height domain ($\nu = z$), then we need to determine $\mathbf{F}_v(\lambda, \phi)$ and $\Lambda_v(\lambda, \phi)$ in Eqs. 34 to 36. These are just the eigenvectors and eigenvalues (respectively) of each local vertical covariance matrix (for each parameter).

- If we choose no rotation, $\mathbf{R}_v = \mathbf{I}$ then we work with states projected onto EOFs of the globally averaged vertical matrices (ν = global EOF index). For this option we require these EOFs, $\mathbf{F}_v$ and their eigenvalues, Λ_v for each parameter. The local covariance matrices, $\mathbf{B}_v(\lambda, \phi)$ (actually averaged zonally and within small (5 °) latitude bands) and the gobal eigenvectors/eigenvalues then define the vertical transform used in this case (Eqs. 45 to 47).

4.5.2 Calibrating the horizontal transform

The equation needed to develop the methodology for calibrating the horizontal transforms is now developed. Start with the definition of $\mathbf{B}$ (Eq. 62) and substitute for $\vec{x}'$ using the $\mathbf{U}$-transform (Eqs. 2 and 30),

$$\langle \vec{x}'\vec{x}'^T \rangle = \langle \mathbf{U}_p\mathbf{U}_v\mathbf{U}_h\vec{v}'\vec{v}'^T\mathbf{U}_h^T\mathbf{U}_v^T\mathbf{U}_p^T \rangle. \tag{66}$$

Use the fact that the transforms are independent of situation (i.e. can take the averaging brackets throught the operators) and that control variables are uncorrelated, $\langle \vec{v}'\vec{v}'^T \rangle = \mathbf{I}$. Also take the parameter transform and the vertical transform, which is known at this stage, to the left hand side,

$$\begin{aligned}
\langle \mathbf{T}_v\mathbf{T}_p^{\psi}\vec{x}'(\mathbf{T}_v\mathbf{T}_p^{\psi}\vec{x}')^T \rangle &= \mathbf{U}_h\mathbf{U}_h^T, && (67) \\
&= \mathbf{F}_h\Lambda_h^{1/2}\Lambda_h^{1/2}\mathbf{F}_h^T. && (68)
\end{aligned}$$

Since there is no covariance between parameters, the above has been written in terms of a single parameter (e.g. ψ) to simplify matters, and, in the last line, the form of the horizontal transform has been substituted from Eq. 55. The spectral transform, $\mathbf{F}_h$ is fixed and so it can be taken to the left hand side. This leaves,

$$\langle \mathbf{F}_h^{-1}\mathbf{T}_v\mathbf{T}_p^{\psi}\vec{x}'(\mathbf{F}_h^{-1}\mathbf{T}_v\mathbf{T}_p^{\psi}\vec{x}')^T \rangle = \Lambda_h. \tag{69}$$

Since we are assuming that the horizontal modes are uncorrelated, the right hand side is a diagonal matrix of variances of the horizontal modes, known as correlation spectra (the left hand side, if calculated in full will not be exactly diagonal, but we are interested in the diagonal elements only). Due to the independence of vertical modes, the left hand side is calculated on a level-by-level basis, the level, ν, being the 'vertical' coordinate used in the control variable (z or global EOF mode, depending upon the choice of rotation used in the vertical transform).

There is a different horizontal correlation spectrum for each vertical mode (and each parameter). These spectra are not actually used directly in the horizontal transform. Instead, characteristic lengthscales for every vertical mode are derived from Eq. 69, and are used to fit functional forms, e.g. second order auto regressive (SOAR) functions as mentioned in section 4.3 (in practice the lengthscales are modified slightly to account for the erroneously large horizontal length scales derived from the NMC method). This is essentially a parametrization of the horizontal transform where the calibration stage helps determine the parameters.

In summary the calibration process is a step-by-step process requiring many forecast differences, and uses only the $\mathbf{T}$-transform. A difficulty is knowing the initial conditions for the 24 and 48 forecasts needed for the NMC method. Since these are forecasts starting from analyses, data assimilation is required, and the data assimilation process requires the background error covariance matrix, which we are trying to find! Hence some 'boot-strapping' is required. A best guess of the background error covariance matrix is used at first with the expectation that a realistic background error covariance matrix is converged upon by repeating the process.

The form of the current transform (Eq. 30) has limited ability to represent realistically the square root of $\mathbf{B}$. In order to outline some features that are desirable for a covariance model to capture, we highlight in this section some properties of 'observed' covariances. We discuss only those features which we believe can be improved using the proposed waveband summation transform (section 6) and take results from [27], which is a more exhaustive survey. This is followed by a list of shortcomings of the current transforms.

5.1 The structure of background errors

Ingleby [27] has used the NMC method to look at covariances of forecast differences between fields (these are the background error covariances in Var.). We extract a number of relevent points below.

5.1.1 Vertical scales, horizontal scales and separability

Analysis of correlations between fields shows that modes with short (long) horizontal length scales are associated with short (long) vertical length scales (for streamfunction, see Fig. 2a of [27], which is reproduced in Fig. 7a later). Further evidence for this kind of correspondence between vertical and horizontal scales is highlighted by two additional features mentioned in [28].

- The horizontal scales for correlations in differences of temperature, T, are shorter than those for geopotential, Φ (see Fig. 1 and table 1 of [27]; the table is reproduced below in Fig. 5a). If Δ_h^T and Δ_h^Φ are such horizontal length scales for T and Φ respectively then mathematicaly,

  $$\Delta_h^T < \Delta_h^\Phi.$$

  T and Φ are related largely via hydrostatic balance. Using the equation of state the hydrostatic relation is,

  $$\frac{\mathrm{d}\Phi}{\mathrm{d}\ln p} = -\frac{R}{g}T, \tag{70}$$

  where p is pressure, R is the specific gas constant, and g is the acceleration due to gravity. The hydrostatic relation implies that the vertical (subscript v) scales will follow,

  $$\Delta_v^T < \Delta_v^\Phi,$$

  because the derivative of Φ on the left hand side of Eq. 70 weights shorter vertical scales more than larger ones. The above two inequalities are consistent if short horizontal and vertical scales are associated. This is evidence that short (long) horizontal scales are associated with short (long) vertical scales.

- Similar arguments are used to demonstrate non-separability by the observation that the vertical length scales for correlations in differences of horizontal wind, $\mathbf{u}$ are less than those for geopotential,

  $$\Delta_v^{\mathbf{u}} < \Delta_v^\Phi.$$

  If winds and geopotential are largely geostrophically related via horizontal gradients,

  $$\mathbf{u} = \frac{1}{f}\mathbf{k} \times \nabla\Phi, \tag{71}$$

  where $\mathbf{k}$ is the local vector pointing to zenith, and f is the Coriolis parameter, then the following follows from Eq. 71,

  $$\Delta_h^{\mathbf{u}} < \Delta_h^\Phi.$$

  The above two inequalities imply an association between horizontal and vertical scales.

The structure functions of an error covariance matrix are its columns (or rows). They are sometimes written as a product of functions - one that is a function of horizontal, but not vertical position, and another that is a function of vertical, but not horizontal position. If this is done, the structure functions are said to be "separable". Furthermore, structure functions that are separable in positional space are also separable in spectral space meaning that the horizontal and vertical space scales are independent. In the light of the points outlined above concerning observed length scales, structure functions of the atmosphere are not separable and consequently this must be represented by the modelled covariances via the control variable transforms.

SELECTED MODEL LEVELS.

level	P	ψ	RKE	χ	DKE	Ap	RH	p	T
28	30	1353	366	1879	418	977		893	756
25	102	1057	277	1343	282	611	234	683	409
19	251	670	273	983	222	490	207	541	297
11	519	599	229	827	179	495	174	500	282
4	867	628	217	1045	195	536	171	499	251
2	961	581	199	724	181	433	164	466	211

Data for July 1998 and January 1999 combined.
The second column gives the average pressures in hPa for the levels. Due t
the vertical grid staggering T and RH are at slightly lower average pressure
than those indicated.

TABLE 2. IMPLIED DIFFERENTIAL LENGTH SCALES (KM), SE LECTED MODEL LEVELS, CF TABLE 1.

level	P	ψ	RKE	χ	DKE	Ap	RH
28	30	356	171	501	200	287	
25	102	306	160	344	169	323	179
19	251	356	171	380	174	357	192
11	519	352	169	320	161	401	199
4	867	339	168	353	167	404	190
2	961	325	164	331	163	373	185

Figure 5: Horizontal length scales for various quantities at a selection of levels in the atmosphere ("RKE" is rotational kinetic energy and "DKE" is divergent kinetic energy). This figure is reproduced from [27]. Panel a is for 'observed' lengths and panel b is for lengths implied from the current covariance model.

5.1.2 Latitudinal variation of vertical errors

Variances and vertical covariances often vary strongly with latitude (and more weakly with longitude). The variation of the tropopause height is an important feature that leaves its mark in the 'observed' error statistics. This is evident in the standard deviation of zonal wind errors (see Fig. 4a of [27], reproduced in Fig. 6a later) and in vertical correlations of temperatures (see Fig. 6c of [27], reproduced in Fig. 8a later).

5.1.3 Latitudinal and height variation of horizontal errors

Horizontal length scales found in the forecast differences are found to be a function of height and latitude. In particular Ingleby found that, generally speaking, errors have longer horizontal scales in the stratosphere than in the troposphere (see table 1 of [27], reproduced as Fig. 5a) and longer scales in the tropics than at high latitudes (see Fig. 1 of [27]).

5.2 The shortcomings of the current covariance model

The current transform imposes restrictions on which features of the 'observed' covariances can be represented adequately in Var. There are known problems with all three components of the transforms (Eq. 30), but we review some important problems with the vertical and horizontal transforms, and their use as a product of transforms in Eqs. 30 and 29.

- In the **T**-transform of the Met Office, the vertical transform is performed before the horizontal transform (see Eq. 29). This makes is possible for the vertical transform to be dependent on horizontal position, and for the horizontal transform to be dependent on vertical mode.
- Two methods of using the vertical transform have been discussed. One with no vertical rotation, giving rise to a vertical-EOF based vertical coordinate for each parameter (ν = EOF-index) and the other that maintains height as the vertical transform (even though in the latter case, local vertical EOFs are used as an intermediate space). These options have different properties and drawbacks depending upon which is used.

ciate different vertical scales with different horizontal scales (this is possible because each vertical EOF has a characteristic vertical length scale, which then can have its own horizontal scale in the horizontal transform). This assocation of vertical and horizontal scales is necessary for non-separable implied structure functions and is captured in the implied covariances. An example involving streamfunction is shown in Fig. 7. The 'observed' covariances (panel a) show a variation of vertical scale with horizontal scale, which is captured in the implied covariances (panel b).

- This first case has its drawbacks. It is not possible to specify a dependence of horizontal scale with height, leading to disagreements between the 'observed' length scales (Fig. 5, panel a) and the implied ones (panel b). Furthermore, this approach does not allow variation of the EOFs with horizontal position (only a variation of their variances with latitude band), leading to a failure to capture the 'observed' latitudinal variation of the tropopause height as manifest, e.g., in the zonal wind standard deviations (Fig. 6, panel a), which do not agree with the implied values (panel b). Failures are evident also in the variation of vertical temperature covariances with latitude. Compare 'observed' values (Fig. 8, panel a) and implied values (panel b).

- For the case with rotation of the vertical EOFs back onto height, the vertical covariance model is allowed to vary with horizontal position. No test cases are available for this option, but would be expected to do well in representing the level dependence on horizontal scale, and the variation of the tropopause with latitude. It should also capture land-sea contrasts.

- For this second case however, it is no longer possible to associate vertical and horizontal scales. This is a serious drawback as the implied structure functions will become nearly separable, and the method will not perform well in capturing the wavenumber dependence shown e.g. in Fig. 7.

The inability of the currect scheme to allow spatial dependencies and to allow associations between horizontal and vertical scales simultaneously is the drive to replace it with a better scheme.

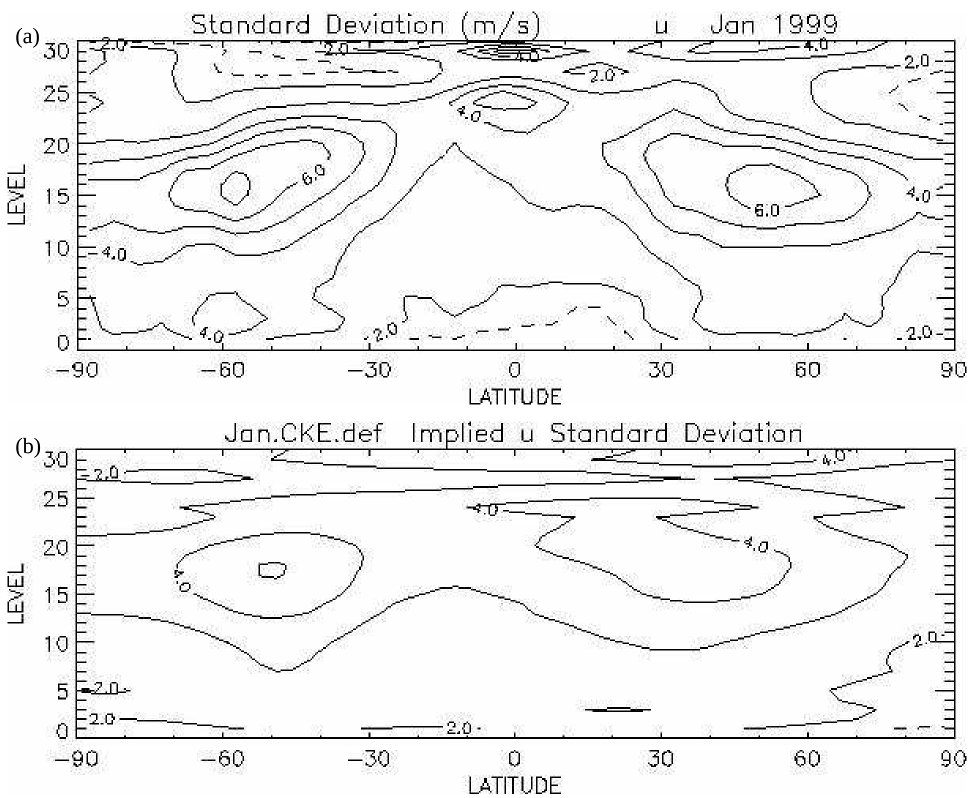

Figure 6: Observed (panel a) and implied (panel b) standard deviation of zonal wind. Reproduced from [27].

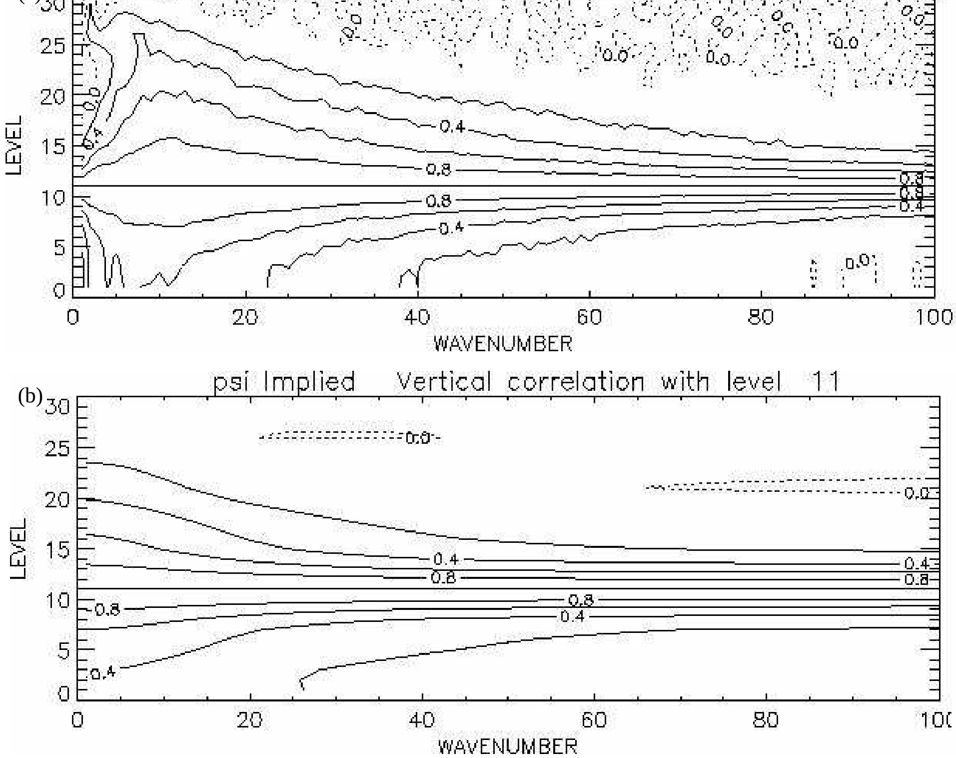

Figure 7: Observed (panel a) and implied (panel b) correlations in streamfunction. The correlation is between streamfunction at each model level with itself at model level 11 ($\sim$ 500 hPa) as a function of horizontal wavenumber. Reproduced from [27].

