U.S. Department of Commerce
National Oceanic and Atmospheric Administration
National Weather Service
National Centers for Environmental Prediction
5200 Auth Road
Camp Springs, MD 20746-4304

Office Note 458

SINGULARITY-FREE GLOBAL TELESCOPING GRIDS SUITABLE FOR LARGE-SCALE GLOBAL COVARIANCE SYNTHESIS BY ANISOTROPIC RECURSIVE FILTERING

Yoshiaki Sato*
NCEP Visiting Scientist on leave from Japan Meteorological Agency

R. James Purser†
Science Applications International Corp., Beltsville, Maryland

April 27th, 2009

* email: y-sato@met.kishou.go.jp
† email: jim.purser@noaa.gov

Abstract

Background error covariances in the Gridpoint Statistical Interpolation (GSI) of NCEP's operational three-dimensional data assimilation schemes are formed by the repeated applications of recursive numerical filters along carefully selected families of parallel lines of the discretization grids used to represent the analysis state. Owing to the need for the analysis grid to be free of coordinate singularities in order that the outputs of the recursive filters remain smooth, the global version of the GSI presently employs a system of three overlapping grids. One grid patch is a cylindrical coordinate projection that covers most of the globe but excludes the poles; then a pair of stereographic projection Cartesian grid patches cover the omitted polar regions, but with a generous overlap across the edges of the cylindrical grid patch. In this way the global GSI can form portions of the analysis increments pertaining to each of the singularity-free patches with the final composition blended to remove any sharp discontinuity between them. However, for the very largest scales of covariances encountered in the highest levels of the atmosphere the width of the blending zone shared by each pair of overlapping patches is not sufficient to avoid an inevitable distortion of the shape of the synthesized covariance from the shape intended.

This note describes a new grid configuration, designed to supplement the existing three-patch configuration, that allows very large scale covariances to be synthesized by the recursive filtering technique without unintended distortion. The basic idea is to map the Earth's atmosphere, regarded as a spherical shell of finite thickness, to the interior of a set of three-dimensional cubic grids in a nested configuration that allows the resolution to be increased (usually by progressive factors of two) towards the common center of the cubes. The sizes of the successive nests, and the radial transformation carrying the atmospheric shell to the cubes, are both under the user's control so that different vertical-to-horizontal aspect ratios of covariances at different altitudes can all be represented by approximately isotropic covariance shapes within the cubic grids. Although the cubic grids are free of singularities, they do possess interior boundaries which must be treated by smooth blending of only the interior contributions constructed on each of these nests, and by blending the totality with the covariance contributions synthesized on the original three-patch grid configuration.

We present the theory for the construction of the nested cubic grids, together with some preliminary results obtained by adopting this approach in an experimental extension of the global GSI.

1. Introduction

The global synthesis of background error covariances in NCEP's GSI is presently done by performing filtering operations on three overlapping domains and merging the results. The domains are:

(i) a southern polar stereographic domain with Cartesian grid coordinates;

(ii) a central domain with cylindrical grid coordinates;

(iii) a northern polar stereographic domain similar to its southern counterpart.

An alternative division of the global domain, which also avoids the polar singularities, is the "Yin-Yang" configuration (Kageyama and Sato 2004, Peng et al. 2004):

(i) a south-pole-straddling rotated lat-long grid whose "equator" occupies Earth's +90 ° and -90 ° meridians;

(ii) a north-pole-straddling rotated lat-long grid whose "equator" occupies Earth's 0 ° and 180 ° meridians.

The purpose of dividing the region in these ways is to ensure that the numerical operations of recursive filters do not encounter coordinate singularities or regions of strong coordinate curvature (for which the filters are not well suited). The Yin-Yang grid configuration would be one way of attaining a modest improvement in computational economy.

Recursive filters generate quasi-Gaussian covariance shapes while real covariances exhibit greater variety. Therefore it is desirable to combine additively several Gaussian contributions of different spatial sizes and appropriate amplitudes to produce the desired covariance. (In practice, it is better to produce the square-root, C, of the covariance, B, by this superposition and form $B = CC^{\mathsf{T}}$ to ensure that the property of 'self-adjointness', necessary for the success of many descent algorithms, is preserved exactly.) It is wasteful to apply coarse-scale recursive (or any other) filters on a fine grid. It is better that coarse-scale filtering be performed on a grid just fine enough to resolve, and to interpolate from, the scales output by the filter. The simpler operation of interpolating to a finer grid, normally in successive stages of binary refinement, costs less than filtering and is a small overhead. When many scales of filtering are combined, the overall cost only slightly exceeds the cost of filtering one or two separate scales at the finest grid of the hierarchy. Nevertheless, the coarser versions of the above-mentioned overlapping grids need correspondingly larger overlaps and, if covariances with contributions of planetary scales are needed, it is no longer possible to ignore the hitherto distant coordinate singularities of these grids. This is particularly likely to pose a problem for stratosphere/mesosphere data assimilation, where the predominant horizontal scales are already very large and not well served by the existing overlapping grid configuration. In the next section we propose a remedy based on a very different kind of global gridding which exploits the full three-dimensional nature of the relevant geometry but abandons the usual convention by which one 'vertical' coordinate approximately adheres to horizontal spherical surfaces.

2. CUBIC GRIDDING

Instead of separating the quasi-horizontal levels by one 'vertical' coordinate and expecting two other 'horizontal' coordinates to cover the spherical surface, we transform the three-dimensional spherical shell occupied by the atmosphere to the interior of an abstract cube which can then be gridded in Cartesian fashion to avoid any coordinate singularity whatsoever. This somewhat radical solution, while woefully inadequate as a strategy for modeling, is perfectly satisfactory as a means for supplying a grid on which anisotropic filtering can proceed unmolested by singularities (which pose a greater threat than does the lack of stratification). Obviously, embedding the atmosphere directly and without metrical distortion is impractical since the atmosphere's effective thickness is tiny (a few tens of kilometers) compared to Earth's radius (a few thousand kilometers). On the other hand, if we radially transform the thin atmospheric shell to a very much thicker one, comparable to the Earth's radius, we make fuller use of

the cubic grid. However, if transforming to a very thick image shell, the lower atmosphere will be resolved relatively poorly (in the horizontal direction) compared to the upper atmosphere. It would be preferable to confer on the upper portion of the atmosphere the lower resolution in the horizontal directions, which suggests that 'inverting' the radial ordering in the transformation would be a better arrangement. Another way of exerting control over the resolution is to refine the inner portion of the cubic grid where the resolution is otherwise deficient and, if necessary, to repeat such refinements in 'telescoping' fashion to create a hierarchic series of successively refined grids nested, each one within its predecessor, like Russian matryoshka dolls. Trimming the corners, we might then have a cubic grid configuration analogous to the two-dimensional square grid depiction given in Fig. 1. Each grid can support a filtering operation with a characteristic scale suitable to that grid and, provided an amplitude-modulating function causes the amplitude of each grid's filter to drop smoothly to zero at or before the grid's boundary, the contributions of all the nested grids can be added up when interpolated back to the principal overlapping analysis grids used to cover the globe.

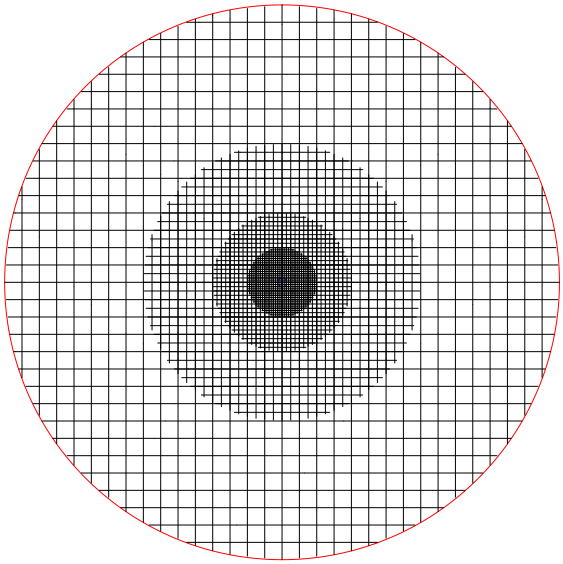

Figure 1. Nested Cartesian grids, with corners trimmed, providing enhanced resolution near the center.

One very natural radial transformation for a telescoping grid is by the logarithm function. If each nested grid has half the total size but twice the resolution of its predecessor in the hierarchy, as is the case shown in Fig. 1, the logarithmic transformation makes this succession exactly periodic in transformed radius across the extent of the 'generations' of these grids. This periodicity can be seen over the four generations of the grids shown in the schematic view of Fig. 2 where the atmospheric shell's thickness is obviously enormously exaggerated. However, the lower part of the 'atmosphere' clearly shows the degraded resolution resulting from the premature termination of the nesting sequence. When we turn the transformation inside-out, as is done in Fig. 3 by placing a minus sign before the logarithm function, we see that the

reduction in resolution, corresponding to the innermost region of the original nested squares, now occurs at the 'high-atmosphere' portion of the cartoon-atmosphere, which is probably more in line with the vertical profile of horizontal resolution that we would intend in practice.