5.3 Structure functions in the current scheme

Structure functions of the atmosphere are not separable and so the modelled covariances must lead to non-separable structure functions. In appendix 3 we work through calculations of implied structure functions under simplified conditions. We find that they are exactly separable only under certain conditions.

- For the first case of no rotation in the vertical transform, the structure functions are exactly separable in the event that $\mathbf{M}$ in the vertical transform (Eqs. 45-47) is independent of position, and Λ_h in the horizontal transform (Eqs. 55-57) is independent of EOF index. The first condition means that variances of the global vertical EOFs are the same over the globe and the second condition implies that the horizontal length scales are the same for all vertical modes. In practice these conditions are not exactly met by the transforms, meaning that the implied structure functions are non-separable as desired. This result is necessary, given that the agreement between the panels in Fig. 7 is so good.

- For the second case using vertical rotation back to height, we find an analagous conclusion. The structure functions are separable only if the vertical modes - $\mathbf{F}_v$, and their variances - are independent of horizontal position and Λ_h is independent of height.

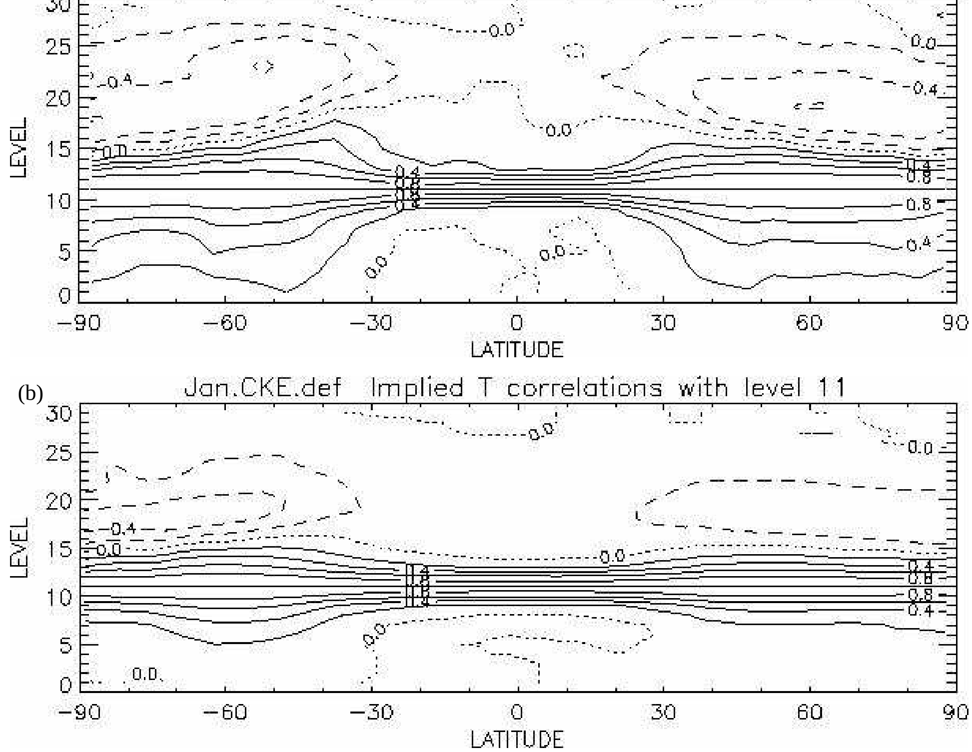

Figure 8: Observed (panel a) and implied (panel b) correlations in temperature. The correlation is between temperature at each model level with itself at model level 11 ($\sim$ 500 hPa) as a function of latitude. Reproduced from [27].

6 THE PROPOSED WAVEBAND SUMMATION TRANSFORM

In order to ease some of the restrictions set by the current control variable transforms, a new model of $\mathbf{B}$ is required with transforms capable of representing the more complicated position and scale dependencies in the covariances. Several alternative transforms have been proposed for use in the Met Office variational assimilation system [28] (see the main part and appendix F of this cited document). These include transforms that use multiple sets of control variables, namely the ECMWF's scheme (labelled "MF" in appendix F of [28]), and a proposed version of this adapted for use with the Met Office's grid-point Unified Model (labelled "PA"). This adapted version keeps many of the ECMWF's conventions intact. A further alternative ("VH") is defined in terms of the Met Office's current vertical and horizontal transforms, applied in a similar way to the current system, but operating (as the above variants do) on different and multiple sets of control variables separately. The final alternative ("WS" - meaning "waveband summation") is along similar lines to the VH transform, but uses only one set of control variables - as in the current approach. The WS transform itself has two variants in [28] - one is discussed in the bulk of this cited document, and the other in appendix F. We will be working with the version introduced in the appendix, at least in the first instance. While it is hoped to yield improved results over the current scheme, it is similar enough in structure to the current operators to facilitate a relatively short development time. This transform is described here.

6.1 The definition of the WS transform

The $\mathbf{U}$-transform for the WS transform is defined as the following (using, where possible, the same notation used beforehand in this paper),

$$\mathbf{U}_{WS} = \mathbf{U}_p \sum_{j=0}^{J} \mathbf{U}_{vj} \mathbf{U}_h \Psi_j^2. \tag{72}$$

Many of these operators are the same as in the current scheme. $\mathbf{U}_p$ is the parameter transform operator (section 4.1), $\mathbf{U}_{vj}$ is the vertical transform operator (section 4.2), $\mathbf{U}_h$ is the horizontal transform operator of total horizontal wavenumber, n, and are described in section 6.2. The summation, j is over $J+1$ bands. In summary this transform, Eq. 72, replaces the current transform, Eq. 30 which acts on the control vector, which represents fields of parameters as functions of horizontal wavenumber and vertical co-ordinate, and generates fields of model increment, which are functions of horizontal position and height. The control vector, acted upon by Eq. 72, has the same structure to that used currently.

The difference between this and the current transform is the summation over bands, the extra 'band index' on the vertical transform and the presence of the spectral band pass filters. Note that the parameter transform is unchanged (the WS transform approach replaces the simple product of vertical and horizontal operators only). Unfortunately, in this formulation, there is no exact inverse transform - only an approximate one. The approximate **T**-transform will be introduced later.

The objective of this new transform is to introduce an extra flexibility into the covariance model to allow it to represent the background errors more realistically than at present. In particular, by virtue of the extra band index on the vertical transform in Eq. 72, it allows a greater degree of non-separability of the horizontal and vertical transforms.

The limitations of the current formulation are discussed in section 5.2 of this document. In particular, we may choose to use one of the following options:

- Allow horizontal length scales to depend upon vertical scales, as required for covariances. The cost is that the shapes of the vertical transforms cannot vary with horizontal position. This is the current working mode.
- Allow vertical covariances to depend fully on horiztonal position, as desirable to capture the variation of the tropopause height and to capture the land-sea contrast. The cost is that the horizontal and vertical length scales cannot be associated. This is not a current working mode, but can be achieved with modification.

The WS transform has the potential to satisfy both of these features (scale and position dependence) simultaneously if the vertical transform has characteristics of the second of the above options.

Each band pass filter is designed to select a different range of horizontal length scales from the control vector, and each can be treated by a different vertical operator (this is the purpose of the extra index 'j' on $\mathbf{U}_{vj}$). This gives both position <u>and scale</u> dependence to the new vertical covariance model. The number of bands, $J+1$, is unspecified at present. The more wavebands that are used, the greater the scope for scale dependence on the vertical covariances. The price for this however is that more computation is needed, and, the spatial dependence becomes less meaningful.

6.1.1 Restrictions

There is a restriction on how $\mathbf{U}_{vj}$ can depend upon band if we choose no rotation of the vertical modes (when the vertical co-odinate is EOF-index - section 4.2.3). In this case the waveband approach does not allow different vertical modes in $\mathbf{F}_v$ for different bands. This is demanded because the results from each operator are added, and so the vector spaces must be the same. The new 'band dependence' on the vertical transform must therefore be made through the vertical variances only, allowing vertical and horizontal scales to be associated.

There is no such restriction if the EOFs are rotated back to height space, as in the vertical scheme proposed in section 4.2.2. It is envisaged that this will be the choice used with the WS transform and so we will probably not have to worry about the above restriction.

In line with [28], there is no band dependence on the horizontal transform in the version of the WS transform given as Eq. (72). The reason is to keep the transform as close as possible to the current scheme, but later (section 6.5) a band dependence will be considered.

6.1.2 Similarity with current transforms

The WS transform of Eq. 72 has the same effective property as the current transforms if the band index is removed from the vertical transform, and if the bandpass filters obey a normalization condition. The physical effect of the transform is revealed through its implied covariance. Substituting Eq. 72 into Eq.

$$
\begin{aligned}
\mathbf{B}_{WS} &= \mathbf{U}_{WS}\mathbf{U}_{WS}^T \\
&= \left(\mathbf{U}_p \sum_{j=0}^{J} \mathbf{U}_v \mathbf{U}_h \Psi_j^2\right)\left(\mathbf{U}_p \sum_{j'=0}^{J} \mathbf{U}_v \mathbf{U}_h \Psi_{j'}^2\right)^T \\
&= \mathbf{U}_p \mathbf{U}_v \mathbf{U}_h \left(\sum_{j=0}^{J} \Psi_j^2\right)\left(\sum_{j'=0}^{J} \Psi_{j'}^2\right) \mathbf{U}_h^T \mathbf{U}_v^T \mathbf{U}_p^T,
\end{aligned} \tag{73}
$$

which is the same as for the current transforms, $\mathbf{U}_p \mathbf{U}_v \mathbf{U}_h \mathbf{U}_h^T \mathbf{U}_v^T \mathbf{U}_p^T$, if, for each total horizontal wavenumber (n), the band pass filters are normalized, ie,

$$
\sum_{j=0}^{J} \Psi_j^2(n) = 1, \tag{74}
$$

(the definition of the total wavenumber is given in section 2 of [29]). In matrix form this is, $\sum_{j=0}^{J} \Psi_j^2 = \mathbf{I}$, where the matrix Ψ_j is diagonal and the total wavenumber dependence of the function $\Psi_j(n)$ runs along its diagonal.

6.2 The spectral band pass filters

The shapes of the functions $\Psi_j(n)$, and their dependence on the band, j, are now specified. The Met Office has proposed using so-called 'optimal' functions that have the following triangular form [28],

$$
\Psi_j^2(n) = \begin{cases} 0 & \text{for} \quad n \leq N_{j-1} \\ 1 & \text{for} \quad n = N_j \\ 0 & \text{for} \quad n \geq N_{j+1} \end{cases}, \tag{75}
$$

with linear dependence between the wavenumbers N_{j-1} to N_j, and N_j to N_{j+1}. (The exception is the 'last' band pass function in the set defined. This has a slightly different form - see below.) An example band pass function is shown in Fig. 9. The wavenumbers, N_j, that help define these functions are found by using the definition of 'optimality' (see below).

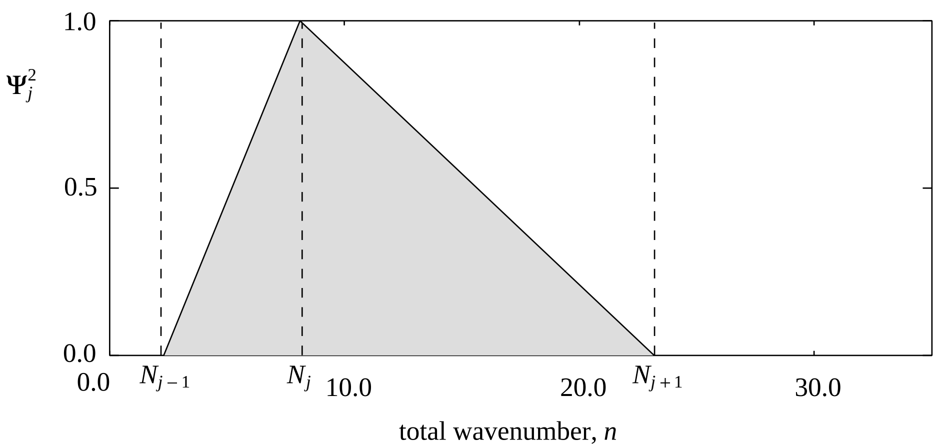

Figure 9: Example triangular band pass function j, as defined in Eq. 75. The three total wavenumbers, N_{j-1}, N_j and N_{j+1}, that define such functions are found from Eq. 78.

6.2.1 Determination of the bounding wavenumbers for 'optimal' functions

The Met Office define the optimal band pass functions as those that have a 'full width at half maximum' (FWHM) equal to the 'typical wavenumber' of the function. For the functions defined in Eq. 75, the FWHM is the following (for band j),

$$
\text{FWHM}_j = \frac{1}{2}(N_{j+1} - N_{j-1}). \tag{76}
$$

[28] is the average of the wavenumbers at the points where the function is half maximum,

$$\begin{aligned}\bar{n}_j &= \frac{1}{2}\left\{(N_{j-1} + \frac{1}{2}(N_j + N_{j-1})) + (N_j + \frac{1}{2}(N_{j+1} + N_j))\right\}, \\ &= \frac{1}{4}\{N_{j-1} + 2N_j + N_{j+1}\}, \end{aligned} \tag{77}$$

which has a more general definition for more general functions in appendix A.2 of [28].

Determining the N_j arises from equating Eqs. 76 with 77, which yields the following recurrence relation,

$$N_{j+1} = 3N_{j-1} + 2N_j. \tag{78}$$

The boundary conditions on N_j are (as in [28]) are $N_0 = 0$ and $N_J = N_{max}$. Equation 78 then represents a set of (tridiagonal) simultaneous equations whose solution is the set of N_j,

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 2 & -1 & 0 & 0 & 0 & 0 \\ & & & & & & & \\ & & & & & & & \\ 0 & 0 & 0 & 0 & 3 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} N_0 \\ N_1 \\ N_2 \\ N_3 \\ \\ N_{J-3} \\ N_{J-2} \\ N_{J-1} \\ N_J \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \\ 0 \\ 0 \\ 0 \\ N_{max} \end{pmatrix}. \tag{79}$$

The determination of the upper limit of the summation, J, is discussed in appendix A.4 of [28]. They start by writing the system of equations, Eq. 78, in a different way by defining a new parameter K_j,

$$K_j = 3K_{j-1}^{-1} + 2, \tag{80}$$

$$\text{where } K_j = N_{j+1}/N_j, \tag{81}$$

and then multiplying K_j^{-1} in Eq. 81 between $j = 1$ and $J - 1$.

$$\begin{aligned}\prod_{j=1}^{J-1} K_j^{-1} &= \prod_{j=1}^{J-1} \frac{N_j}{N_{j+1}}, \\ &= \left(\prod_{j=1}^{J-1} N_j\right) \times \left(\prod_{j=1}^{J-1} \frac{1}{N_{j+1}}\right), \\ &= \left(\prod_{j=1}^{J-1} N_j\right) \times \left(\prod_{j=2}^{J} \frac{1}{N_j}\right), \\ &= N_1 \left(\prod_{j=2}^{J-1} N_j\right) \times \left(\prod_{j=2}^{J-1} \frac{1}{N_j}\right) \frac{1}{N_J}, \\ &= \frac{N_1}{N_J}. \end{aligned} \tag{82}$$

If we choose to restrict $N_1 \geq 1$, then Eq. 82 leads to,

$$N_J \prod_{j=1}^{J-1} K_j^{-1} \geq 1. \tag{83}$$

Equation 80 tells us what the set of K_j are (given that $K_1 = 2$), which then allows computation of the product in Eq. 83. By evaulating the left hand side of Eq. 83 for trial values of J, we use the largest J that the inequality in Eq. 83 holds.

The example band pass functions shown in [28] are reproduced in Fig. 10. Note that the condition of Eq. 74 is satisfied at all wavenumbers, as required. Importantly, for the triangular functions, no more than two are non-zero at any one wavenumber. For a particular maximum wavenumber, there is a choice of two conventions in how the last function (highest band index) is represented, and the two possibilities require different numbers of functions, differing by one. The panels in Fig. 10 cover the two schemes. The last function of panel (a) is a so-called 'oversized' function and in panel (b) it is an 'undersized' function. These alternatives each have good and bad features.

- The typical wavenumber of the oversize function in panel (a) is smaller than its FWHM. The length scales of this function therefore have a greater range than they should have, including smaller scales. This is not a problem with the undersized option in panel (b).
- One less band pass function is needed in panel (a) than in panel (b) to fill the wavenumber range and to satisfy Eq. 74. Implementing the oversized option therefore presents the possibility for a more efficient solution.

The discussion in [28] recommends use of the oversized scheme for reasons of efficiency, and that any inadequacies at small length scales are probably swamped by problems in the model anyway. Also discussed in appendix D of [28] are other shapes of band pass filters.

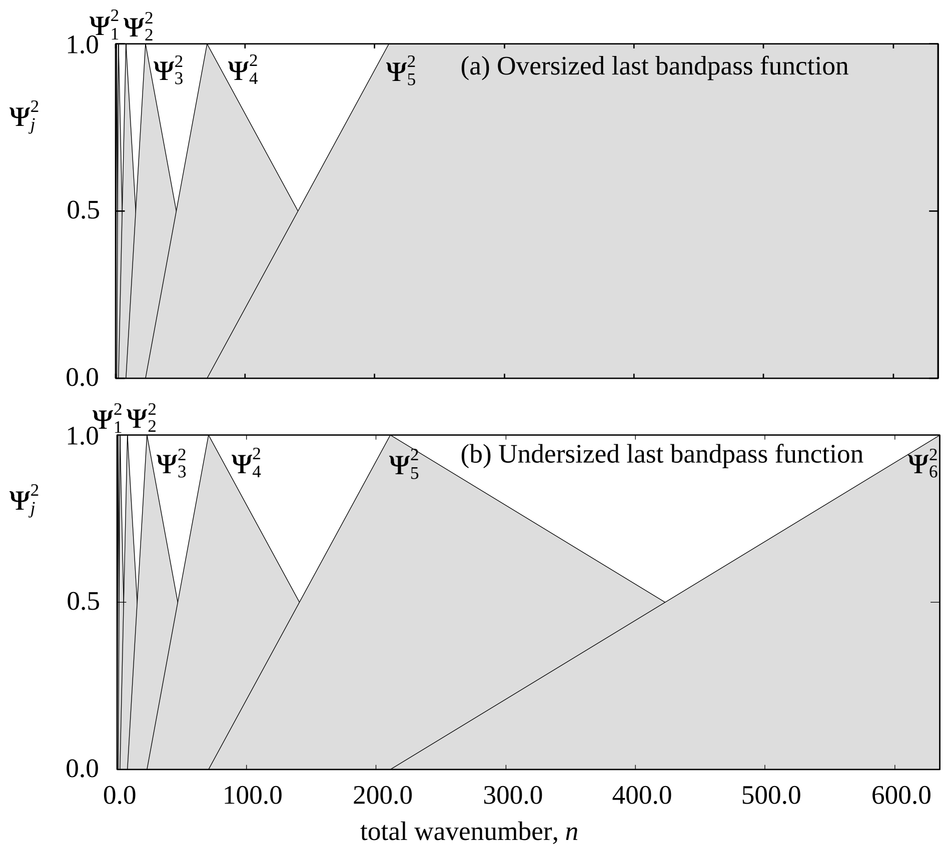

Figure 10: Example triangular band pass functions showing the two alternatives for specification of the final function required to fill the same total wavenumber domain in each case. Panel (a) uses an oversize function and panel (b) uses an undersize function. The oversize option requires one less function than the undersize option.

6.3 The T-transform for the WS transform

A difficulty of the waveband summation transform is that it has no exact inverse. The foremost cause of this is the non-orthogonality of the band pass functions. Before suggesting an approximate inverse

Eq. 72.

6.3.1 Commutation between the vertical and horizontal transforms

There are two alternatives for the configuration of the vertical transform and we consider separately the commutation of each with the horizontal transform and with the bandpass filters (proofs are given in appendix 4).

Considering firstly the configuration with no vertical rotation (section 4.2.3), the combination $\mathbf{U}_v\mathbf{U}_h$ can be written with reference to Eqs. 45 (with an extra band index, $\mathbf{M}_j$) and 55 as,

$$\begin{aligned}\mathbf{U}_v\mathbf{U}_h &= \mathbf{Z}^{-1}\mathbf{M}_j^{1/2}\mathbf{F}_h\Lambda_h^{1/2}, \\ &= (\mathbf{Z}^{-1}\mathbf{M}_j^{1/2}\mathbf{F}_h\Lambda_h^{1/2}\mathbf{F}_h^{-1})\mathbf{F}_h. \end{aligned} \tag{84}$$

The bracketted term in Eq. 84 is the combined vertical $\mathbf{U}$-transform and the horizontal $\mathbf{U}$-transform written in real space. Using square brackets, $[A,B]$, to denote the commutator, $AB-BA$, the commutation between the vertical and horizontal transforms can be written as,

$$[\mathbf{Z}^{-1}\mathbf{M}_j^{1/2}, \mathbf{F}_h\Lambda_h^{1/2}\mathbf{F}_h^{-1}] = 0 \text{ if } \mathbf{M}_j \text{ ind. of horiz. pos. and } \Lambda_h \text{ ind. of } z, \tag{85}$$

$$\neq 0 \text{ otherwise.} \tag{86}$$

With vertical rotation (Eq. 34 with an extra band index, $\mathbf{F}_{vj}$, Λ_{vj}), the commutator is,

$$[\mathbf{F}_{vj}\Lambda_{vj}\mathbf{F}_{vj}^T, \mathbf{F}_h\Lambda_h^{1/2}\mathbf{F}_h^{-1}] = 0 \text{ if } \Lambda_{vj},\ \mathbf{F}_{vj} \text{ ind. of horiz. pos. and } \Lambda_h \text{ ind. of EOF index,} \tag{87}$$

$$\neq 0 \text{ otherwise.} \tag{88}$$

6.3.2 Commutation between part of the horizontal transform and waveband filters

Since each are diagonal operators in spectral space, the part of the horizontal transform, $\Lambda_h^{1/2}$, commutes with each bandpass function, Ψ_j^2,

$$[\Lambda_h^{1/2}, \Psi_j^2] = 0. \tag{89}$$

6.3.3 Commutation between the vertical transform and waveband filters

Once more, we consider the two vertical transform configurations. For the case with no rotation, we examine the commutivity between the vertical transform, $\mathbf{U}_{vj} = \mathbf{Z}^{-1}\mathbf{M}_j^{1/2}$, and the waveband operator translated into grid space, $\mathbf{F}_h\Psi_j^2\mathbf{F}_h^{-1}$,

$$[\mathbf{Z}^{-1}\mathbf{M}_j^{1/2}, \mathbf{F}_h\Psi_j^2\mathbf{F}_h^{-1}] = 0. \text{ if } \mathbf{M}_j \text{ ind. of horiz. pos.} \tag{90}$$

$$\neq 0. \text{ otherwise} \tag{91}$$

which is the same result as in section 6.3.1 for commutation between the vertical and horizontal transforms. This commutivity property does not hold strictly if $\mathbf{M}_j^{1/2}$ is not independent of position. It is however assumed to hold approximately if its positional dependence is weak.

With vertical rotation in the simple vertical scheme outlined in section 4.2.2, the commutator is,

$$[\mathbf{F}_{vj}\Lambda_{vj}^{1/2}\mathbf{F}_{vj}^T, \mathbf{F}_h\Psi_j^2\mathbf{F}_h^{-1}] = 0. \text{ if } \Lambda_{vj},\ \mathbf{F}_{vj} \text{ ind. of horiz. pos.} \tag{92}$$

$$\neq 0. \text{ otherwise,} \tag{93}$$

which is the same result as in section 6.3.1 for commutation between the vertical and horizontal transforms.

The commutation relations stated above can be used to rewrite $\mathbf{U}_{WS}$ in the following way, starting from Eq. 72 and following the essence of [28],

$$\begin{aligned}
\mathbf{U}_{WS} &= \mathbf{U}_p \sum_{j=0}^{J} \mathbf{U}_{vj} \mathbf{U}_h \Psi_j^2, \\
&= \mathbf{U}_p \sum_{j=0}^{J} \mathbf{U}_{vj} \mathbf{F}_h \Lambda_h^{1/2} \Psi_j^2, \\
&= \mathbf{U}_p \sum_{j=0}^{J} \mathbf{U}_{vj} \mathbf{F}_h \Psi_j^2 \Lambda_h^{1/2}, \\
&= \mathbf{U}_p \sum_{j=0}^{J} \mathbf{U}_{vj} (\mathbf{F}_h \Psi_j^2 \mathbf{F}_h^{-1}) \mathbf{F}_h \Lambda_h^{1/2}, \\
&\approx \mathbf{U}_p \sum_{j=0}^{J} (\mathbf{F}_h \Psi_j^2 \mathbf{F}_h^{-1}) \mathbf{U}_{vj} \mathbf{F}_h \Lambda_h^{1/2}, \\
&\approx \mathbf{U}_p \sum_{j=0}^{J} \mathbf{F}_h \Psi_j^2 \mathbf{F}_h^{-1} \mathbf{U}_{vj} \mathbf{U}_h. \qquad (94)
\end{aligned}$$

Equation 55 has been used in the second and sixth lines, Eq. 89 in the third, and Eq. 90 or Eq. 92 (depending upon the vertical transform, and applied approximately) in the fifth. The $\mathbf{U}_{WS}$ transform has been rewritten like this, solely in order to write an approximate left inverse transform. The proposed left inverse transform is found from Eq. 94,

$$\mathbf{T}_{WS} \approx \sum_{j=0}^{J} \mathbf{T}_h \mathbf{T}_{vj} \mathbf{F}_h \Psi_j^2 \mathbf{F}_h^{-1} \mathbf{T}_p, \qquad (95)$$

where Ψ_j^2 is left intentionally with a positive power of two. It can be demonstrated that Eq. 95 is the approximate left inverse of Eq. 94 by showing that $\mathbf{T}_{WS}\mathbf{U}_{WS} \approx \mathbf{I}$, as below,

$$\begin{aligned}
\mathbf{T}_{WS}\mathbf{U}_{WS} &= \left(\sum_{j=0}^{J} \mathbf{T}_h \mathbf{T}_{vj} \mathbf{F}_h \Psi_j^2 \mathbf{F}_h^{-1} \mathbf{T}_p \right) \left(\mathbf{U}_p \sum_{j=0}^{J} \mathbf{F}_h \Psi_j^2 \mathbf{F}_h^{-1} \mathbf{U}_{vj} \mathbf{U}_h \right), \\
&= \left(\sum_{j=0}^{J} \mathbf{T}_h \mathbf{T}_{vj} \mathbf{F}_h \Psi_j^2 \right) \left(\sum_{j=0}^{J} \Psi_j^2 \mathbf{F}_h^{-1} \mathbf{U}_{vj} \mathbf{U}_h \right), \\
&= \left(\sum_{j=0}^{J} \Lambda_h^{-1/2} \mathbf{F}_h^{-1} \mathbf{T}_{vj} \mathbf{F}_h \Psi_j^2 \right) \left(\sum_{j=0}^{J} \Psi_j^2 \mathbf{F}_h^{-1} \mathbf{U}_{vj} \mathbf{F}_h \Lambda_h^{1/2} \right), \qquad (96)
\end{aligned}$$

making use of Eqs. 57 and 55 in the last line. Appendix 4 demonstrates that $\mathbf{F}_h^{-1}\mathbf{T}_{vj}\mathbf{F}_h$ and $\mathbf{F}_h^{-1}\mathbf{U}_{vj}\mathbf{F}_h$, as appearing above, which are the vertical $\mathbf{T}$ and $\mathbf{U}$ transforms in spectral space, are approximately block diagonal in wavenumber if the components of the vertical transform (irrespective of the kind of transform used - Eqs. 34-36 or Eqs. 45-47) are only weak functions of horizontal position.

This has two consequences. Firstly that $\Lambda_h^{1/2}$ will commute approximately with $\mathbf{F}_h^{-1}\mathbf{U}_{vj}\mathbf{F}_h$. This, and the commutator in Eq. 89 allows the $\Lambda_h^{1/2}$ from each summation in Eq. 96 to be brought together, outside of the summations, allowing them to cancel. The second consequence exploits the property that at any total wavenumber, n, no more than two bandpass functions are non-zero (see Fig. 10). Hence each summation in the combination $\mathbf{T}_{WS}\mathbf{U}_{WS}$, written for a single total wavenumber only, reduces to the sum of two terms. If, for a given n, wavebands $j = \alpha$ and $j = \beta$ (where $\beta = \alpha + 1$) contribute, then

$$
\begin{aligned}
\mathbf{T}_{WS}\mathbf{U}_{WS}(n) &\approx \left(\mathbf{F}_h^{-1}\mathbf{T}_{\mathbf{v}\alpha}\mathbf{F}_h\Psi_\alpha^2(n) + \mathbf{F}_h^{-1}\mathbf{T}_{\mathbf{v}\beta}\mathbf{F}_h\Psi_\beta^2(n)\right) \times \\
&\quad \left(\Psi_\alpha^2(n)\mathbf{F}_h^{-1}\mathbf{U}_{\mathbf{v}\alpha}\mathbf{F}_h + \Psi_\beta^2(n)\mathbf{F}_h^{-1}\mathbf{U}_{\mathbf{v}\beta}\mathbf{F}_h\right) \\
&\approx \mathbf{F}_h^{-1}\mathbf{T}_{\mathbf{v}\alpha}\mathbf{U}_{\mathbf{v}\alpha}\mathbf{F}_h\Psi_\alpha^4(n) + \mathbf{F}_h^{-1}\mathbf{T}_{\mathbf{v}\beta}\mathbf{U}_{\mathbf{v}\alpha}\mathbf{F}_h\Psi_\beta^2(n)\Psi_\alpha^2(n) + \\
&\quad \mathbf{F}_h^{-1}\mathbf{T}_{\mathbf{v}\alpha}\mathbf{U}_{\mathbf{v}\beta}\mathbf{F}_h\Psi_\alpha^2(n)\Psi_\beta^2(n) + \mathbf{F}_h^{-1}\mathbf{T}_{\mathbf{v}\beta}\mathbf{U}_{\mathbf{v}\beta}\mathbf{F}_h\Psi_\beta^4(n).
\end{aligned} \tag{97}
$$

In Eq. 97, $\mathbf{T}_{\mathbf{v}\alpha}\mathbf{U}_{\mathbf{v}\alpha} = \mathbf{I}$, and $\mathbf{T}_{\mathbf{v}\beta}\mathbf{U}_{\mathbf{v}\beta} = \mathbf{I}$. Since α and β are neighbouring total wavenumbers, the corresponding vertical transforms will probably not be very different, and hence we may take, $\mathbf{T}_{\mathbf{v}\alpha}\mathbf{U}_{\mathbf{v}\beta} \approx \mathbf{I}$ and $\mathbf{T}_{\mathbf{v}\beta}\mathbf{U}_{\mathbf{v}\alpha} \approx \mathbf{I}$. Then, Eq. 97 becomes,

$$
\begin{aligned}
\mathbf{T}_{WS}\mathbf{U}_{WS}(n) &\approx \Psi_\alpha^4(n) + 2\Psi_\beta^2(n)\Psi_\alpha^2(n) + \Psi_\beta^4(n) \\
&\approx \Psi_\alpha^2(n)(\Psi_\alpha^2(n) + \Psi_\beta^2(n)) + \Psi_\beta^2(n)(\Psi_\alpha^2(n) + \Psi_\beta^2(n)) \\
&\approx 1,
\end{aligned} \tag{98}
$$

the last line being due to the normalization condition, Eq. 74. This demonstrates that $\mathbf{T}_{WS}$ (Eq. 95) is the approximate left inverse of $\mathbf{U}_{WS}$ (Eq. 72).