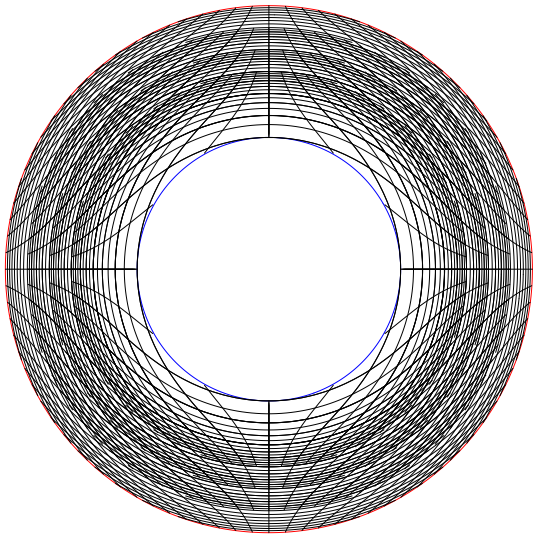

Figure 2. A direct logarithmic radial transformation of the nested Cartesian grids of Fig. 1, to the finite-thickness spherical shell of the global atmosphere, including an exaggeration of the vertical thickness in order to see how the transformed grid resolution varies with height. The region of degraded resolution caused by terminating the grid nesting hierarchy too soon, occurs near the bottom of the 'atmosphere'. Additional nests would clearly be needed to maintain the higher resolution down to the bottom.

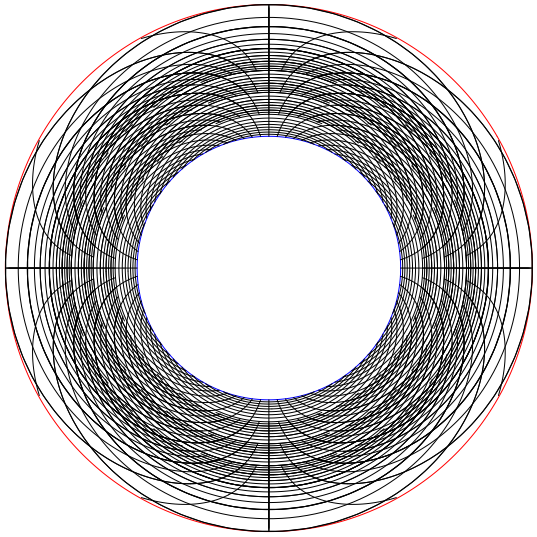

Figure 3. An inverted (negative) logarithmic radial transformation of the nested Cartesian grids of Fig. 1, to the same 'atmosphere' as shown in Fig. 2. The region of degraded resolution is now near the top, which is probably more acceptable, since the resolution at the highest levels of the atmosphere need not be as fine as at lower levels.

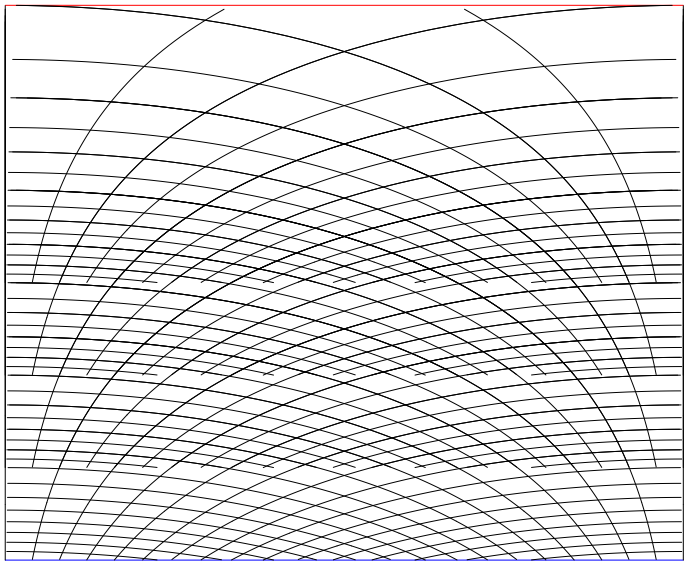

Figure 4. The same inverted (negative) logarithmic radial transformation of the nested Cartesian grids as shown in Fig. 3, but now displayed for one quadrant in a more conventional horizontal-vertical projection.

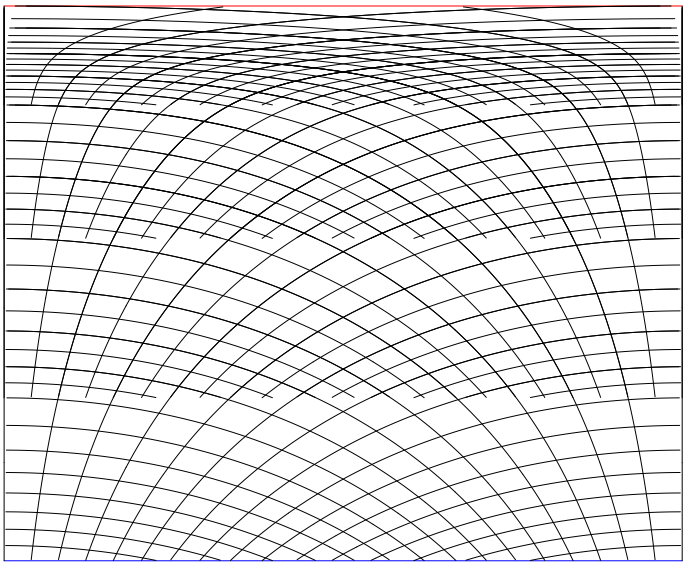

Figure 5. A non-logarithmic radial transformation of the nested Cartesian grids of Fig. 1, showing the profile of the vertical resolution can be controlled to some extent by the choice of the radial transformation function.

A more conventional horizontal-vertical plot of the same inverted mapping of Fig. 3 is shown for a single quadrant in Fig. 4. While the high altitude degradation of horizontal resolution seen in Fig. 4 may be acceptable, it may be desirable to maintain, or even enhance, the vertical resolution at altitudes of strong stratification. This can be done by departing from the strictly logarithmic transformation to some appropriate modification of it, as shown in this case in Fig.

5. But more generally, it is desirable to exert some degree of control on both the vertical *and* horizontal resolution profiles with height, for which we need another control degree of freedom.

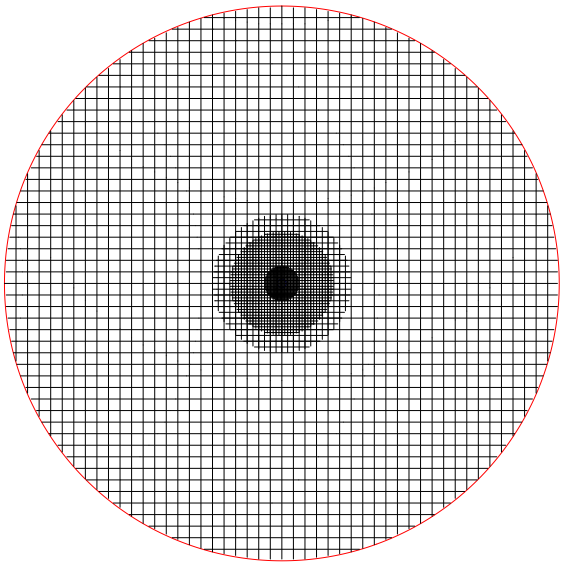

Figure 6. Nested Cartesian grids, of unequal spans (widths in their own grid-spacing units)

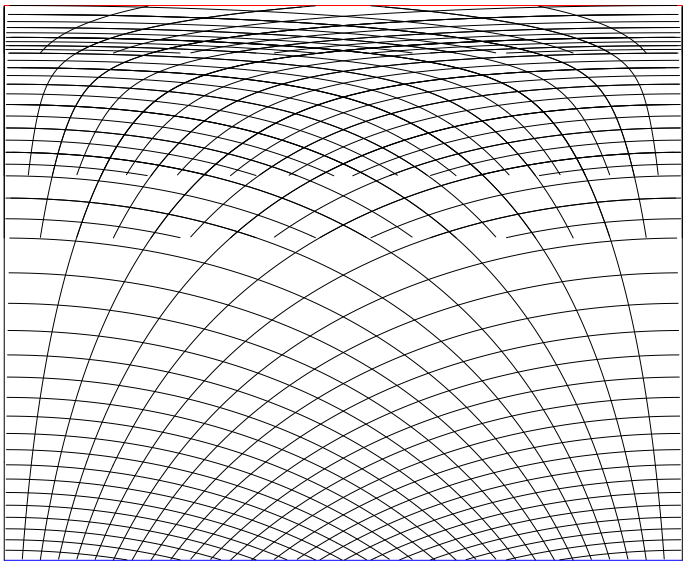

Figure 7. A non-logarithmic radial transformation of the nested Cartesian grids of Fig. 6, showing a vertically nonuniform resolution in both the vertical and horizontal directions.