6.4 The $\mathbf{U}^T$-transform for the WS transform

The transpose of the waveband summation transform, $\mathbf{U}_{WS}^T$, follows in a straightforward manner from Eq. 72. Taking the transpose yields,

$$
\mathbf{U}_{WS}^T = \sum_{j=0}^{J} \Psi_j^2 \mathbf{U}_h^T \mathbf{U}_{vj}^T \mathbf{U}_p^T. \tag{99}
$$

The individual transposes $\mathbf{U}_h^T$, $\mathbf{U}_{vj}^T$ and $\mathbf{U}_p^T$ should follow from the standard code that already exists (with appropriate modifications to $\mathbf{U}_{vj}^T$ given the extra band index). The Ψ_j^2 operator is diagonal in spectral space and so is trivially self-adjoint.

6.5 Calibrating the WS transform

The covariances statistics for the waveband summation (WS) transform are calibrated in a similar way to the current transforms (see section 4.5), but now there is more information to deal with due to the multiple bands. It is here that we make use of the approximate inverse of $\mathbf{U}_{WS}$ derived in section 6.3.4. Recall that in the current transforms, we have to determine the following.

- The vertical covariance matrices at each position, $\mathbf{B}_v(\lambda, \phi)$ are required. When averaged globally, these will yield the eigenvectors, $\mathbf{F}_v$ and eigenvalues, Λ_v, which will allow calculation of the diagonal matrices $\mathbf{M}(\lambda, \phi)$. This is the information required for the current vertical transform in the case of no rotation (see section 4.2.3).
- The horizontal variances, Λ_h, can then be computed for each vertical mode. These are then used to determine a length scale for each vertical mode.

The WS transform calibration (below) is more complicated than this and there are a number of alternative approaches that can be considered. Options include the use of top-hat and triangular band-pass functions and the introduction of a band-dependent horizontal transform - something that has not been considered until now. For the WS transform, we will use a specific choice of rotation in the vertical transform (see below).

The structure of the WS calibration procedure is similar to that for the current transform (section 4.5). It is proposed that the NMC method will be continued and so in the following, $\vec{x}'$ refers to the forecast differences as in Eq. 64 and the angled braces, $\langle\rangle$, refer to an average over situations. As in the current scheme, the calibration is done in stages and we can consider each parameter (ψ, χ, ${}^{A}p$, etc.) separately (we consider ψ below). There are a number of choices that have to be made.

- Full benefit of the WS transform is gained when the vertical transform is a function of horizontal position (variances and EOFs). In the light of the discussion in section 4.2.1, this requires the version of the vertical transform discussed in section 4.2.2 which works by projecting fields onto vertical transform is simpler than the one currently implemented and is the version used in the description below.

- The spectral bandpass filters overlap (Fig. 10). This creates difficulties when dealing with calibration of the horizontal transform. Many treatments of this problem are discussed below - extending the discussion given in section 7 of [28]. Four options are discussed, labelled as 1a, 1b, 2a and 2b (section 6.5.2). Considerations include the use of non-overlapping top-hat bandpass functions and an approximate treatment of the band coupling present when using the triangular bandpass functions.

6.5.1 Calibrating the vertical transform

The vertical transform statistics depend upon the waveband. Analagous to Eq. 65, the vertical error covariance matrix for ψ at position $\vec{r}$ in waveband j is,

$$\mathbf{B}_{vj}(\lambda,\phi)=\langle\mathbf{C}(\vec{r})\mathbf{F}_h\Psi_j^2\mathbf{F}_h^{-1}\mathbf{T}_p^{\psi}\vec{x}'(\mathbf{C}(\vec{r})\mathbf{F}_h\Psi_j^2\mathbf{F}_h^{-1}\mathbf{T}_p^{\psi}\vec{x}')^T\rangle. \tag{100}$$

The string of operators $\mathbf{C}(\vec{r})\mathbf{F}_h\Psi_j^2\mathbf{F}_h^{-1}\mathbf{T}_p^{\psi}\vec{x}'$ does the following: (i) $\mathbf{T}_p^{\psi}$ generates streamfunction from $\vec{x}'$, (ii) $\mathbf{F}_h\Psi_j^2\mathbf{F}_h^{-1}$ filters through a range of wavenumbers as specified by the bandpass function Ψ_j^2, and (iii) $\mathbf{C}(\vec{r})$ isolates the vertical column at position $\vec{r}$. This string of operators (except for $\mathbf{C}(\vec{r})$) is the same as that string appearing in the approximate form of the WS T-transform in Eq. 95.

The waveband (and position) dependent vertical transforms are, from Eqs. 34 to 36 (and with an extra band index, j),

$$\begin{aligned}\mathbf{U}_{vj} &= \mathbf{P}_v^{-1/2}\mathbf{F}_{vj}\Lambda_{vj}^{1/2}\mathbf{F}_{vj}^T, & (101)\\ \mathbf{U}_{vj}^T &= \mathbf{F}_{vj}\Lambda_{vj}^{1/2}\mathbf{F}_{vj}^T\mathbf{P}_v^{-T/2}, & (102)\\ \mathbf{T}_{vj} &= \mathbf{F}_{vj}\Lambda_{vj}^{-1/2}\mathbf{F}_{vj}^T\mathbf{P}_v^{1/2}. & (103)\end{aligned}$$

The matrices $\mathbf{F}_{vj}(\vec{r})$ and $\Lambda_{vj}(\vec{r})$ are, respectively, the eigenvectors and eigenvalues of the matrix $\mathbf{B}_{vj}$ in Eq. 100 (they are functions of position $\vec{r}=(\lambda,\phi)$). This process is repeated for each position and band (and parameter) and in total describes the vertical part of the WS transform.

Note: in the case of the above equations, the bandpass functions are the triangular functions shown in Fig. 10. For some options in the discussion to follow, the transforms will use the top-hat functions shown in Fig. 11 instead, which are denoted by a 'hat', $\hat{\Psi}_j^2$. In these cases, replace Ψ_j^2 with $\hat{\Psi}_j^2$ in Eq. 100.

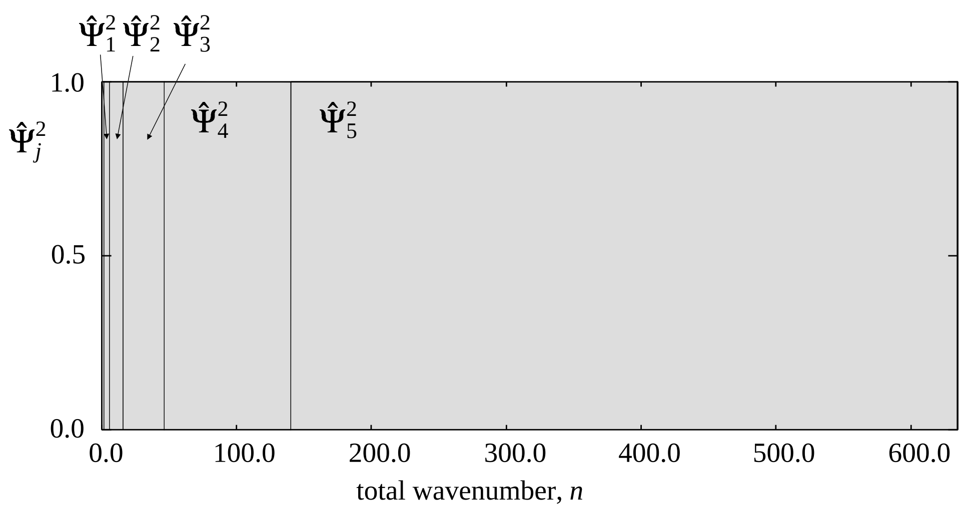

Figure 11: Non-overlapping top-hat band pass functions, $\hat{\Psi}_j^2$, approximating the triangular ones of Fig. 10. The sides of the top-hat $\hat{\Psi}_j^2$ functions is taken to be midway between the peaks of neighbouring Ψ_j^2 functions.

The equations needed to calibrate the horizontal transforms are now developed with four different options for treating the band overlap.

Calibration option 1a: non-overlapping top-hat bandpass functions with a band-independent horizontal transform

After finding the vertical transform matrices (refer to Eq. 100 with Ψ_j^2 replaced with $\hat{\Psi}_j^2$) the horizontal transform needs to be determined. By analogy with section 4.5.2, start with the definition of $\mathbf{B}$ (Eq. 62) and substitute for $\vec{x}'$ using the approximate $\mathbf{U}_{WS}$-transform (Eqs. 2 and 94),

$$\langle \vec{x}' \vec{x}'^T \rangle \approx \langle \mathbf{U}_p \sum_{j=0}^{J} \mathbf{F}_h \hat{\Psi}_j^2 \mathbf{F}_h^{-1} \mathbf{U}_{vj} \mathbf{U}_{hj} \vec{v}' \vec{v}'^T (\mathbf{U}_p \sum_{j'=0}^{J} \mathbf{F}_h \hat{\Psi}_{j'}^2 \mathbf{F}_h^{-1} \mathbf{U}_{vj'} \mathbf{U}_{hj'})^T \rangle. \tag{104}$$

There are a number of differences between the approximate $\mathbf{U}_{WS}$-transform used above and that shown in Eq. 94. In this first option, 1a, we treat the band overlap by ignoring it, ie top-hat bandpass functions have been used instead of triangular ones (only for calibration). There is also an additional band index on the horizontal transform, even though we will soon drop this (see below for why this index has been added). The angled brackets can be taken to the $\vec{v}' \vec{v}'^T$, giving $\langle \vec{v}' \vec{v}'^T \rangle = \mathbf{I}$ (control vectors are uncorrelated). Even though the bandpass functions are non-overlapping, there can be a contribution to $\langle \vec{x}' \vec{x}'^T \rangle$ even when $j \neq j'$. This inter-band coupling can arise from scales introduced by position dependences in the vertical transform. For simplicity though we will proceed by assuming that such coupling is negligible and take contributions to $\langle \vec{x}' \vec{x}'^T \rangle$ from $j = j'$ only, thus elliminating one of the summations,

$$\langle \vec{x}' \vec{x}'^T \rangle \approx \mathbf{U}_p \sum_{j'=0}^{J} \mathbf{F}_h \hat{\Psi}_{j'}^2 \mathbf{F}_h^{-1} \mathbf{U}_{vj'} \mathbf{U}_{hj'} (\mathbf{U}_p \mathbf{F}_h \hat{\Psi}_{j'}^2 \mathbf{F}_h^{-1} \mathbf{U}_{vj'} \mathbf{U}_{hj'})^T. \tag{105}$$

Take the parameter transform and the outer spectral transforms to the left-hand side. Considering only one parameter (ψ),

$$\langle \mathbf{F}_h^{-1} \mathbf{T}_p^{\psi} \vec{x}' (\mathbf{F}_h^{-1} \mathbf{T}_p^{\psi} \vec{x}')^T \rangle \approx \sum_{j'=0}^{J} \hat{\Psi}_{j'}^2 \mathbf{F}_h^{-1} \mathbf{U}_{vj'} \mathbf{U}_{hj'} (\hat{\Psi}_{j'}^2 \mathbf{F}_h^{-1} \mathbf{U}_{vj'} \mathbf{U}_{hj'})^T. \tag{106}$$

Multiply (to the left and right) by the operator $\hat{\Psi}_j^2$. The only member of the summation over j' that survives is where $j' = j$,

$$\langle \hat{\Psi}_j^2 \mathbf{F}_h^{-1} \mathbf{T}_p^{\psi} \vec{x}' (\hat{\Psi}_j^2 \mathbf{F}_h^{-1} \mathbf{T}_p^{\psi} \vec{x}')^T \rangle \approx \mathbf{F}_h^{-1} \mathbf{U}_{vj} \mathbf{U}_{hj} (\mathbf{F}_h^{-1} \mathbf{U}_{vj} \mathbf{U}_{hj})^T. \tag{107}$$

Take the inverse spectral transforms, and the vertical transforms, which are now known from subsection 6.5.1, to the left hand side,

$$\langle \mathbf{T}_{vj} \mathbf{F}_h \hat{\Psi}_j^2 \mathbf{F}_h^{-1} \mathbf{T}_p^{\psi} \vec{x}' (\mathbf{T}_{vj} \mathbf{F}_h \hat{\Psi}_j^2 \mathbf{F}_h^{-1} \mathbf{T}_p^{\psi} \vec{x}')^T \rangle \approx \mathbf{U}_{hj} (\mathbf{U}_{hj})^T. \tag{108}$$

The horizontal transform is defined by $\mathbf{F}_h \Lambda_{hj}^{1/2}$ (Eq. 55). Substituting this into Eq. 107 gives the horizontal spectrum for band j,

$$\langle \mathbf{F}_h^{-1} \mathbf{T}_{vj} \mathbf{F}_h \hat{\Psi}_j^2 \mathbf{F}_h^{-1} \mathbf{T}_p^{\psi} \vec{x}' (\mathbf{F}_h^{-1} \mathbf{T}_{vj} \mathbf{F}_h \hat{\Psi}_j^2 \mathbf{F}_h^{-1} \mathbf{T}_p^{\psi} \vec{x}')^T \rangle \approx \Lambda_{hj}. \tag{109}$$

Λ_{hj} is the diagonal matrix of variances as a function of wavenumber and height and Eq. 109 can be used to determine these diagonal elements. Even though there is a band label on this horizontal spectrum, it does not mean that the horizontal transform has to be band-dependent. The label here identifies the part of the spectrum (range of wavenumbers) that is found from Eq. 109 for a given j. Due to the spectral filtering in Eq. 109, the horizontal variances for band j are expected to be non-zero predominantly inside the range of wavenumbers associated with band j only (via the filter $\hat{\Psi}_j^2$ - see Fig. 11). For option 1a, only this weight inside the top-hat is preserved (it is set to zero outside). By applying this procedure over all bands j ($0 \leq j \leq J$), the entire spectrum of variances at each level is found for each parameter,

$$\Lambda_h = \sum_{j=0}^{J} \Lambda_{hj}. \tag{110}$$

horizontal transform

There are two instances where inter-band coupling is neglected in option 1a. The first is with the removal of one of the band summations to give Eq. 105, and the other is made near the end to give a band-independent horizontal transform. It is possible to not neglect the weight outside of each top-hat function in the second instance and we may wish to investigate its effect by capturing the entire spectrum for Λ_{hj}. This gives rise to a different spectrum for each band and thus introduces a band dependence on the horizontal transform.