An extra control, though crude, is available in the choice of the *size* of each nest in the

original telescoping grid configuration. Each nested Cartesian grid can be given its own half-span of grid intervals, a number independent of the size of the next smaller or next larger grid except for the lenient constraint that the finer grid does not extend beyond the next coarser one in the sequence. Fig. 6 shows such an unequal progression of the Cartesian grids and Fig. 7 shows the image of this new nesting corresponding to the same domain as shown in Fig. 5. But we need to take a systematic approach before we can reliably choose the half-spans and design an appropriate radial transformation to achieve horizontal and vertical resolutions that meet the requirements without exceeding them to an unnecessary degree.

The following section provides the systematic framework for the design of the nesting structure and radial transformation. Section 3 refines the design of the cubic grid nesting to ensure an economic gridding is constructed when both the horizontal and vertical scales of the covariances become nontrivial functions of altitude. Section 4 provides a discussion of alternatives to the cubic gridding that might be advantageous at very high resolutions. Preliminary tests of the proposed nested cubic gridding method are described, and the results presented, in Section 5 and, finally, a discussion section summarizes the ideas and experiments presented in this note.

3. Designing the nested cubic grid structure and radial transformation

The aim of this section is to set out a systematic procedure by which the desired horizontal and vertical resolution at each altitude is assured by an economical choice of the combination of the smooth radial transformation and the numbers of half-spans of each of the grid nests. Since the grid refinement occurs in discrete jumps by a factor of two each time a radial trajectory enters a new nest, the desired resolution must be exceeded by this factor in places, so the control we seek is of a rather crude kind.

For reasons articulated in the previous section we only consider the 'inverting' class of radial transformations (resembling Fig. 3 rather than Fig. 2) and the generations, G, of the successive grid nests are assumed to commence with the coarsest, $G = 1$, with each refinement incrementing G by one. The half-span of each grid is denoted N_G and the spacing is scaled arbitrarily but conveniently to:

$$\Delta R_G = 2^{-G}, \tag{3.1}$$

so that the half-width of this grid is,

$$R_G = N_G \Delta R_G = N_G 2^{-G}. \tag{3.2}$$

The angular resolution, $\Delta\phi$, at radius R from the center of Cartesian grid G is:

$$\Delta\phi = \frac{\Delta R_G}{R}, \tag{3.3}$$

which is unchanged on transforming to the Earth's atmosphere where, at height z, we assume the required horizontal resolution (in Earth radians) to be $s_\phi(z)$. Hence, we require that $\Delta\phi \geq s_\phi$ for the finest grid extending to the horizontal (spherical) surface at altitude z when mapped by the radial transformation. In practice, this means that the critical resolution for grid G occurs (except for the very last G) right at the smallest radial distance, R_{G+1}, of the *next* grid in the nesting (at which point the resolution jumps). In Fig. 8 the graphs of $\log_2(\Delta\phi)$ for each of

the grids of Fig. 6 is plotted against $H = -\log_2(R)$. If the corresponding logarithm of s_ϕ, the intended angular resolution, mapped to the H variable follows the dashed graph, then we see, first that the grid configuration is not adequately achieving the target resolution at between about $H = -3$ and $H_{(2)} \approx -3.3$, where only grid $G = 1$ extends. Second, we see that, at other places, the given gridding is quite inefficient, being significantly finer than the criterion requires.

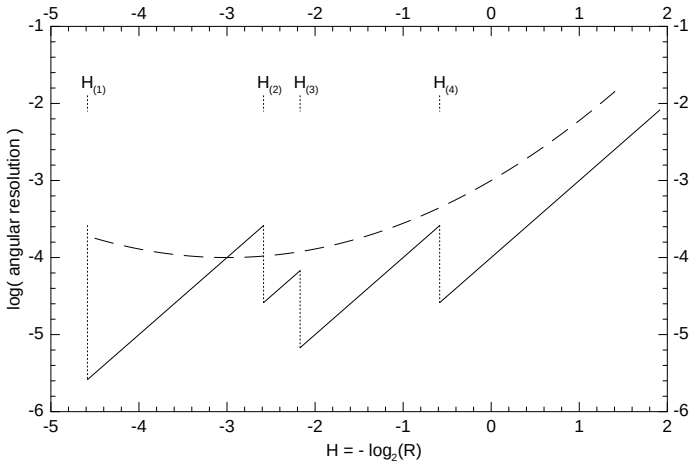

Figure 8. Graphs of the $\log_2$ of the angular resolutions attained by the nested grids at different radial distances, also plotted logarithmically as $H = -\log_2(R)$, in the case of the nested configuration shown in Fig. 6. A graph (dashed) of the $\log_2$ of a hypothetical profile of intended resolution is also plotted. The intended resolution is not met by the configuration just outside the extent of the second nest, but is over-generously exceeded in other places.

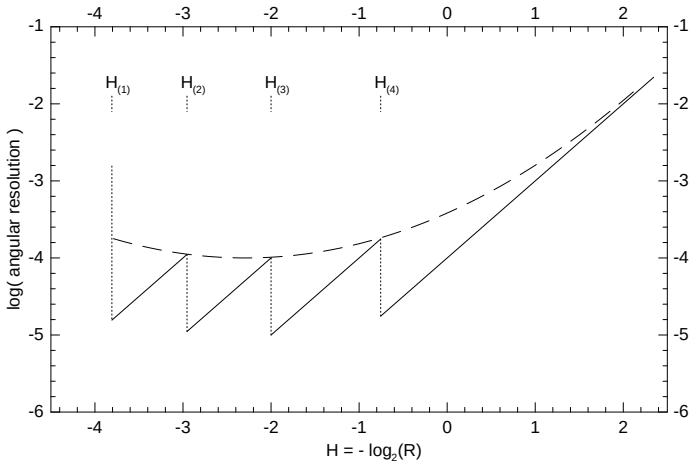

Figure 9. Graphs of the $\log_2$ of the angular resolutions attained by the nested grids (forming the saw-tooth series) that collectively more efficiently achieve the intended resolution, given by the dashed curve, equivalent to the curve shown in Fig. 8.

To maximize economy when the intended resolution maps to a smooth curve in H, it is

usually appropriate to choose the equality of $\Delta\phi$ and s_ϕ (at least approximately) at the critical altitudes corresponding to the tops of the saw-tooth graph shown in Fig. 8. That is:

$$s_\phi(z_{G+1}) = 2/N_{G+1}, \tag{3.4}$$

(so long as this N_{G+1} does not exceed $2N_G$ or fall short of $N_{G+2}/2$, of course). Since the slanting portions of the saw-tooth graph are segments of a family of fixed and equally-spaced parallel lines, this task would be very easy if the graph of $\log_2(s_\phi)$ remained fixed and was known. But we are initially given s_ϕ as a function of z, not H, and the functional relationship connecting z and h is yet to be determined. The solution to this obstacle is simply to require that the desired resolution in both the horizontal ('angular') direction, which is s_ϕ, and in the vertical direction, which we likewise call s_z, maps isotropically to the Cartesian domain of nested cubes. This criterion, is equivalent to

$$|\frac{dz}{dH}| = \frac{s_z}{s_\phi}\log_e(2), \tag{3.5}$$

which can be integrated directly:

$$R \propto \exp\left(\pm\int \frac{s_\phi\, dz}{s_z}\right). \tag{3.6}$$

If we assume this radial transformation has already been done for the curve of intended resolution shown in Fig. 8, then a change in the configuration of nested cubes (Fig.9) attains this resolution with an economy of total grid points fewer by a factor of about ten. The constant of proportionality in (3.6) is first chosen to allow the finest ($G = 4$) grid's resolution graph to just attain tangency with the graph of intended resolution at the largest H corresponding to the highest altitudes. This grid is extended until it over-resolves by at least a factor of two, whereupon the next coarsest grid assumes responsibility, and so on, down to the Earth's surface.

4. Alternative gridding

Given approximately isotropic structures of a given scale, a triangular grid in two dimensions is slightly more efficient at resolving them than a square grid. Correspondingly, in three dimensions, a close-packing grid, such as that known to crystallographers as the 'face-centered cubic' (fcc) lattice is more efficient than the simple cubic lattice (for example, see Ziman 1979). The set of resolved wave-vectors of a given lattice can be taken to be those that lie within the Voronoi cell of the reciprocal lattice that contains the origin, a region of wave-vector space known in solid-state physics as the first 'Brillouin zone'. (For any scatter of points in space, each Voronoi cell is the region of space closer to a given point of the scatter than to any other.) In the case of the fcc lattice, the reciprocal lattice is of the 'body-centered cubic' (bcc) kind and the Brillouin zone is consequently in the form of a truncated octahedron made up of eight regular hexagon and six squares (see illustration in Fig. 10).

The reciprocal of a cubic lattice is another cubic lattice with a Brillouin zone in the form of a cube; the fcc lattice, formed by interleaving four cubic lattices so that each cube's face has another lattice's point at its center, leads to the cuboctahedral Brillouin zone having the corners of the former (cube) Brillouin zone at the centers of the eight hexagonal faces. For an

expenditure of a factor of four invested in increasing the number of lattice points, the effective resolution, measured by the radius of the largest inscribed sphere that the Brillouin zone can support, increases in the linear sense by a factor of $\sqrt{3}$ or, in volume sense, by a factor of $3^{3/2} \approx 5.2$ implying a theoretical efficiency advantage over the cubic lattice of about 20%.

It is clearly awkward to have, as the numerical grid on which filtering is carried out, the four interleaved lattices of the fcc arrangement in comparison with the single cubic lattice. (There is an oblique basis of generators from which all the points of the fcc lattice have integral coordinates, but then an inconveniently large array of values would be needed to contain the cubic domain of interest, defeating the purpose of the more exotic lattice.) But a more serious difficulty is extracting by interpolation the filtered information from such an fcc lattice without seriously degrading (by inadvertent numerical damping or aliasing) the higher wavenumber components that this lattice is supposed to resolve. On balance, it looks as if the potential gains of such an exotic computational lattice do not presently justify the associated costs and inconvenience, although the computational economy may tilt the argument in favor of the exotic computational lattice at very high resolution.

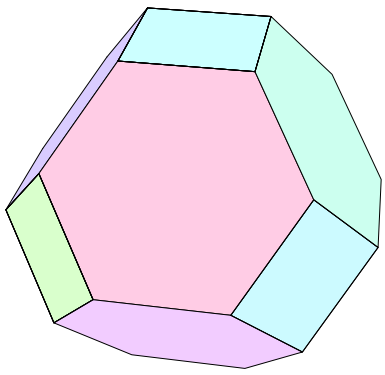

Figure 10. A view of a truncated octahedron.

5. Test implementation of the cubic grids onto the GSI

(*a*) *Design of the nested cubic grids*

As a first trial, a set of nested cubic grids was designed to be able to cover the stratosphere and mesosphere, where the horizontal correlation scale lengths of the stream function (SF) and unbalanced velocity potential (VP) are very long. The following very rough assumptions were used for this definition:

(i) The background error statistics for the T62-GSI, which were estimated by the so-called "NMC method" (Parrish and Derber, 1992), were used to provide baseline parameters in this test.

(ii) "Very long correlation length" was defined as 1,000km.

(iii) The minimum requirement for the angular resolution in the cubic grid space (horizontal resolution in the original lat/lon grid space) was defined as half of the minimum horizontal correlation scale length in the error statistics at the same level.

(iv) The minimum requirement for the radial resolution in the cubic grid space (vertical resolution in the original lat/lon grid space) is defined as the same length as the minimum vertical correlation length in the error statistics at the same level.

The reason for the assumptions (iii) and (iv) is that the generated covariance is a smoothing field and the features small compared to these correlation scale lengths tend to be smoothed out. Since the horizontal derivatives, which emphasize smaller scales of the SF and VP, are used to derive the physical values (i.e., wind components), the horizontal resolution criterion is twice as demanding as the vertical resolution criterion. However, it is not yet clear to what extent the resolution may be safely reduced, the stated definitions being somewhat arbitrary, so these requirements are not always enforced strictly.

Fig. 11 shows the maximum correlation scale length in the background error statistics for the SF, VP, unbalanced temperature and specific humidity, all plotted against sigma levels. The very long correlation length (1,000km) is found at $-3.3 < \log_2(\sigma) < -2.3$ and $\log_2(\sigma) < -8.0$ for the VP and at $\log_2(\sigma) < -4.0$ for the SF. On the other hand, the scale lengths for the temperature and humidity are much shorter than the ones for the SF and VP. Thus the nested cubic grids are designed only for the SF and VP at $\log_2(\sigma) < -2.0$. Since the SF and VP are strongly coupled, the same cubic grid configurations are used for both variables.

Fig. 12 shows the minimum horizontal correlation lengths for the SF and VP. The scale lengths at $\log_2(\sigma) = -2.0$ are around 500km and increase to 700km by $\log_2(\sigma) = -3.0$. The scales show only smaller changes around $-9.0 < \log_2(\sigma) < -4.0$, and much longer scale lengths at the top two model levels.

Fig. 13 shows the minimum vertical correlation lengths. The scale length for the VP shows local minima around $-4.0 < \log_2(\sigma) < -3.0$. Except for these levels, the length is almost constant at $-5.0 < \log_2(\sigma) < -2.0$ and gradually decreases at $\log_2(\sigma) < -5.0$. The length for the SF is longer than the one for the VP.

Based on this information, the following set of vertical level conversion equations was defined for three nested cubic grids with an edge length of 68.

$$\begin{cases} R = 40\sigma^{0.35} & : \text{Cube1} & (-5.28\ (26\text{mb}) < \log_2(\sigma) < -1.15\ (456\text{mb}), & R = 8\text{—}34) \\ R = 80\sigma^{0.35} & : \text{Cube2} & (\log_2(\sigma) < -3.63\ (81.8\text{mb}), & R = 8\text{—}34) \\ R = 160\sigma^{0.35} & : \text{Cube3} & (\log_2(\sigma) < -6.53\ (11.0\text{mb}), & R = 12\text{—}34) \end{cases} \tag{5.1}$$

where R is radius in each cube (i.e. distance from the center of the cube) and σ is the sigma level in the lat/lon grid space.

Fig. 14 shows a schematic view of these conversion equations. At the model top, Cube3 is mainly used. The horizontal resolution of Cube3 at the model top is about 500km, which is about half of the horizontal correlation length of the VP ($\sim$1,000km, Fig. 12). The horizontal resolution is increased rapidly for $\log_2(\sigma) \leq -6.53$. Since $\log_2(\sigma) > -6.53$ is out of the domain of Cube3, it is the Cube2 resolution ($\sim$350km) that is applied; the same transition was repeated around $\log_2(\sigma) > -3.63$ from Cube2 to Cube1.

Fig. 15 shows radial resolution changes in the nested cubes against sigma levels. The Cube3 resolution is applied at $\log_2(\sigma) \leq -6.53$, Cube2 at $-6.53 < \log_2(\sigma) \leq -3.63$, and Cube1 at $-3.63 < \log_2(\sigma) \leq -1.15$. Though the vertical resolution is insufficient between $-3.5 < \log_2(\sigma) \leq -3.0$ (1.3 in the Cube1 but 1.0 in the lat/lon grid, Fig. 13), this deficit is ignored

in this test. For a smooth transition of the covariances at each cube boundary, the weighting functions shown in Fig. 16 are applied for each cubic grid space and in the original three patch space to synthesize the covariances which are estimated in each grid independently. Therefore, the effective resolution near the boundary is much lower than the one mentioned above, because it is composited with the lower resolution cubes.

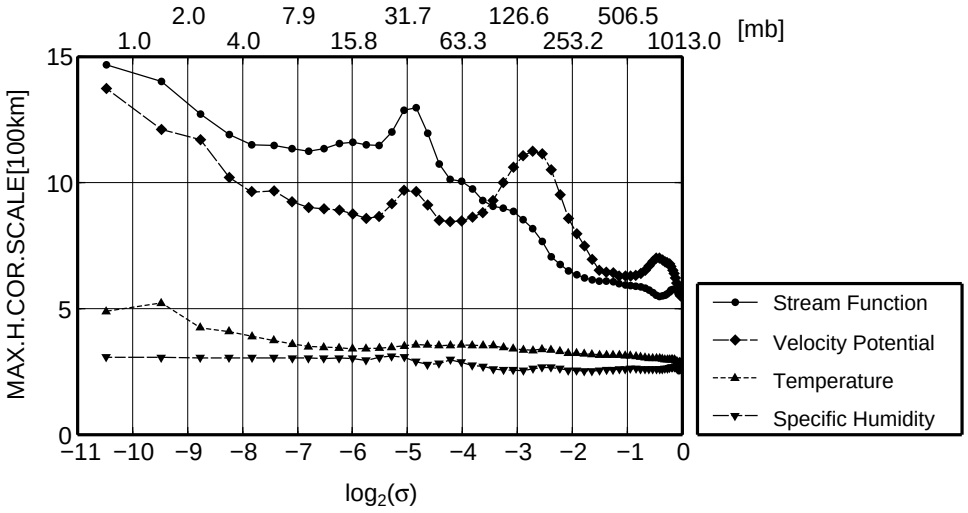

Figure 11. Maximum horizontal correlation scale length in the T62-GSI background error statistics against sigma levels. Symbols show the model levels. Standard pressure levels corresponding to the grid lines, also shown in the figure, are labeled above the graph.

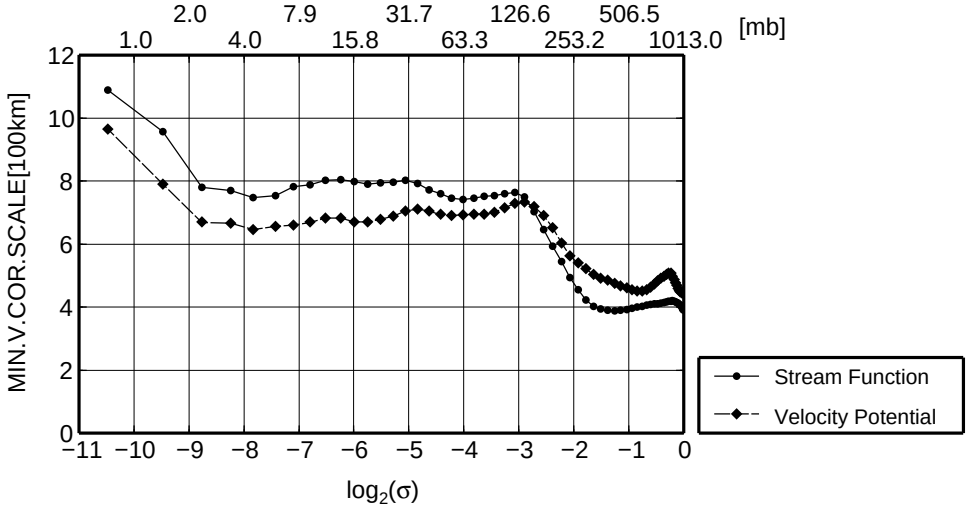

Figure 12. Same as Fig. 11 but for minimum horizontal correlation length.