Calibration option 2a: triangular band pass functions with a band-independent horizontal transform

It is claimed that it is possible to calibrate the WS transform without the need to approximate the band pass functions with non-overlapping top-hat functions. This procedure uses the same triangular band pass functions that are present in the original definition of $\mathbf{U}_{WS}$ (Fig. 9), but still uses the approximate $\mathbf{U}_{WS}$-transform (Eq. 94). Just as for the top-hat function case, there are two options - with and without a band-dependent horizontal transform. Of these two triangular cases, the band-independent case (option 2a) is by far the simpler method, but treats the overlap in an approximate way. It is left to option 2b for an outline of the exact formulation.

Apart from the use of triangular functions in the calculation of the vertical and horizontal transforms, the procedure for triangular functions here differs from option 1a in another way. Some of the transforms used in the calibration may require a modification to correct for double counting of variances; the correction is in the form of a spectral correction function, $\Gamma(n)$. Double counting occurs because the triangular band pass functions overlap and so each wavenumber will be considered twice. At any given wavenumber, at least two band pass functions are non-zero - see Fig. 10. This approach is not a formal treatment, and so its effectivness will need to be demonstrated.

For now we will take $\Gamma(n)$ as a given (a form of $\Gamma(n)$, similar to that proposed in [28], is derived in appendix 6) and consider it to be a diagonal matrix in spectral space, used in association with the band pass functions, ie $\Psi_j^2 \rightarrow \Psi_j^2\Gamma$ when used in the $\mathbf{T}$-transform sense and $\Psi_j^2 \rightarrow \Gamma^{-1}\Psi_j^2$ in the $\mathbf{U}$-transform sense. It is not clear at this stage whether this modification should appear in both the vertical and horizontal calibration steps, or just one. We will assume for now that the correction should be used in both. Then, the formula for the vertical transform, Eq. 100, becomes,

$$\mathbf{B}_{vj}(\lambda,\phi) = \langle \mathbf{C}(\vec{r})\mathbf{F}_h\Psi_j^2\Gamma\mathbf{F}_h^{-1}\mathbf{T}_p^{\psi}\vec{x}'(\mathbf{C}(\vec{r})\mathbf{F}_h\Psi_j^2\Gamma\mathbf{F}_h^{-1}\mathbf{T}_p^{\psi}\vec{x}')^T\rangle, \tag{111}$$

and for the horizontal transform, Eq. 109, becomes,

$$\langle \mathbf{F}_h^{-1}\mathbf{T}_{vj}\mathbf{F}_h\Psi_j^2\Gamma\mathbf{F}_h^{-1}\mathbf{T}_p^{\psi}\vec{x}'(\mathbf{F}_h^{-1}\mathbf{T}_{vj}\mathbf{F}_h\Psi_j^2\Gamma\mathbf{F}_h^{-1}\mathbf{T}_p^{\psi}\vec{x}')^T\rangle \approx \Lambda_{hj}. \tag{112}$$

Calibration option 2b: triangular band pass functions with a band-dependent horizontal transform

The following approach makes no approximations to the calibration process other than making use of the approximate transform, Eq. 94 as before. Treating the band coupling properly gives rise to equations that are difficult to solve (i.e. a large rank problem that cannot be solved entirely by off-the-shelf numerical routines).

The vertical transforms are found as before by application of Eq. 100. The vertical transform matrices follow from Eqs. 101 to 103. The difficulty starts with calibration of the horizontal transform, which, incidently, is now necessarily band dependent.

Start with the definition of $\mathbf{B}$ (Eq. 62) and substitute for $\vec{x}'$ using the approximate $\mathbf{U}_{WS}$-transform (with an additional band index on the horizontal transform), Eqs. 2 and 94,

$$\langle \vec{x}'\vec{x}'^T\rangle \approx \langle \mathbf{U}_p\sum_{j=0}^{J}\mathbf{F}_h\Psi_j^2\mathbf{F}_h^{-1}\mathbf{U}_{vj}\mathbf{U}_{hj}\vec{v}'\vec{v}'^T(\mathbf{U}_p\sum_{j'=0}^{J}\mathbf{F}_h\Psi_{j'}^2\mathbf{F}_h^{-1}\mathbf{U}_{vj'}\mathbf{U}_{hj'})^T\rangle. \tag{113}$$

The angled brackets on the right hand side can be taken to the $\vec{v}'\vec{v}'^T$, giving $\langle\vec{v}'\vec{v}'^T\rangle = \mathbf{I}$,

$$\langle \vec{x}'\vec{x}'^T\rangle \approx \mathbf{U}_p\sum_{j=0}^{J}\mathbf{F}_h\Psi_j^2\mathbf{F}_h^{-1}\mathbf{U}_{vj}\mathbf{U}_{hj}(\mathbf{U}_p\sum_{j'=0}^{J}\mathbf{F}_h\Psi_{j'}^2\mathbf{F}_h^{-1}\mathbf{U}_{vj'}\mathbf{U}_{hj'})^T. \tag{114}$$

one parameter (ψ),

$$\langle \mathbf{F}_h^{-1}\mathbf{T}_p^{\psi}\vec{x}'(\mathbf{F}_h^{-1}\mathbf{T}_p^{\psi}\vec{x}')^T\rangle \approx \sum_{j=0}^{J}\Psi_j^2\mathbf{F}_h^{-1}\mathbf{U}_{vj}\mathbf{U}_{hj}(\sum_{j'=0}^{J}\Psi_{j'}^2\mathbf{F}_h^{-1}\mathbf{U}_{vj'}\mathbf{U}_{hj'})^T. \tag{115}$$

Multiply from the left with the operator $\hat{\Psi}_l^2$ and from the right with $\hat{\Psi}_{l'}^2$,

$$\langle \Psi_l^2\mathbf{F}_h^{-1}\mathbf{T}_p^{\psi}\vec{x}'(\Psi_{l'}^2\mathbf{F}_h^{-1}\mathbf{T}_p^{\psi}\vec{x}')^T\rangle \approx \sum_{j=0}^{J}\Psi_l^2\Psi_j^2\mathbf{F}_h^{-1}\mathbf{U}_{vj}\mathbf{U}_{hj}(\sum_{j'=0}^{J}\Psi_{l'}^2\Psi_{j'}^2\mathbf{F}_h^{-1}\mathbf{U}_{vj'}\mathbf{U}_{hj'})^T. \tag{116}$$

Equation 116 has a similar structure to a matrix equation, but where we have 'matrices of matrices' instead of the more familiar situation of matrices of numbers. It is apparent that the left-hand side is a $(J+1)\times(J+1)$ matrix of $N\times N$ matrices (or 'blocks'), where N is the number of grid-points in the model domain. Each block is found by varying l and l'.

It will simplify the procedure used to solve Eq. 116 if matrix blocks in this equation are defined as the following,

$$\begin{aligned} \text{let}: \; \tilde{R}_{ll'} &= \langle \Psi_l^2\mathbf{F}_h^{-1}\mathbf{T}_p^{\psi}\vec{x}'(\Psi_{l'}^2\mathbf{F}_h^{-1}\mathbf{T}_p^{\psi}\vec{x}')^T\rangle, & (117)\\ \tilde{\Psi}_{ll'} &= \Psi_l^2\Psi_{l'}^2, & (118)\\ \tilde{Q}_{jj'} &= \mathbf{F}_h^{-1}\mathbf{U}_{vj}\mathbf{U}_{hj}\mathbf{U}_{hj'}^T\mathbf{U}_{vj'}^T\mathbf{F}_h^{-T}. & (119)\end{aligned}$$

$\tilde{R}_{ll'}$, $\tilde{\Psi}_{ll'}$ and $\tilde{Q}_{jj'}$ are all $N\times N$ matrices. By varying l and l', Eq. 116 then resembles a matrix equation (of matrices),

$$\tilde{R}_{ll'} = \sum_{j=0}^{J}\sum_{j'=0}^{J}\tilde{\Psi}_{lj}\tilde{Q}_{jj'}\tilde{\Psi}_{j'l'}. \tag{120}$$

The $\tilde{\Psi}_{ll'}$ forms a tridiagonal block of diagonal matrices (it is tridiagonal because bands further than nearest neighbours do not overlap). It is beyond the scope of this report to discuss the method of solution of this equation, but the aim is to calculate the $\tilde{Q}_{jj'}$ matrices by solving Eq. 120. For the remainder of this report it is assumed that these matrices are now known. In fact, to compute the band-dependent horizontal spectra, the diagonal blocks of $\tilde{Q}$ will suffice. Continuing from the definition of $\tilde{Q}_{jj}$ from Eq. 119, we have,

$$\tilde{Q}_{jj} = \mathbf{F}_h^{-1}\mathbf{U}_{vj}\mathbf{U}_{hj}\mathbf{U}_{hj}^T\mathbf{U}_{vj}^T\mathbf{F}_h^{-T}. \tag{121}$$

The spectral transform on the left of the right-hand side and the adjacent vertical transform (which is now known from the previous subsection) can be taken to the left-hand side, followed by a transpose of the whole equation,

$$\begin{aligned} \mathbf{T}_{vj}\mathbf{F}_h\tilde{Q}_{jj} &= \mathbf{U}_{hj}\mathbf{U}_{hj}^T\mathbf{U}_{vj}^T\mathbf{F}_h^{-T}, & \\ \left(\mathbf{T}_{vj}\mathbf{F}_h\tilde{Q}_{jj}\right)^T &= \mathbf{F}_h^{-1}\mathbf{U}_{vj}\mathbf{U}_{hj}\mathbf{U}_{hj}^T. & (122)\end{aligned}$$

The remaining horizontal and vertical transforms, on the right-hand side in Eq. 122, can then be transferred to the left-hand side as above,

$$\mathbf{T}_{vj}\mathbf{F}_h\left(\mathbf{T}_{vj}\mathbf{F}_h\tilde{Q}_{jj}\right)^T = \mathbf{U}_{hj}\mathbf{U}_{hj}^T. \tag{123}$$

The left-hand side is known and the horizontal spectrum for band j is found by substituting the form for the horizontal transform from Eq. 55 (with an extra band index on the spectrum) on the right-hand side. Rearranged this gives,

$$\mathbf{F}_h^{-1}\left(\mathbf{F}_h^{-1}\mathbf{T}_{vj}\mathbf{F}_h\left(\mathbf{T}_{vj}\mathbf{F}_h\tilde{Q}_{jj}\right)^T\right)^T = \Lambda_{hj}. \tag{124}$$

Repeating this for every block matrix, $\tilde{Q}_{jj}$, along the diagonal of $\tilde{Q}$ for $0\le j\le J$ derives the horizontal spectra.

The role of an operator is to map one vector space to another. Typically in data assimilation, such operators fit into the 'forward part' of the data assimilation scheme, where each operator is part of a sequence whose ultimate aim is to predict the observations that are consistent with a control vector, $\vec{v}'$. Information is passed back to the control vector in reverse by the adjoint counterparts of each operator. The information that is passed back is in the form of gradients with respect to each vector, and the ultimate aim of this 'adjoint part' is to provide a gradient of the cost function with respect to components of $\vec{v}'$. This is then passed to a descent algorithm which decides how to adjust $\vec{v}'$.

In order to see how we need adjoint operators, consider the following example which uses the $\mathbf{U}$ operator (see Eq. 2). Since $\mathbf{U}$ is a linear operator, it may be represented as a matrix, and Eq. 2 may be expanded for the ith component of $\vec{x}'$,

$$\begin{aligned} \vec{x}' &= \mathbf{U}\vec{v}', \\ x_i' &= \sum_j \mathbf{U}_{ij}\vec{v}_j', \\ \text{thus} \quad \frac{\partial x_i'}{\partial v_j'} &= \mathbf{U}_{ij}. \end{aligned} \tag{125}$$

If we are provided with a vector of partial derivatives of a scalar, J, performed with respect to each component of $\vec{x}'$, then the chain rule as follows can be used to find the vector of partial derivatives of J with respect to each component of $\vec{v}'$,

$$\begin{aligned} \frac{\partial}{\partial v_j'} J &= \sum_i \frac{\partial x_i'}{\partial v_j'} \frac{\partial}{\partial x_i'} J, \\ &= \sum_i \mathbf{U}_{ij} \frac{\partial}{\partial x_i'} J, \\ &= \sum_i \mathbf{U}_{ji}^T \frac{\partial}{\partial x_i'} J, \\ \text{thus} \quad \left(\frac{\partial}{\partial \vec{v}'}\right)^T &= \mathbf{U}^T \left(\frac{\partial}{\partial \vec{x}'}\right)^T. \end{aligned} \tag{126}$$

The transpose, which is the same as the adjoint operator under certain conditions (see below), propagates the gradient information in the opposite direction to the forward operator.

The inner product between two vectors in $\vec{x}'$-space and in $\vec{v}'$-space are defined as the following [23],

$$\langle \vec{x}_1', \vec{x}_2' \rangle_x = \vec{x}_1'^T \mathbf{P_x} \vec{x}_2', \tag{127}$$

$$\langle \vec{v}_1', \vec{v}_2' \rangle_v = \vec{v}_1'^T \mathbf{P}_v \vec{v}_2', \tag{128}$$

where $\mathbf{P_x}$ and $\mathbf{P_y}$ are inner product matrices. In the left hand side of Eq. 127, substitute $\vec{x}_2' = \mathbf{U}\vec{v}_2'$. This can be written as follows, which can be regarded as the definition of the adjoint operator,

$$\langle \vec{x}_1', \mathbf{U}\vec{v}_2' \rangle_x = \langle \mathbf{U}^* \vec{x}_1', \vec{v}_2' \rangle_v \,. \tag{129}$$

In Eq. 129, $\mathbf{U}^*$ is the adjoint of $\mathbf{U}$. A similar inner product can be written for practically any operator, allowing its adjoint to be defined. Completing the inner products using Eqs. 127 and 128 yields,

$$\begin{aligned} \vec{x}_1'^T \mathbf{P_x} \mathbf{U} \vec{v}_2' &= (\mathbf{U}^* \vec{x}_1')^T \mathbf{P}_v \vec{v}_2', \\ \text{leading to} \quad \mathbf{U}^* &= \mathbf{P}_v^{-1} \mathbf{U}^T \mathbf{P_x}, \end{aligned} \tag{130}$$

which is the relationship between the adjoint and transpose operators (when deriving the second line from the first line, remember that the inner product matrices are symmetric). It is only in the event that the inner products are identity matrices that the adjoint is the same as the transpose, even though the terminology is often used interchangeably.

Let J be a scalar that is a functional of $\vec{x}'$. We note that the change, δJ, that follows from the change $\delta\vec{x}'$ has the form of a scalar product,

$$\delta J = \frac{\partial J}{\partial \vec{x}'} \delta \vec{x}', \tag{131}$$

already. Alternatively, we may write δJ in terms of the functional derivative, $\nabla_{\vec{x}'} J$, which is expressed as the following inner product equation,

$$\delta J = \langle \nabla_{\vec{x}'} J, \delta \vec{x}' \rangle_x = (\nabla_{\vec{x}'} J)^T \mathbf{P_x} \delta \vec{x}', \tag{132}$$

which has been written using the definition of the inner product in Eq. 127. The δJ has to be the same in Eqs. 132 and 131, and so $\nabla_{\vec{x}'} J$ must have a different meaning from $\partial J/\partial \vec{x}'$. Comparing the two equations yields the relationship between the functional and partial derivatives,

$$\nabla_{\vec{x}'} J = \mathbf{P_x}^{-1} \left(\frac{\partial J}{\partial \vec{x}'} \right)^T . \tag{133}$$

A similar relation exists between the functional and partial derivatives for vectors in $\vec{v}$-space,

$$\nabla_{\vec{v}'} J = \mathbf{P}_v^{-1} \left(\frac{\partial J}{\partial \vec{v}'} \right)^T . \tag{134}$$

Substituting the partial derivatives into the chain rule, Eq. 126, yields the following,

$$\begin{aligned} \mathbf{P}_v \nabla_{\vec{v}'} J &= \mathbf{U}^T \mathbf{P_x} \nabla_{\vec{x}'} J, \\ \nabla_{\vec{v}'} J &= \mathbf{P}_v^{-1} \mathbf{U}^T \mathbf{P_x} \nabla_{\vec{x}'} J, \\ \nabla_{\vec{v}'} J &= \mathbf{U}^* \nabla_{\vec{x}'} J, \end{aligned} \tag{135}$$

where in the last line, we have used the relationship between the adjoint and transpose operators (Eq. 130). Equation 135 is the chain rule for functional derivatives and is analogous to Eq. 126 for simple partial derivatives.

8 APPENDIX 2 - Covariance and convolution

It is often quoted that errors in the atmosphere have a horizontal structure which is approximately isotropic and homogeneous. This approximation is exploited in the horizontal transform, specifically in the form of Eq. 51. Another way of making this statement is that given that errors are isotropic and homogeneous in the horizontal, then modes (spherical harmonics in the case of a global model and plane waves in the case of a limited area model - the rows in $\mathbf{F}_h^{-1}$) are uncorrelated and the variances are the diagonal elements of Λ_h. Since this point is not explained in the main body of this paper, we reserve this appendix to give a plausibility argument towards its validity.

At the heart of the argument is the convolution theorem for fourier transforms. For simplicity we proceed to show our workings using the ordinary fourier transform in one dimension only, and assume that the result holds in the plane and for other bases. In particular we assume that there is a corresponding convolution theorem for a sperical harmonics-based transform.

A convolution is written as the following,

$$\tilde{g}(x) = \int_{x'} \mu(x' - x) g(x') \mathrm{d}x', \tag{136}$$

where $g(x)$ is the function being convoluted, and $\mu(x)$ is the function that specifies how $g(x)$ is to be smoothed. Fourier transforming each side of Eq. 136 and assuming that $\mu(x)$ is even, results in the following,

$$\tilde{G}(k) = M(k) G(k), \tag{137}$$

$$\text{where} \quad \tilde{G}(k) = \int_{x'} \tilde{g}(x) \exp(-\mathrm{i}kx) \mathrm{d}x, \tag{138}$$

$$M(k) = \int_{x'} \mu(x) \exp(-\mathrm{i}kx) \mathrm{d}x, \tag{139}$$

$$\text{and} \quad G(k) = \int_{x'} g(x) \exp(-\mathrm{i}kx) \mathrm{d}x, \tag{140}$$

where $\tilde{G}(k)$, $M(k)$ and $G(k)$ denote the fourier transforms of $\tilde{g}(x)$, $\mu(x)$ and $g(x)$ respectively. Equation 137 is called the convolution theorem and says that the convolution process is diagonal in fourier space.

function $g(x')$ is represented by the vector $\vec{g}$, then Eq. 136 is equivalent to acting on $\vec{g}$ with a covariance matrix, $\mathbf{M_x}$, whose rows are thought of as vectors, each being the function $\mu(x'-x)$, with x varying in each row (ie the rows are representations of the function $\mu(x')$, shifted in space by a different amount in each row of $\mathbf{M_x}$). The matrix form of Eq. 136 is,

$$\vec{\tilde{g}} = \mathbf{M_x}\vec{g}. \tag{141}$$

Let the vectors $\vec{G}$ and $\vec{\tilde{G}}$ be the fourier space vector representations of $\vec{g}$ and $\vec{\tilde{g}}$, respectively,

$$\vec{G} = \mathbf{F}_h^{-1}\vec{g}, \tag{142}$$

$$\vec{\tilde{G}} = \mathbf{F}_h^{-1}\vec{\tilde{g}}, \tag{143}$$

where $\mathbf{F}_h^{-1}$ is the horizontal spectral transform operator introduced in section 2.3, which represents the integrals in Eqs. 138, 139 and 140. Substituting the above two equations into Eq. 141 yields,

$$\vec{\tilde{G}} = \mathbf{F}_h^{-1}\mathbf{M_x}\mathbf{F}_h\vec{G}. \tag{144}$$

Comparing this equation to Eq. 137 tells us that $\mathbf{F}_h^{-1}\mathbf{M_x}\mathbf{F}_h$ (an operator in spectral space) is diagonal, and that the diagonal elements are found from the fourier transform of the smoothing function, μ. The equivalence of $\mathbf{F}_h^{-1}\mathbf{M_x}\mathbf{F}_h$ with $M(k)$ is essentially Eq. 51 in section 2.3 with $\mathbf{S}_h = \mathbf{M_x}$, and $\Lambda_h = M(k)$. Note that the property of errors being isotropic is reflected in our analysis by μ being an even function, and the homogeneous aspect is a natural consequence of the convolution.

9 APPENDIX 3 - Example of the implied covariances in the current transform

The control variable transforms, even in their current form, can be conceptually non-trivial to understand. In this appendix we analyse the combined horizontal and vertical transforms in the absense of the parameter transform, and their implied covariances.

Even though the current transforms are written as products (Eqs. 29 and 30), generally they imply non-separable structure functions (i.e. the structure functions cannot, in general, be written as a product of functions - one depending on height but not on horizontal position, and another on horizontal position but not on height). Some circumstances do lead to separable structure functions. We examine these below in the cases of two options for rotation.

9.1 Implied structure functions when $\mathbf{R}_v = \mathbf{I}$

Taking the $\mathbf{U}$-transform to be made up of the horizontal and vertical parts only, the $\mathbf{U}$-transform is written,

$$\begin{aligned} \vec{x}' &= \mathbf{U}_v\mathbf{U}_h\vec{v}', \\ &= \mathbf{Z}^{-1}\mathbf{M}^{1/2}\mathbf{F}_h\Lambda_h^{1/2}\vec{v}', \end{aligned} \tag{145}$$

where the same notation as in the main body of the text is used here (see in particular Eqs. 45 and 55), except that (due to the absence of the parameter transform in our arguments here) $\vec{x}'$ now refers to a parameter rather than a model variable. Equation 145 is more clearly written in integral notation,

$$x'(r,z) = \int \mathrm{d}\nu \mathbf{Z}^{-1}(z,\nu)\mathbf{M}^{1/2}(r,\nu) \int \mathrm{d}k \mathbf{F}_h(r,k)\Lambda_h^{1/2}(k,\nu)v'(k,\nu), \tag{146}$$

where r represents horizontal position, z height, k horizontal mode, and ν vertical mode index (in this integral form the symbols that were matrices are now functions of the two variables that the rows and columns of the matrix represented). The implied covariance, $\mathbf{UU}^T$ is written,

$$\begin{aligned} x'(r',z') = & \int \mathrm{d}\nu \mathbf{Z}^{-1}(z',\nu)\mathbf{M}^{1/2}(r',\nu) \int \mathrm{d}k \mathbf{F}_h(r',k)\Lambda_h^{1/2}(k,\nu) \times \\ & \int \mathrm{d}r \Lambda_h^{1/2}(k,\nu)\mathbf{F}_h^*(r,k) \int \mathrm{d}z \mathbf{M}^{1/2}(r,\nu)\mathbf{Z}^{-1}(z,\nu)x''(r,z). \end{aligned} \tag{147}$$

the horizontal modes use the complex conjugate, $\mathbf{F}_h^*$. The implied covariance maps each point in model space (at r and z in x'') to every other (at r' and z' in x'), just like a convolution (see appendix 2).

In order to see how this integral operator spreads information, let it act on the delta function $\delta(r-\tilde{r}, z-\tilde{z})$. Equation 147 then simplifies to,

$$\begin{aligned} x'(r',z') &= \int d\nu \mathbf{Z}^{-1}(z',\nu)\mathbf{M}^{1/2}(r',\nu) \int dk \mathbf{F}_h(r',k)\Lambda_h(k,\nu)\mathbf{F}_h^*(\tilde{r},k)\mathbf{M}^{1/2}(\tilde{r},\nu)\mathbf{Z}^{-1}(\tilde{z},\nu), \\ &= \int d\nu \mathbf{Z}^{-1}(z',\nu)\mathbf{M}^{1/2}(r',\nu)\mathbf{M}^{1/2}(\tilde{r},\nu)\mathbf{Z}^{-1}(\tilde{z},\nu) \times \\ &\quad \int dk \mathbf{F}_h(r',k)\Lambda_h(k,\nu)\mathbf{F}_h^*(\tilde{r},k), \end{aligned} \tag{148}$$

which has effectively traced-out a column of the implied covariance matrix - represented as a function - that represents the covariance between the point $(\tilde{r},\tilde{z})$ and every other (r',z') (such a function is called a structure function).

As it stands, Eq. 148 is non-separable. If, however, the r dependence of $\mathbf{M}$ and the ν dependence of Λ_h are removed, then the structure function becomes separable. This means that Eq. 148 can be written as a product of two terms - one that is a function of height but not of horizontal position, and another that is a function of horizontal position but not of height,

$$x'(r',z') = \int d\nu \mathbf{Z}^{-1}(z',\nu)\mathbf{M}(\nu)\mathbf{Z}^{-1}(\tilde{z},\nu) \times \int dk \mathbf{F}_h(r',k)\Lambda_h(k)\mathbf{F}_h^*(\tilde{r},k). \tag{149}$$

9.2 Implied structure functions when $\mathbf{R}_v = \mathbf{F}_v^T$

The other choice of $\mathbf{R}_v$ discussed involves a variation of EOF and rotation with horizontal position, $\mathbf{R}_v(\lambda,\phi) = \mathbf{F}_v^T(\lambda,\phi)$. The vertical co-ordinate output by the $\mathbf{T}_v$ transform is now height, and the vertical transforms (Eqs. 34 to 36) are used to investigate the nature of the structure functions. For simplified notation let the horizontal position dependence $(\lambda,\phi) = r$ of the vertical matrices be implicit in Eq. 150. The analogue here of Eq. 145 is,

$$\begin{aligned} \vec{x}' &= \mathbf{U}_v\mathbf{U}_h\vec{v}', \\ &= \mathbf{F}_v\Lambda_v^{1/2}\mathbf{F}_v^T\mathbf{F}_h\Lambda_h^{1/2}\vec{v}', \end{aligned} \tag{150}$$

In integral form, this $\mathbf{U}$-transform (with no parameter transform) is,

$$x'(r,z') = \int d\mu \mathbf{F}_v(z',\mu,r)\Lambda_v^{1/2}(\mu,r) \int dz \mathbf{F}_v^*(z,\mu,r) \int dk \mathbf{F}_h(r,k)\Lambda_h^{1/2}(k,z)v'(k,z), \tag{151}$$

where, due to the rotation, v' and Λ_h are now functions of z where they were a function of EOF mode in the previous section. The paraneter μ is a new variable for the local EOF index (ν was used before, which has become to mean height in this report when discussing the choice of rotation used here). The implied covariance, UU^T, is,

$$\begin{aligned} x'(r',z') &= \int d\mu' \mathbf{F}_v(z',\mu',r')\Lambda_v^{1/2}(\mu',r') \int dz'' \mathbf{F}_v^*(z'',\mu',r') \int dk \mathbf{F}_h(r',k)\Lambda_h^{1/2}(k,z'') \times \\ &\quad \Lambda_h^{1/2}(k,z'') \int dr \mathbf{F}_h^*(r,k) \int d\mu \mathbf{F}_v(z'',\mu,r)\Lambda_v^{1/2}(\mu,r) \int dz \mathbf{F}_v^*(z,\mu,r)x''(r,z). \end{aligned} \tag{152}$$

To derive the structure function, let this integral act on the same delta function as before. Let $x''(r,z) = \delta(r-\tilde{r}, z-\tilde{z})$ giving,

$$\begin{aligned} x'(r',z') &= \int d\mu' \mathbf{F}_v(z',\mu',r')\Lambda_v^{1/2}(\mu',r') \int dz'' \mathbf{F}_v^*(z'',\mu',r') \int dk \mathbf{F}_h(r',k)\Lambda_h^{1/2}(k,z'') \times \\ &\quad \Lambda_h^{1/2}(k,z'')\mathbf{F}_h^*(\tilde{r},k) \int d\mu \mathbf{F}_v(z'',\mu,\tilde{r})\Lambda_v^{1/2}(\mu,\tilde{r})\mathbf{F}_v^*(\tilde{z},\mu,\tilde{r}), \\ &= \int d\mu' \mathbf{F}_v(z',\mu',r')\Lambda_v^{1/2}(\mu',r') \int dz'' \mathbf{F}_v^*(z'',\mu',r') \int d\mu \mathbf{F}_v(z'',\mu,\tilde{r})\Lambda_v^{1/2}(\mu,\tilde{r})\mathbf{F}_v^*(\tilde{z},\mu,\tilde{r}) \times \\ &\quad \int dk \mathbf{F}_h(r',k)\Lambda_h(k,z'')\mathbf{F}_h^*(\tilde{r},k). \end{aligned} \tag{153}$$

the r dependence of Λ_v and $\mathbf{F}_v$ and the z dependence of Λ_h are removed, then the structure function becomes separable,

$$\begin{aligned} x'(r', z') &= \int \mathrm{d}\mu' \mathbf{F}_v(z', \mu') \Lambda_v^{1/2}(\mu') \int \mathrm{d}z'' \mathbf{F}_v^*(z'', \mu') \int \mathrm{d}\mu \mathbf{F}_v(z'', \mu) \Lambda_v^{1/2}(\mu) \mathbf{F}_v^*(\tilde{z}, \mu) \times \\ & \int \mathrm{d}k \mathbf{F}_h(r', k) \Lambda_h(k) \mathbf{F}_h^*(\tilde{r}, k). \end{aligned} \tag{154}$$

9.3 Variation of horizontal scale with horizontal position

It is believed - although more difficult to show - that the horizontal length scale is a (probably weak) function of horizontal position, at least in the first of the above cases. Although the horizontal transform is meant to be homogeneous (section 4.3), the presence of latitudinal dependences in the vertical transform gives rise to inhomogeneous structure functions when the horizontal and vertical transforms are combined. This may be shown, e.g., by looking at the width of the structure functions, Eq. 148 or 153 as the delta-function is moved around (in particular varied with latitude). This probably requires simplified numerical examples to show this in a straightforward manner.

10 APPENDIX 4 - Demonstration of some relations

10.1 Commutation between the vertical and horizontal transforms when $\mathbf{R}_v = \mathbf{I}$

Does the vertical U-transform in the case of no rotation of the vertical modes, $\mathbf{Z}^{-1}\mathbf{M}^{1/2}$, commute with the horizontal U-transform translated into grid space, $\mathbf{F}_h \Lambda_h^{1/2} \mathbf{F}_h^{-1}$? Let $v'(r', z)$ be the result of acting with the translated horizontal transform, and then the vertical transform on the state $v'(r, \nu)$,

$$v'(r', z) = \int \mathrm{d}\nu \mathbf{Z}^{-1}(z, \nu) \mathbf{M}^{1/2}(r', \nu) \int \mathrm{d}k \mathbf{F}_h(r', k) \Lambda_h^{1/2}(k, \nu) \int \mathrm{d}r \mathbf{F}_h^*(r, k) v'(r, \nu). \tag{155}$$

The vertical and horizontal U-transforms commute only if (i) Λ_h is no longer a function of ν and (ii) $\mathbf{M}$ is no longer a function of r'. In this case, the order of the transformations can be reversed,

$$\begin{aligned} v'(r', z) &= \int \mathrm{d}\nu \mathbf{Z}^{-1}(z, \nu) \mathbf{M}^{1/2}(\nu) \int \mathrm{d}k \mathbf{F}_h(r', k) \Lambda_h^{1/2}(k) \int \mathrm{d}r \mathbf{F}_h^*(r, k) v'(r, \nu), \\ &= \int \mathrm{d}k \mathbf{F}_h(r', k) \Lambda_h^{1/2}(k) \int \mathrm{d}r \mathbf{F}_h^*(r, k) \int \mathrm{d}\nu \mathbf{Z}^{-1}(z, \nu) \mathbf{M}^{1/2}(\nu) v'(r, \nu). \end{aligned} \tag{156}$$

10.2 Commutation between the vertical and bandpass functions when $\mathbf{R}_v = \mathbf{I}$

Does the vertical U-transform in the case of no rotation of the vertical modes, $\mathbf{Z}^{-1}\mathbf{M}^{1/2}$, commute with the bandpass functions translated into grid space, $\mathbf{F}_h \Psi_j^2(k) \mathbf{F}_h^{-1}$? Since this is the same analysis as the above (with $\Lambda_h^{1/2}$ replaced with Ψ_j^2), a similar conclusion applies. The vertical and bandpass functions commute only if $\mathbf{M}^{1/2}$ is no longer a function of r'.