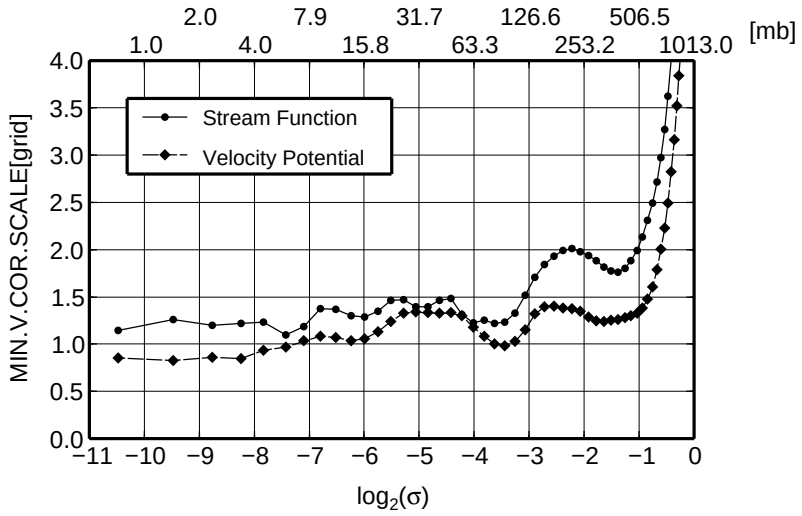

Figure 13. Same as Fig. 11 but for minimum vertical correlation length.

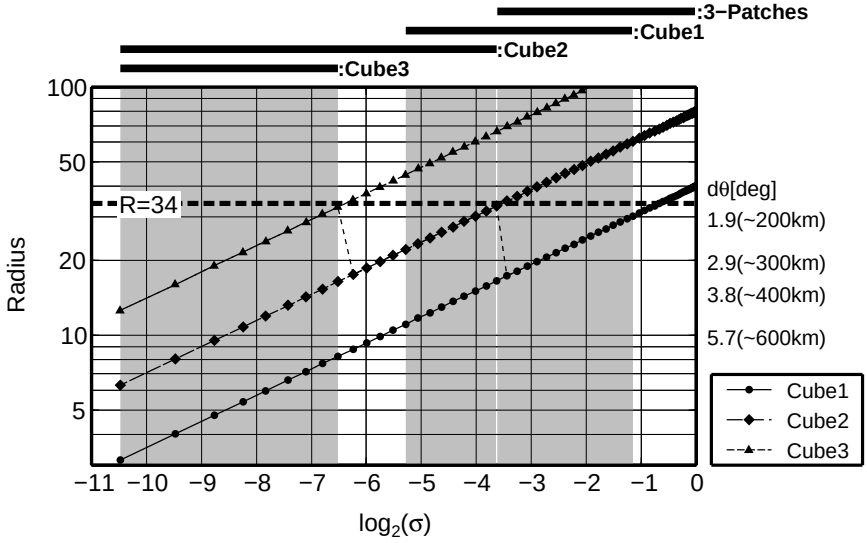

Figure 14. The horizontal axis shows sigma level in the original lat/lon grid space and the vertical axis shows radius in the cube (distance from the cube center). The radius corresponds to the angular resolution $d\theta$ (and horizontal resolution in the lat/lon grid space). The lines of symbols show the model grid level positions mapped to the radial grid index of each of the three cubes at positions along the cubes' principal axes. Thick bars over the plot area show the grid coverage. Gray shades show transition zones from Cube3 to Cube2, from Cube2 to Cube1, and from Cube1 to original three patch system. Dashed thin lines show the transition of the resolution at the cube boundary.

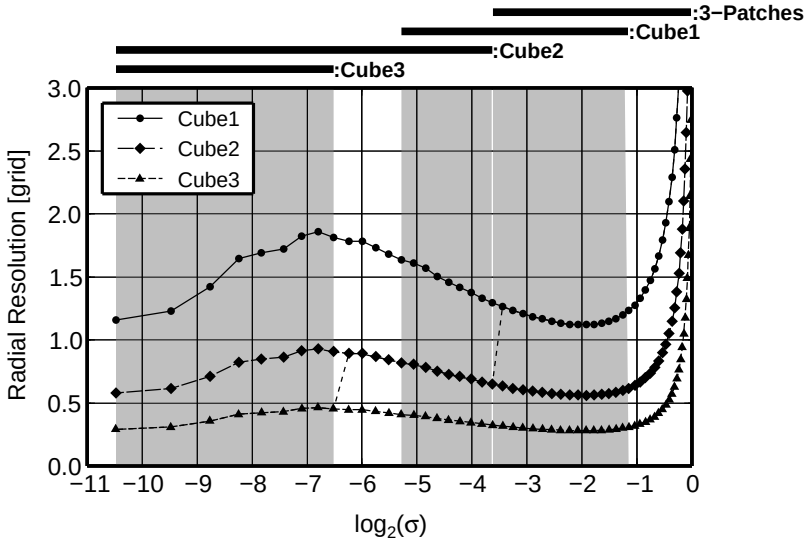

Figure 15. Radial resolutions in each cubic grid. Dashed lines show the transition of the resolution at the cube boundary.

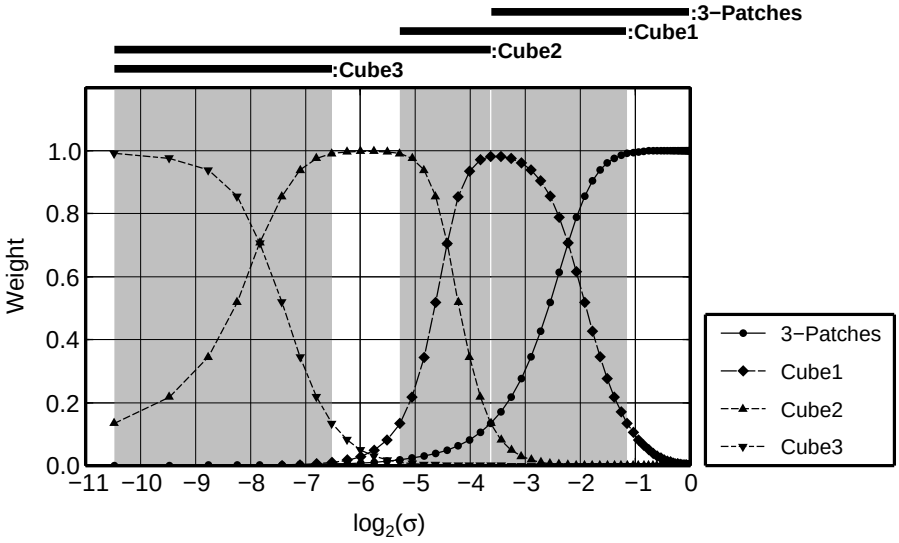

Figure 16. Weighting functions for the each cubic grid space and original three patch system. The square-sum of the weighting function is one.

(*b*) *Map projection validation*

The interpolation between the lat/lon grid and the nested cubic grids are performed by using so-called "diamond interpolation" method (Purser, 2007). A temperature field projection test was performed for validation of this interpolation before constructing the covariances. In this test, a temperature field in the lat/lon grid space is projected into the nested cubic grids

and the projected temperature fields in the cubic grids are projected back into the lat/lon grid space.

In Fig. 17 the upper panels show the temperature field cross section at 180 °E above $\log_2(\sigma) = -1.0$(a) before, and (b) after projection, while panel (c) shows the difference. The lower panels show the projected temperature in the cubic grids. The left hand sides of these panels corresponds to 180 °E. The map projection works fine and the interpolation error was not large except at the top level ($\log_2(\sigma) = -10.48$ (0.7mb)). The relatively larger interpolation error was found around $\log_2(\sigma) = -4.0$ (63mb), $\log_2(\sigma) = -6.5$ (11mb) and at higher levels. Both these levels are close to cubic grid boundaries, which explains why the error was relatively large. Fig. 18 shows the horizontal cross section of the interpolation error.

The contour lines show the temperature field before projection. At $\log_2(\sigma) = -4.02$ (62.4mb) and -6.24 (13.4mb), the interpolation error was concentrated over the Tropics where many small scale features are prominent. At the second model top level, panel (c) shows the larger projection errors, and the original temperature field shows the small scale waves. At the model top level, panel (d) shows much larger projection errors. At the top level, the map projection is performed not by interpolation, but by extrapolation, and the vertical temperature gradient is large between the second top level and the top level. These two factors explain the large error. Evidently, the interpolation at the top (and bottom) of the model domain needs to be improved in the future.