10.3 Commutation between the vertical and horizontal transforms when $\mathbf{R}_v = \mathbf{F}_v^T$

Does the vertical U-transform in the case of rotation of the vertical modes back to height co-ordinates, $\mathbf{F}_v \Lambda_v^{1/2} \mathbf{F}_v^T$, commute with the horizontal U-transform translated into grid space, $\mathbf{F}_h \Lambda_h^{1/2} \mathbf{F}_h^{-1}$? Let $v'(r', z')$ be the result of acting with the translated horizontal transform, and then the vertical transform on the state $v'(r, z)$,

$$v'(r', z) = \int \mathrm{d}\mu \mathbf{F}_v(z', \mu, r') \Lambda_v^{1/2}(r', \mu) \int \mathrm{d}z \mathbf{F}_v^*(z, \mu, r') \int \mathrm{d}k \mathbf{F}_h(r', k) \Lambda_h^{1/2}(k, z) \int \mathrm{d}r \mathbf{F}_h^*(r, k) v'(r, z), \tag{157}$$

no longer a function of z and (ii) Λ_v and $\mathbf{F}_v$ are no longer functions of r'. In this case, the order of the transformations can be reversed,

$$\begin{aligned} v'(r',z') &= \int \mathrm{d}\mu \mathbf{F}_v(z',\mu)\Lambda_v^{1/2}(\mu) \int \mathrm{d}z \mathbf{F}_v^*(z,\mu) \int \mathrm{d}k \mathbf{F}_h(r',k)\Lambda_h^{1/2}(k) \int \mathrm{d}r \mathbf{F}_h^*(r,k) v'(r,z), \\ &= \int \mathrm{d}k \mathbf{F}_h(r',k)\Lambda_h^{1/2}(k) \int \mathrm{d}r \mathbf{F}_h^*(r,k) \int \mathrm{d}\mu \mathbf{F}_v(z',\mu)\Lambda_v^{1/2}(\mu) \int \mathrm{d}z \mathbf{F}_v^*(z,\mu) v'(r,z). \quad (158) \end{aligned}$$

10.4 Commutation between the vertical and bandpass functions when $\mathbf{R}_v = \mathbf{F}_v^T$

Does the vertical U-transform in the case of rotation of the vertical modes, $\mathbf{F}_v\Lambda_v^{1/2}\mathbf{F}_v^T$, commute with the bandpass functions translated into grid space, $\mathbf{F}_h\Psi_j^2(k)\mathbf{F}_h^{-1}$? Since this is the same analysis as the above (with $\Lambda_h^{1/2}$ replaced with Ψ_j^2), a similar conclusion applies. The vertical and bandpass functions commute only if Λ_v and $\mathbf{F}_v$ are no longer functions of r'.

10.5 The combination of operators $\mathbf{F}_h^{-1}\mathbf{U}_{vj}\mathbf{F}_h$ and $\mathbf{F}_h^{-1}\mathbf{T}_{vj}\mathbf{F}_h$

10.5.1 This combination when $\mathbf{R}_v = \mathbf{I}$

In section 6.3.4 we used the result that the combinations of operators $\mathbf{F}_h^{-1}\mathbf{U}_{vj}\mathbf{F}_h$ and $\mathbf{F}_h^{-1}\mathbf{T}_{vj}\mathbf{F}_h$ are diagonal in spectral space if $\mathbf{M}_j$ is independent of position. The $\mathbf{U}$ counterpart (first one of these) is demonstrated here by writing the combination in integral form, and acting on the function $x'(k,\nu)$. Using the form of the vertical transform from Eq. 45 in integral form (including the extra band index 'j' required for the WS transform),

$$\begin{aligned} x'(k',z) &= \int \mathrm{d}r \mathbf{F}_h^*(r,k') \int \mathrm{d}\nu \mathbf{Z}^{-1}(z,\nu)\mathbf{M}_j^{1/2}(\nu) \int \mathrm{d}k \mathbf{F}_h(r,k) x'(k,\nu), \\ &= \int \mathrm{d}k \int \mathrm{d}\nu \mathbf{Z}^{-1}(z,\nu)\mathbf{M}_j^{1/2}(\nu) \left(\int \mathrm{d}r \mathbf{F}_h^*(r,k')\mathbf{F}_h(r,k) \right) x'(k,\nu). \end{aligned} \quad (159)$$

The manipulation shown in the last line can be done only if $\mathbf{M}_j$ is independent of position. The functions comprising the $\mathbf{F}_h$, and inverse, operators will be mutually orthogonal. This means that the integral enclosed in brackets can be evaluated as,

$$\left(\int \mathrm{d}r \mathbf{F}_h^*(r,k')\mathbf{F}_h(r,k) \right) = \delta(k-k'), \quad (160)$$

where δ is the dirac delta function. This leaves Eq. 159 as,

$$x'(k',z) = \int \mathrm{d}\nu \mathbf{Z}^{-1}(z,\nu)\mathbf{M}_j^{1/2}(\nu) x'(k',\nu), \quad (161)$$

which is block diagonal in spectral space. The combination $\mathbf{F}_h^{-1}\mathbf{T}_{vj}\mathbf{F}_h$ can similarly be shown to be diagonal under the same circumstances.

In practice, $\mathbf{M}_j$, will be a function of position. As long as this function varies only slowly with position, then we would expect that the combinations of operators in question would be approximately diagonal, that is, any transfer of weight between different wavenumbers as a result of these operators would be negligible.

10.5.2 This combination when $\mathbf{R}_v = \mathbf{F}_v^T$

We repeat the analysis in section 10.5.1, but now for the case with rotation and using the vertical transform from Eq. 34. In integral form, with the added 'j' index required for the WS transform, and with no horizontal position dependence on components on the vertical transform, the combination of operators, $\mathbf{F}_h^{-1}\mathbf{U}_{vj}\mathbf{F}_h$ is,

$$\begin{aligned} x'(k',z) &= \int \mathrm{d}r \mathbf{F}_h^*(r,k') \int \mathrm{d}\mu \mathbf{F}_{vj}(z',\mu)\Lambda_{vj}^{1/2}(\mu) \int \mathrm{d}z \mathbf{F}_{vj}^*(z,\mu) \int \mathrm{d}k \mathbf{F}_h(r,k) x'(k,z), \\ &= \int \mathrm{d}k \int \mathrm{d}\mu \mathbf{F}_{vj}(z',\mu)\Lambda_{vj}^{1/2}(\mu) \int \mathrm{d}z \mathbf{F}_{vj}^*(z,\mu) \left(\int \mathrm{d}r \mathbf{F}_h^*(r,k')\mathbf{F}_h(r,k) \right) x'(k,z). \quad (162) \end{aligned}$$

position. The functions comprising the $\mathbf{F}_h$, and inverse, operators will be mutually orthogonal. This means that the integral enclosed in brackets can be evaluated as the delta-function in Eq. 160 above. This leaves Eq. 162 as,

$$x'(k',z) = \int d\mu \mathbf{F}_{vj}(z',\mu)\Lambda_{vj}^{1/2}(\mu) \int dz \mathbf{F}_{vj}^{*}(z,\mu)x'(k,z), \tag{163}$$

which is diagonal. The combination $\mathbf{F}_h^{-1}\mathbf{T}_{vj}\mathbf{F}_h$ can similarly be shown to be diagonal under the same circumstances.

In practice, $\mathbf{F}_{vj}$ and Λ_v, will be functions of position. As long as they vary only slowly with position, then we would expect that the combinations of operators in question would be approximately diagonal, that is, any transfer of weight between different wavenumbers as a result of these operators would be negligible.

11 APPENDIX 5 - Example of the implied covariances in the WS scheme

In appendix 3 we showed that under the current set of transforms the implied covariances are nearly or exactly separable, depending upon whether the vertical transform is dependent or independent of horizontal position. In the first section of this appendix, we perform a similar analysis with the proposed waveband summation (WS) transform. This is followed by a rough examination of the interplay between the positional and scale resolutions implied by the WS scheme.

11.1 Non-separability of the WS transform

The implied covariances, Eq. 8, can be calculated for the proposed form of the WS transform, Eq. 72. In a similar way to the method of appendix 3, we consider the implied covariances for one parameter only for the two possibilities of vertical transform discussed in this report.

11.1.1 Implied structure functions in WS when $\mathbf{R}_v = \mathbf{I}$

In an analagous way to Eq. 147, which is written in integral notation, the following holds for the WS transform when there is no rotation of the vertical transform, Eq. 45,

$$\begin{aligned} x'(r',z') \quad = \quad & \sum_{j'=0}^{J} \int d\nu \mathbf{Z}^{-1}(z',\nu)\mathbf{M}_{j'}^{1/2}(r',\nu) \int dk \mathbf{F}_h(r',k)\Lambda_h^{1/2}(k,\nu)\Psi_{j'}^2(n) \\ & \sum_{j=0}^{J} \Psi_j^2(n)\Lambda_h^{1/2}(k,\nu) \int dr \mathbf{F}_h^{*}(r,k)\mathbf{M}_j^{1/2}(r,\nu) \int dz \mathbf{Z}^{-1}(z,\nu)x''(r,z), \end{aligned} \tag{164}$$

where j, j' are band indices, J is the top band index, ν is the EOF index, z, z' label vertical coordinate, r, r' label horizontal position, k is horizontal wavenumber and n is total horizontal wavenumber (a function of k). The operators are described in sections 4.2 and 4.3. The implied covariance maps each point in model space (at r and z in x'') to every other (at r' and z' in x'), just like a convolution (see appendix 2).

Acting on a delta function spike at r'', z'' yields the structure functions implied by the WS transform,

$$\begin{aligned} x'(r',z') \quad = \quad & \sum_{j,j'=0}^{J} \int d\nu \mathbf{Z}^{-1}(z',\nu)\mathbf{M}_{j'}^{1/2}(r',\nu)\mathbf{M}_j^{1/2}(r'',\nu)\mathbf{Z}^{-1}(z'',\nu) \\ & \int dk \mathbf{F}_h(r',k)\Lambda_h(k,\nu)\Psi_{j'}^2(n)\Psi_j^2(n)\mathbf{F}_h^{*}(r'',k). \end{aligned} \tag{165}$$

The structure function of Eq. 165 is not separable, even if $\mathbf{M}$ is not a function of r and Λ_h is not a function of ν (unlike for the current scheme - see Eq. 149). Under these conditions, Eq. 165 becomes,

$$x'(r',z') = \sum_{j,j'=0}^{J} \int d\nu \mathbf{Z}^{-1}(z',\nu)\mathbf{M}_{j'}(\nu)\mathbf{Z}^{-1}(z'',\nu) \int dk \mathbf{F}_h(r',k)\Lambda_h(k)\Psi_{j'}^2(n)\Psi_j^2(n)\mathbf{F}_h^{*}(r'',k). \tag{166}$$

separable functions, however, is not itself a separable function in general. Hence the WS transform leads to fundamentally non-separable functions.

11.1.2 Implied structure functions in WS when $\mathbf{R}_v = \mathbf{F}_v^T$

In an analagous way to Eq. 147, which is written in integral notation, the following holds for the WS transform when there is rotation of the vertical transform, Eq. 34,

$$
\begin{aligned}
x'(r',z') \quad = \quad & \sum_{j'=0}^{J} \int \mathrm{d}\mu' \mathbf{F}_{vj'}(z',\mu',r')\Lambda_{vj'}^{1/2}(\mu',r') \int \mathrm{d}z''' \mathbf{F}_{vj'}^{*}(z''',\mu',r') \int \mathrm{d}k \mathbf{F}_h(r',k)\Lambda_h^{1/2}(k,z''')\Psi_{j'}^{2}(n) \\
& \sum_{j=0}^{J} \Psi_j^2(n)\Lambda_h^{1/2}(k,z''') \int \mathrm{d}r \mathbf{F}_h^{*}(r,k) \int \mathrm{d}\mu \mathbf{F}_{vj}(z''',\mu,r)\Lambda_{vj}^{1/2}(\mu,r) \\
& \int \mathrm{d}z \mathbf{F}_{vj}^{*}(z,\mu,r)x''(r,z), \qquad (167)
\end{aligned}
$$

where j,j' are band indices, J is the top band index, μ is the EOF index, z,z',z''' label vertical coordinate, r,r' label horizontal position, k is horizontal wavenumber and n is total horizontal wavenumber (a function of k). The operators are described in sections 4.2 and 4.3. The implied covariance maps each point in model space (at r and z in x'') to every other (at r' and z' in x'), just like a convolution (see appendix 2).

Acting on a delta function spike at r'',z'' yields the structure functions implied by the WS transform,

$$
\begin{aligned}
x'(r',z') \quad = \quad & \sum_{j,j'=0}^{J} \int \mathrm{d}\mu' \mathbf{F}_{vj'}(z',\mu',r')\Lambda_{vj'}^{1/2}(\mu',r') \int \mathrm{d}z''' \mathbf{F}_{vj'}^{*}(z''',\mu',r') \\
& \int \mathrm{d}k \mathbf{F}_h(r',k)\Lambda_h(k,z''')\Psi_{j'}^{2}(n)\Psi_j^2(n) \\
& \mathbf{F}_h^{*}(r'',k) \int \mathrm{d}\mu \mathbf{F}_{vj}(z''',\mu,r'')\Lambda_{vj}^{1/2}(\mu,r'')\mathbf{F}_{vj}^{*}(z'',\mu,r''). \qquad (168)
\end{aligned}
$$

The structure function of Eq. 168 is not separable, even if $\mathbf{F}_{vj}$ and Λ_{vj} are not functions of r and Λ_h is not a function of z (unlike for the current scheme - see Eq. 149). Under these conditions, Eq. 168 becomes,

$$
\begin{aligned}
x'(r',z') \quad = \quad & \sum_{j,j'=0}^{J} \int \mathrm{d}\mu' \mathbf{F}_{vj'}(z',\mu')\Lambda_{vj'}^{1/2}(\mu') \int \mathrm{d}z''' \mathbf{F}_{vj'}^{*}(z''',\mu') \int \mathrm{d}\mu \mathbf{F}_{vj}(z''',\mu)\Lambda_{vj}^{1/2}(\mu)F_{vj}^{*}(z'',\mu) \\
& \int \mathrm{d}k \mathbf{F}_h(r',k)\Lambda_h(k)\Psi_{j'}^{2}(n)\Psi_j^2(n)\mathbf{F}_h^{*}(r'',k). \qquad (169)
\end{aligned}
$$

Each component inside the double summation comprises a separable function in this case. A sum of separable functions, however, is not itself a separable function in general. Hence the WS transform leads to fundamentally non-separable functions.

11.2 Interplay between positional and scale resolutions

In the current transform, there is no explicit horizontal scale dependence of the vertical transform - just horizontal position dependence. Although the new WS transform introduces a scale dependence to the structure functions, we show here that this comes - unsurprisingly - at the expence of the ability to capture positional dependent features. The analysis here is rough, and is presented only for illustrative purposes.

We examine the structure function associated with a δ-function at position r'',z'', as in Eq. 168, but - for simplicity - in the case of a single vertical level. This has the consequence that $\mathbf{F}_{vj} = 1$ (or $\mathbf{Z} = 1$, depending upon the choice of vertical transform), and that there is no vertical position or mode dependence (and hence no vertical integrals). The presence of only one level means that the result that follows can be interpreted in terms of either of the vertical transform options considered in this report. The analysis below is written in the framework of the simple vertical transform in section 4.2.2 involving Λ_{vj}. For the current vertical transform (section 4.2.3), simply replace $\Lambda_{vj}^{1/2}$ with $\mathbf{M}_{vj}^{1/2}$ below. We make to and from spectral space is achieved through a fourier transform. We maintain, however, the positional dependence of Λ_{vj}. Equation 168 is then,

$$x'(r') = \sum_{j,j'=0}^{J} \Lambda_{vj'}^{1/2}(r') \int \mathrm{d}k \mathbf{F}_h(r',k)\Lambda_h(k)\Psi_{j'}^2(n)\Psi_j^2(n)\mathbf{F}_h^*(r'',k)\Lambda_{vj}^{1/2}(r''). \tag{170}$$

In order to reduce this equation, let the bandpass functions approximate to non-overlapping top-hat functions. For function j, let it have the value unity inside the bounds N_j to N_{j+1}, and zero elsewhere. This orthogonality leads to Eq. 170 simplifying to,

$$x'(r') = \sum_{j=0}^{J} \Lambda_{vj}^{1/2}(r')\Lambda_{vj}^{1/2}(r'') \int_{N_j}^{N_{j+1}} \mathrm{d}k \mathbf{F}_h(r',k)\mathbf{F}_h^*(r'',k)\Lambda_h(k). \tag{171}$$

There is only one band summation now because the product of bandpass functions $\Psi_{j'}^2(n)\Psi_j^2(n)$ are non-zero only when $j = j'$. Using fourier transforms, Eq. 171 becomes,

$$x'(r') = \sum_{j=0}^{J} \Lambda_{vj}^{1/2}(r')\Lambda_{vj}^{1/2}(r'') \int_{N_j}^{N_{j+1}} \mathrm{d}k \exp ik(r'-r'')\, \Lambda_h(k). \tag{172}$$

Let $y = r' - r''$, and assume that the the range $N_{j+1} - N_j = \Delta k$ is so small that the horizontal variances, $\Lambda_h(k)$, vary very little over that range,

$$\begin{aligned}
x'(y) &\approx \sum_{j=0}^{J} \Lambda_{vj}^{1/2}(y+r'')\Lambda_{vj}^{1/2}(r'')\Lambda_h(k_j)\frac{1}{ix}[\exp ikx]_{N_j}^{N_{j+1}}, \\
&\approx -\sum_{j=0}^{J} \Lambda_{vj}^{1/2}(y+r'')\Lambda_{vj}^{1/2}(r'')\Lambda_h(k_j)\frac{i\exp i\Delta kx}{x}, \\
Re(x'(y)) &\approx \sum_{j=0}^{J} \Lambda_{vj}^{1/2}(y+r'')\Lambda_{vj}^{1/2}(r'')\Lambda_h(k_j)\frac{\sin \Delta kx}{x}, \\
&\approx \sum_{j=0}^{J} \Lambda_{vj}^{1/2}(y+r'')\Lambda_{vj}^{1/2}(r'')\Lambda_h(k_j)\Delta k,
\end{aligned} \tag{173}$$

where k_j is the typical wavenumber inside the range N_j to N_{j+1} (e.g., $k_j = (N_{j+1} - N_j)/2$), and where we assume in the last line that $\Delta kx \ll 1$.

This expression for the structure function (Eq. 173) is valid only when the number of bands is large. Equation 173 has the form of a linear combination of functions $\Lambda_{vj}^{1/2}(y + r'')$, with expansion coefficients $\Lambda_{vj}^{1/2}(r'')\Lambda_h(k_j)\Delta k$. If each $\Lambda_{vj}^{1/2}(y + r'')$ has a different spacial variation, then the linear combination will smear out the contributions from each scale, with a greater smearing effect with the greater number of bands.

12 APPENDIX 6 - The spectral correction function

When calibrating the WS transform, the overlap of the triangular band pass functions complicates the process. One way of dealing with the overlap that has been proposed is to treat the calibration equations as though there is no overlap, apart from the inclusion of a spectral correction factor, Γ. This is the approach of calibration option 2a, which is described in section 6.5.2. A proposed form of Γ is developed in this appendix, based on that shown in [28]. We end this appendix with a note on whether we do actually need to use a correction function at all.

The 'Andrews and Lorenc' method

Andrews and Lorenc [28] derive similar functions which here are called the spectral correction functions. These are designed to correct for the 'double counting' that occurs due to the overlap of the triangular spectral band pass functions.

each wavenumber within band j according to the weighted sum,

$$\text{Variance of } \psi \text{ for band } j = \int_{N_{j-1}}^{N_{j+1}} \mathrm{d}n \; w_j(n) \times \text{Variance of } \psi \text{ at wavenumber } n, \tag{174}$$

where $w_j(n)\mathrm{d}n$ is the weight applied to each individual wavenumber in the integral. This weight is given by the assumed form,

$$w_j(n) = \frac{\Psi_j^2(n)\Gamma(n)}{\int_{N_{j-1}}^{N_{j+1}} \mathrm{d}n' \Psi_j^2(n')}. \tag{175}$$

(Note that in [28], $\Psi_j^4(n)$ has been used instead of $\Psi_j^2(n)$ as in the above. It is not clear which is correct.) For any given total wavenumber, n, the weights must sum to unity and $\Gamma(n)$ in Eq. 175 is there as a 'fudge-factor' to make this happen. Enforcing this gives a formula for $\Gamma(n)$,

$$\begin{aligned} \sum_{j=0}^{J} w_j(n)\delta n &= 1, \\ \therefore \Gamma(n) &= \left(\sum_{j=0}^{J} \frac{\Psi_j^2(n)\delta n}{\int_{N_{j-1}}^{N_{j+1}} \mathrm{d}n' \Psi_j^2(n')} \right)^{-1}. \end{aligned} \tag{176}$$

In the above, δn represents the area of wavenumber-space occupied by wavenumber n. The fact that this is δn, rather than $\mathrm{d}n$ reflects the fact that our problem is discrete (strictly speaking the normalization integrals in the denominator should be summations rather than integrals). At any given total wavenumber, there are a maximum of two bands that are non-zero,

$$\Gamma(n) = \begin{cases} \left(\frac{\Psi_{j-1}^2(n)\delta n}{\int_{N_{j-2}}^{N_j} \mathrm{d}n' \Psi_j^2(n')} + \frac{\Psi_j^2(n)\delta n}{\int_{N_{j-1}}^{N_{j+1}} \mathrm{d}n' \Psi_j^2(n')} \right)^{-1} & \text{for} \quad N_{j-1} \leq n \leq N_j \\ \left(\frac{\Psi_j^2(n)\delta n}{\int_{N_{j-1}}^{N_{j+1}} \mathrm{d}n' \Psi_j^2(n')} + \frac{\Psi_{j+1}^2(n)\delta n}{\int_{N_j}^{N_{j+2}} \mathrm{d}n' \Psi_j^2(n')} \right)^{-1} & \text{for} \quad N_j < n \leq N_{j+1} \end{cases}, \tag{177}$$

Γ is used when calibrating the transforms and is used in Eq. 112. There are two possible problems with this result:

- We are correcting for overcounting and so it seems sensible that $\Gamma < 1$. This is found not to be the case (see Fig. 12, solid curve).
- If the trangular band pass functions in the above procedure are replace by non-overlaping functions (such as top-hat functions), then there is no correction required and Γ should be unity at all wavenumbers. As a test, we find that this is not the case (see Fig. 12, dashed curve).

It is not certain whether these findings prevent the spectral correction factor from being used. Certainly Γ in its current form is not suitable, but may give better behaviour with modifications.

Do we need a spectral correction factor?

The above derivation is not definitive. The assumed form of $w_j(n)$, Eq. 175, is subjective. If the width of a band is broadened, it should as a whole carry more variance in Eq. 174, yet the increased denominator in Eq. 175 will 'normalize-away' this effect (approximately). In other words, the above definitions do not allow the variances to scale properly with band width. Here we argue that perhaps we do not need a correction function at all.

Let us repeat the derivation, but with a different definition of $w_j(n)$ than Eq. 175. Replace Eq. 175 with the following simpler form, that does allow band scaling in a qualitatively reasonable fashion,

$$w_j(n) = \Psi_j^2(n)\Gamma(n). \tag{178}$$

Once again, we find the $\Gamma(n)$ that gives a unit sum of weights,

$$\begin{aligned} \sum_{j=0}^{J} w_j(n) &= 1, \\ \therefore \Gamma(n) &= \left(\sum_{j=0}^{J} \Psi_j^2(n) \right)^{-1}. \end{aligned} \tag{179}$$

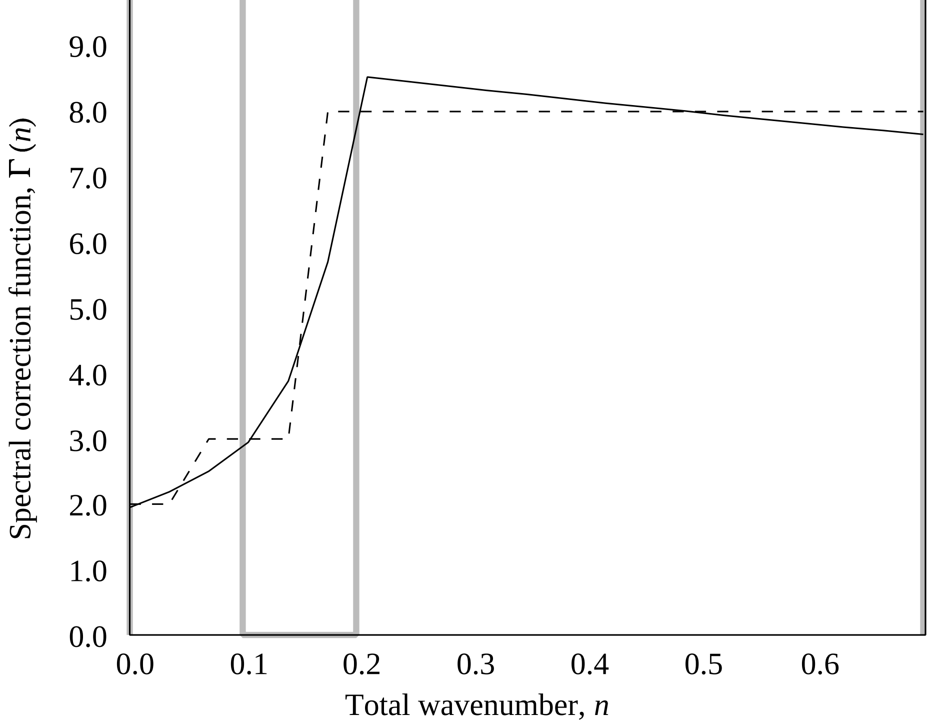

Figure 12: Example calculation of the $\Gamma(n)$ spectral correction functions for the case of triangular band pass functions (solid curve) and top-hat band pass functions (dashed curve). For simplicity, this example has been derived from a two-dimensional situation, with just latitude as the horizontal coordinate; total wavenumber has been replaced by meridional wavenumber. There are three bands with the N_j positions shown as the light, thick vertical lines (these lie also at the left- and right-most end of the plot, and are partially obscured by the left and right boundaries of the graph).

Thus, due to Eq. 74, Γ is equal to unity. It can be argued then that the double counting is compensated for, approximately, by the band pass functions themselves, leaving no need for an additional factor.

13 ACKNOWLEDGEMENTS

I would like to thank the following people for useful discussions which have helped me, in one way or another, with this document: Mike Cullen, Sarah Dance, Mark Dixon, Bruce Ingleby, Andrew Lorenc, Nancy Nichols, David Pearson, Ian Roulstone, Olaf Stiller and Andy White, plus anyone who I have failed to include!

References

[1] Lorenc A.C., Ballard S.P., Bell R.S., Ingleby N.B., Andrews P.L.F., Barker D.M., Bray J.R., Clayton A.M., Dalby T., Li D., Payne T.J., Saunders F.W., The Met Office global 3-dimensional variational data assimilation scheme, Q.J.R.Meteor.Soc. **126**, pp. 2991-3012 (2000).

[2] Barker D.M., The use of synoptically dependent background error structures in 3d-Var, Var. Scientific Documentation Paper No. 26 (1999).

[3] Semple A., A Meteorological Assessment of the Geostrophic Co-ordinate Transform and Error Breeding System when used in 3D Variational Data Assimilation JCMM Internal Report 130 and NWP Technical Report 357 (2000/2001).

[4] Desroziers G., A coordinate change for data assimilation in spherical geometry of frontal structures, Mon. Wea. Rev. **125**, pp. 3030-3038 (1997).

[5] Dubal M.R., Use of a geostrophic coordinate transform in 3d-Var (2001).

Report (2003).

[7] Andrews P.L.F., An introduction to the UKMO variational data analysis scheme, available via the Var homepage (1998).

[8] Ingleby N.B., Estimation of background error statistics, Var. Scientific Documentation Paper 20.

[9] Horizonal transforms, Var. Scientific Documentation Paper 14 (2000).

[10] Lanczos C., Applied analysis, Dover Publications Inc., New York (1988).

[11] Bannister R.N., Finite vector space representations, "http://www.met.rdg.ac.uk/ ross/DARC/LinearVectorSpaces.html" (2003).

[12] Tarantola A., Inverse problem theory - methods for data fitting and model parameter estimation (Elsevier, Amsterdam, 1994), 2nd ed., section xx.

[13] Parrish D.F., Derber J.C., The National Meteorological Center's spectral statistical-interpolation analysis system, Mon. Wea. Rev. **120**, pp. 1747-1793 (1992).

[14] Swinbank R., Ingleby N.B., Boorman P.M., Renshaw R.J., A 3D variational data assimilation system for the stratosphere and troposphere, Forecasting Research Paper 71 (2002).

[15] Control variable transforms - Parameters, Var. Scientific Documentation Paper 11 (1995).

[16] Roulstone I., White A., Byrom M., Control variable transforms, available via "http://vulcan/ apir/home.html" (2002).

[17] Andrews P.L.F., Covariance development strategy: a three point plan for improving the estimation and representation of background error covariances in Var., Var. working paper 20 (2002).

[18] Ingleby N.B., Options for improving and modelling of background errors in the Met Office variational analysis (2002).

[19] Payne T.J., Advantages of PV as principal control variable, available via "http://fr2000/ frtp/PV.html" (2000).

[20] Control variable transforms - Vertical transforms, Var. Scientific Documentation Paper 13 (1999).

[21] Lorenc A.C., Iterative analysis using covariance functions and filters, Q.J.R.Meteor.Soc. **118**, 569-591 (1992).

[22] Barker D., The Spectral Representation of Control Variables in Var. (1995).

[23] Top level, Var. Scientific Documentation Paper 2 (1995).

[24] Bannister R.N., Multiple outer loops in Var. without the T-transform (2004). Available via

`"http://www.met.rdg.ac.uk/~ross/DARC/No_T_trans.ps.gz"`

[25] Parrish D.F., Derber J.C., The National Meteorological Centers' spectral statistical interpolation analysis system, Mon. Wea. Rev. **120**, pp. 1747-1763 (1992).

[26] Fisher M., Background error covariance modelling, in ECMWF seminar proceedings (2003).

[27] Ingleby N.B., The statistical structure of forecast errors and its representation in the Met Office global 3-dimensional variational data assimilation scheme, Q.J.R.Meteor.Soc. **127**, 209-231 (2001). Also available via "http://www-nwp/ frbi/papers/home.html".

[28] Andrews P.L.F., Lorenc A.C., On introducing a multi-scale waveband-summation covariance model into the Met Office Var system, Var. Working Paper 19 (2002).

[29] Ingleby N.B., Global horizontal spectra for 3d-Var, available via Var. Scientific Documentation Paper 20 (2002).