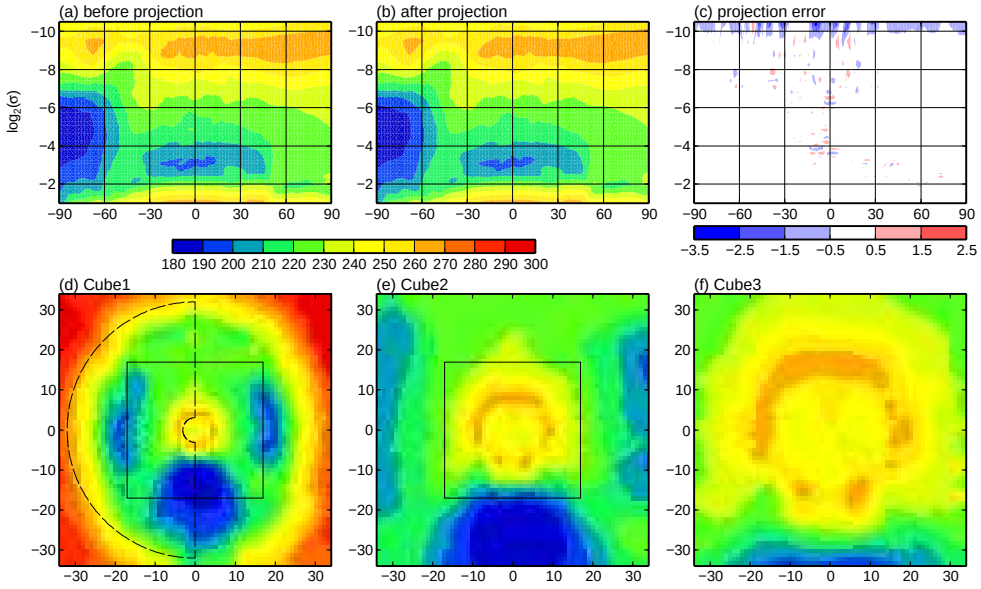

Figure 17. (a) Vertical cross section of the temperature field before projection at 180 °E, above $\log_2(\sigma) = -1.0$, (b) Same as (a) but after projection, (c) The difference. (d, e, f) Temperature field cross sections in the nested cubic grids ((d) Cube1 / (e) Cube2 / (f) Cube3). The square in (d) shows Cube2 domain and the one in (e) shows Cube3 domain. The dashed line in (d) shows the domain plotted in (a,b,c). The left hand side of the panel (d, e, f) corresponds to 180 °E and the other side corresponds to 0 °E.

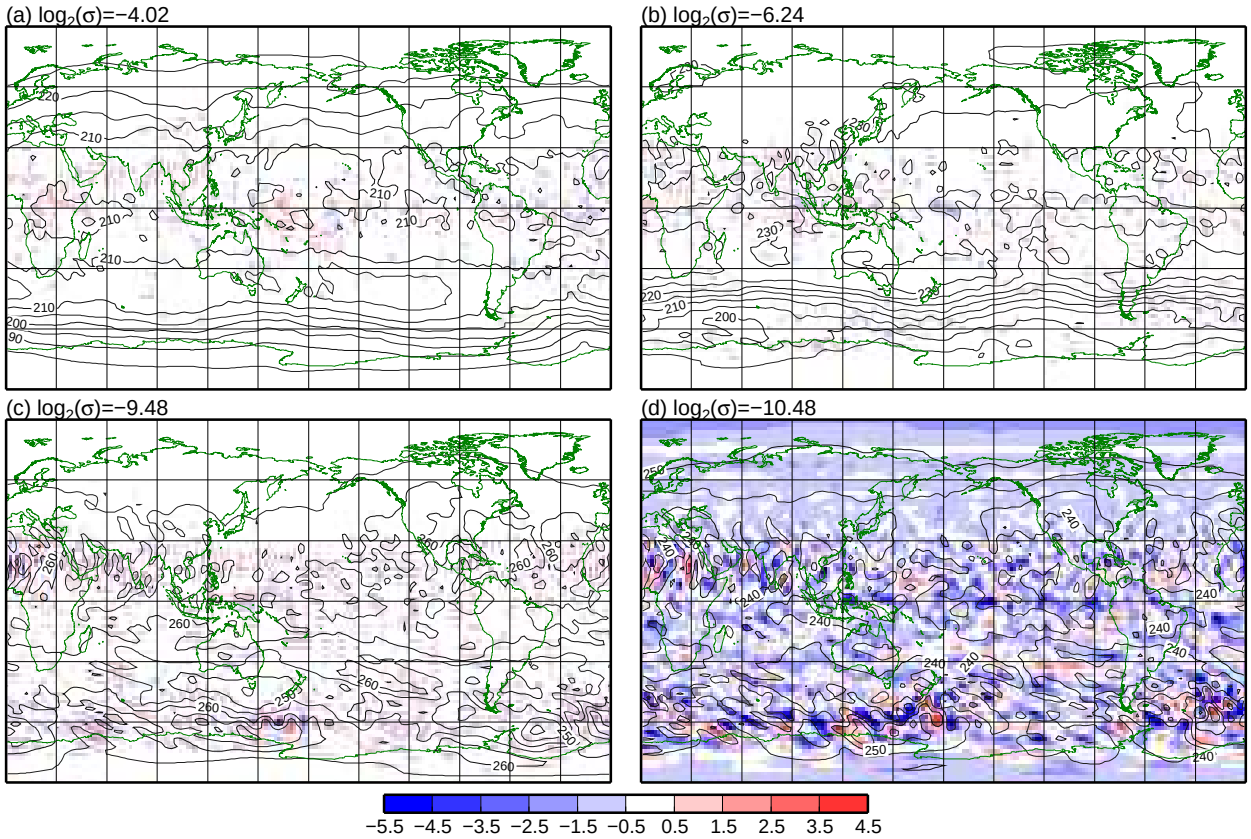

Figure 18. Temperature difference before and after projection at (a) $\log_2(\sigma) = -4.02$ (62.4mb, Lv45(the 45th model level)), (b) $\log_2(\sigma) = -6.24$ (13.4mb, Lv55), (c) $\log_2(\sigma) = -9.48$ (1.4mb, second model top level (Lv63)) and (d) $\log_2(\sigma) = -10.48$ (0.7mb, model top level(Lv64)). Contour line shows the temperature fields before projection, and the interval is 5K.

(*c*) *Aspect tensor definition*

The covariance used in the present standard GSI has horizontally isotropic shape, and its correlation scale length depends on the variable, latitude and height. Constant sigma levels in the lat/lon grid space are projected to constant radius spheres embedded in the cubic grid, and the horizontally isotropic covariance is projected into a covariance that is isotropic in subtended central angle at a constant radius. Since the covariance shape in the cubic grid is not three-dimensionally isotropic, the anisotropic recursive filter is needed for the covariance construction.

The anisotropic recursive filter involves an extension of the basic recursive filter, that enable it to stretch or compress the covariance in arbitrarily defined directions by selecting sets of possibly non-Cartesian grid lines dictated by the "hexad method" (Purser et al. 2003) along which numerical smoothing is then applied. The directions of local dilatation or compression and the associated correlation scale lengths are defined by the symmetric second-rank "aspect tensor" (approximately a normalized tensor of the local second-moments).

The inverse of the aspect tensor for an isotropic covariance is defined as follows:

$$\mathbf{A}_{\rm iso}{}^{-1} = \begin{pmatrix} \frac{1}{L_{\rm iso}^2} & 0 & 0 \\ 0 & \frac{1}{L_{\rm iso}^2} & 0 \\ 0 & 0 & \frac{1}{L_{\rm iso}^2} \end{pmatrix}, \tag{5.2}$$

where $\mathbf{A}_{\rm iso}$ is the aspect tensor and $L_{\rm iso}$ is the correlation scale length.

By stretching or compressing the isotropic covariance along the radial direction, the covariance will follow the constant radius sphere. The inverse of the aspect tensor for such a covariance at the position (i, j, k) is defined as follows:

$$\mathbf{A}^{-1} = \mathbf{A}_{\rm iso}^{-1} + \left(\frac{1}{L_{\rm rd}^2} - \frac{1}{L_{\rm iso}^2} \right) \mathbf{P}\mathbf{P}^{\mathsf{T}}, \tag{5.3}$$

$$\mathbf{P} = \begin{pmatrix} \left(\sqrt{(i+1)^2 + j^2 + k^2} - \sqrt{(i-1)^2 + j^2 + k^2} \right) /2 \\ \left(\sqrt{i^2 + (j+1)^2 + k^2} - \sqrt{i^2 + (j-1)^2 + k^2} \right) /2 \\ \left(\sqrt{i^2 + j^2 + (k+1)^2} - \sqrt{i^2 + j^2 + (k-1)^2} \right) /2 \end{pmatrix}, \tag{5.4}$$

where $\mathbf{A}$ is the aspect tensor, $L_{\rm rd}$ is the correlation scale length for the radial direction (vertical correlation scale length in the original lat/lon grid space), $\mathbf{P}$ is a radial vector and (i, j, k) shows the position relative to the center of the sphere. The $L_{\rm iso}$ in this equation works as a correlation scale length for the directions following the constant radius sphere, and it corresponds to the horizontal correlation scale length in the original lat/lon grid space. Note that all the variables in Eq.(5.3)–(5.4) are functions of the position (i, j, k). When $L_{\rm iso}$ is smaller than $L_{\rm rd}$, the covariance is stretched along the radial direction and when $L_{\rm iso}$ is larger than $L_{\rm rd}$, the covariance is compressed along the radial direction.

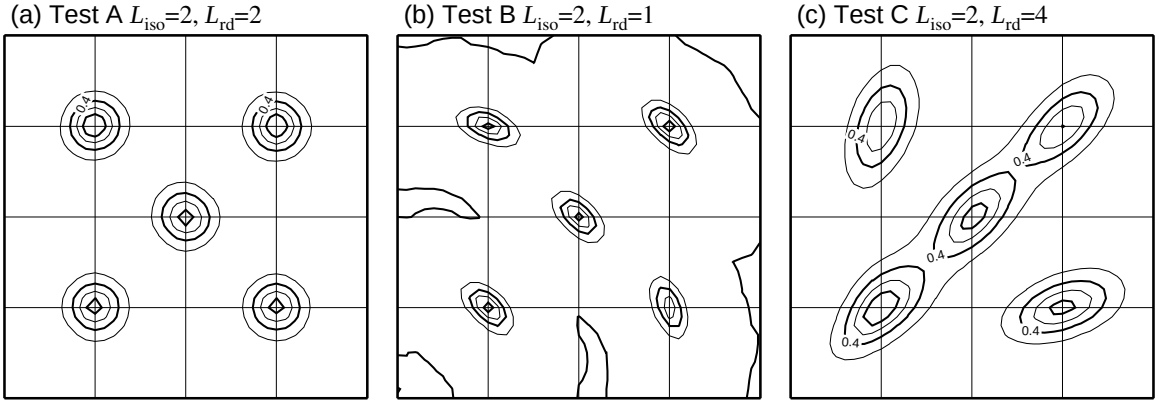

Figure 19. Idealized 2D covariance test results. Dimensions of the domains are 41 by 41, and the bottom left corner is (1,1). The grid lines are drawn with 10 grid interval. Contour intervals are 0.2.

To confirm this formulation, three idealized tests are performed. Test domain is 2D grid space (41x41) and the circle (that is a projection of the constant radius 3D sphere to the 2D

field) center is defined at the bottom left corner (1,1). Five signals are set at the positions of (11,11), (31,11), (11,31), (31,31) and (21,21). Test A defines the parameter $L_{\rm iso} = L_{\rm rd} = 2$, Test B uses $L_{\rm iso} = 2$, $L_{\rm rd} = 1$, and Test C uses $L_{\rm iso} = 2$, $L_{\rm rd} = 4$. The anisotropic recursive filter is applied with these parameters. Fig. 19 shows the filtering results. Test A shows isotropic covariance patterns at all five points, but the amplitudes and the patterns are not identical because of imperfections in the present method of normalization used for the anisotropic filter. This problem is expected to be solved in the near future when normalization based upon the asymptotic "Parametrix Expansion" method will be implemented (Purser 2008a,b). As expected, Test B shows that the covariance patterns are compressed along the directions toward the left bottom corner by the factor of half, and Test C shows that the covariance patterns are stretched along the directions by the factor of two. These results show that the covariance shapes in the cubic grid can be controlled arbitrarily by this formulation.

In the GSI, the $L_{\rm iso}$ and $L_{\rm rd}$ are defined by the input background error statistics with the following unit conversion.

$$L_{\rm iso} = C_1 \, L_{\rm h} \, R/R_{\rm earth}, \tag{5.5}$$

$$L_{\rm rd} = C_1 \, L_{\rm v} \, dR/d{\rm lev}, \tag{5.6}$$

where $L_{\rm h}$ is a horizontal correlation scale length in the background error statistics, R is the radius in the cubic grids at the position (i, j, k), $R_{\rm earth}$ is the radius of the Earth, $L_{\rm v}$ is a vertical correlation scale length in the error statistics, dR/dlev is a derivative of the R against original model levels at the position (i, j, k). C_1 is a tuning parameter and the value is set to the $\sqrt{2}$ in this test. Both $L_{\rm h}$ and $L_{\rm v}$ are converted to the $L_{\rm iso}$ and $L_{\rm rd}$ after the map projection from the lat/lon grid space to the cubic grids, and the anisotropic recursive filter is initialized with the aspect tensor.

(*d*) *Covariance construction test in the GSI with the cubic grids*

The covariance construction steps in the GSI with the cubic grids are:
(i) set signals in the lat/lon grid space;
(ii) execute the adjoint of the "grid-merge-projection" process;
(iii) apply the anisotropic recursive filter and its adjoint in the cubic grids;
(iv) execute the "grid-merge-projection" process.

The grid-merge-projection process is a map projection from the cubic grids to the lat/lon grid, followed by simple weighting average with the grid data in the original three patch system.

In this test, the covariance shapes for point-I (0 °E, 60 °S) and point-II (180 °E, 30 °N) at $\log_2(\sigma) = -6.24$ (13.4mb) and $\log_2(\sigma) = -2.23$ (216mb) are compared. Fig 20 shows the covariance construction test result at $\log_2(\sigma) = -6.24$ by the orthogonal projection with the projection center at the signal points. The upper panels show the response to the point-I and the lower panels show the one to the point-II. In this test, the variances are not used. The filtering result is normalized to one. The left column shows the standard GSI (isotropic mode) result, the center column shows the result in anisotropic mode, but without actual anisotropy, and the right panels show the result of the anisotropic mode with the cubic grids. The anisotropic mode uses the three-patch system as standard GSI, but the anisotropic recursive filter is applied instead of the isotropic recursive filter.

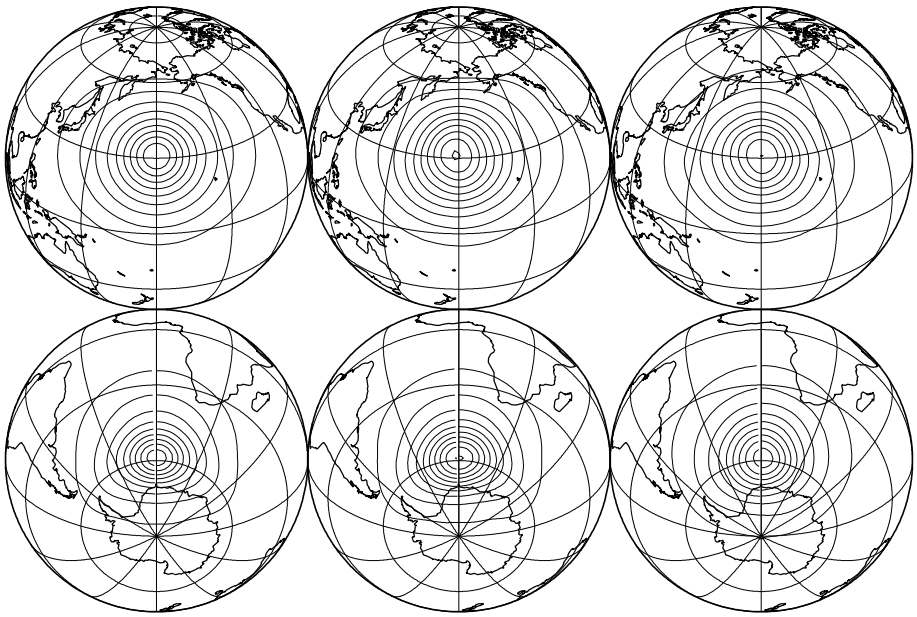

Figure 20. Covariance construction test result at $\log_2(\sigma) = -6.24$ (13.4mb) with the contour interval of 0.1. The initial signals are put at (180 °E, 30 °N) (upper panels) and at (0 °E, 60 °S) (lower panels). The left column shows the standard GSI (isotropic mode) result, the center column shows the anisotropic mode GSI result, and the right column shows the anisotropic mode GSI with cubic grid results.

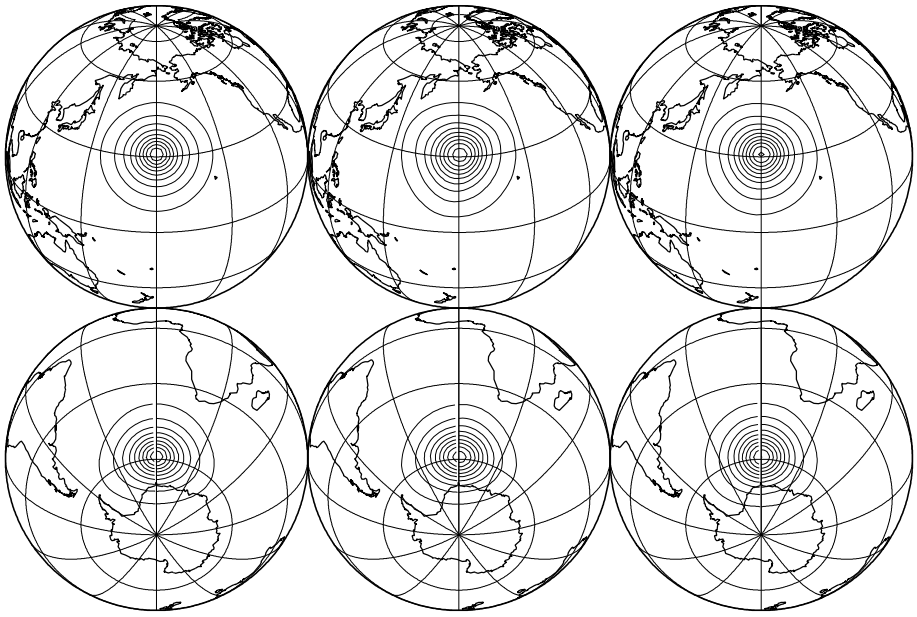

Figure 21. Same as Fig. 20 but at $\log_2(\sigma) = -2.23$ (216mb).

The covariance shapes at point-I are not so different among the three tests. On the other hand, the covariance shapes at point-II show the larger differences. While the covariance

constructed in the GSI with the cubic grids shows pretty isotropic shape, the covariances in the other tests are distorted near the 60 °S parallel where they suffer from a spurious triangular shape. Since the horizontal correlation scale length in the GSI depends on the latitude, the estimated covariance shape will not have an exactly isotropic shape. However, the distortion is quite large and is located at the boundary of the three-patch system (60 °S/60 °N). Thus this distortion clearly comes artificially from the covariance composition and the cubic grid application can solve this issue. Fig. 21 shows the same plots as Fig. 20 but at $\log_2(\sigma) = -2.23$ (216mb). Since the horizontal correlation scale length are much shorter than the one at $\log_2(\sigma) = -6.24$ (13.4mb), the differences among the three tests are much smaller.

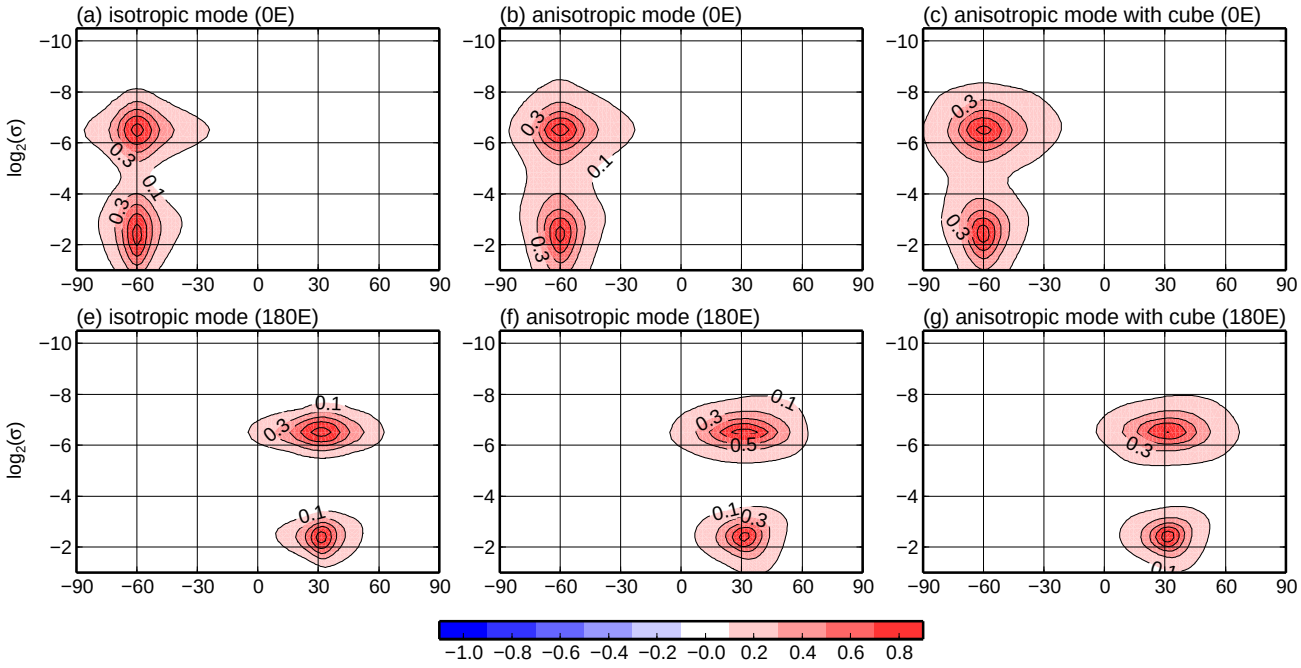

Figure 22. The vertical cross section of the Covariance construction test result. The upper panels show the cross section at 0 °E (for Point-II) and the lower panels show the ones at 180 °E (for Point-I).

Fig. 22 shows the vertical cross section of this test result. Compared to the horizontal cross section, the difference among the tests is much larger. One of the causes is the decoupling of the horizontal/vertical covariance construction in the standard GSI. Although the standard GSI decouples the horizontal covariance construction and the vertical covariance construction, the anisotropic mode does not decouple them. With this decoupling, the standard GSI can use a different configuration of the filtering for the horizontal and vertical directions (a composition of empirical functions being used in the vertical). On the other hand, the anisotropic mode cannot use such a special treatment and can use only one configuration for both directions. In this test, the numbers of the Gaussians in the anisotropic mode and the cubic grids are set to three and this configuration is the same as the one for horizontal filtering in the standard GSI, so the horizontal cross section shows much better agreements. But the basic pattern, which means the larger vertical extent of the covariance at 0 °E and the thin structure at 180 °E,

looks reasonable.

As a result of this test, while there are some characteristic differences from the covariance constructed in the standard GSI, the covariance constructed in the cubic grid using the GSI statistics and projected into the lat/lon grid space is quite reasonable and it does not suffer from the horizontal patch boundary problems.

(*e*) *Data assimilation test*

Some data assimilation tests were performed and the innovation fields estimated by the isotropic mode GSI with the cubic grids (isocb-test) are compared to the ones estimated by the standard GSI (iso-test).

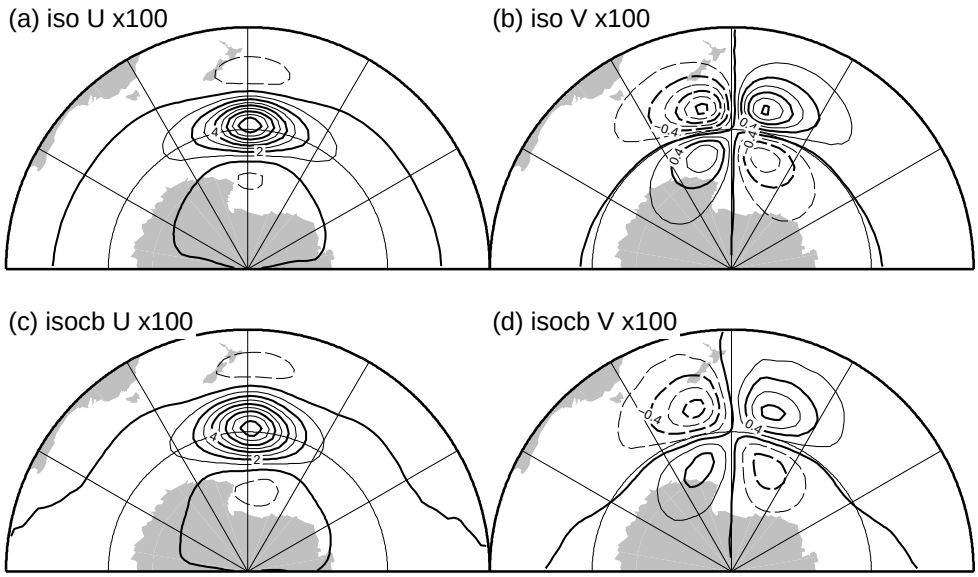

Figure 23. Innovation fields in the single observation data assimilation test at $\log_2(\sigma) = -6.24$ (13.4mb). The upper panels show the result in the standard GSI, and the lower panels show the one in the isotropic GSI with cubic grids. The left panels show the U-wind innovation fields with the contour interval of 0.01 m/s and the right panels show the V-wind innovation with the contour interval of 0.002 m/s.

Fig. 23 shows innovation fields in the single observation data assimilation test. The type of the assimilated data is U-wind (zonal wind) and the data is put at (180 °E,60 °S, $\log_2(\sigma) = -6.64$ (10mb)). The innovation value and the observation error are set to 0.1 m/s. Looking at the result, the U-wind innovation intensity for isocb-test is comparable to that for iso-test, but the extent is different. While the innovation pattern in the iso-test shows a narrow shape on the constant latitude (60 °S) band, the one in the isocb-test shows a much wider shape. As shown in Fig. 20, the stream function covariance in the iso-test is distorted around 60 °S, and it makes the gradient perpendicular to the patch boundary larger. As a result, the wind innovation field (derivative of the stream function) becomes narrower. On the other hand, the V-wind (meridional wind) innovation field in the isocb-test shows slight bumps in the contour lines. These bumps should be removed by the future employment of the "blended hexads algorithm" described in some detail in Purser (2005).

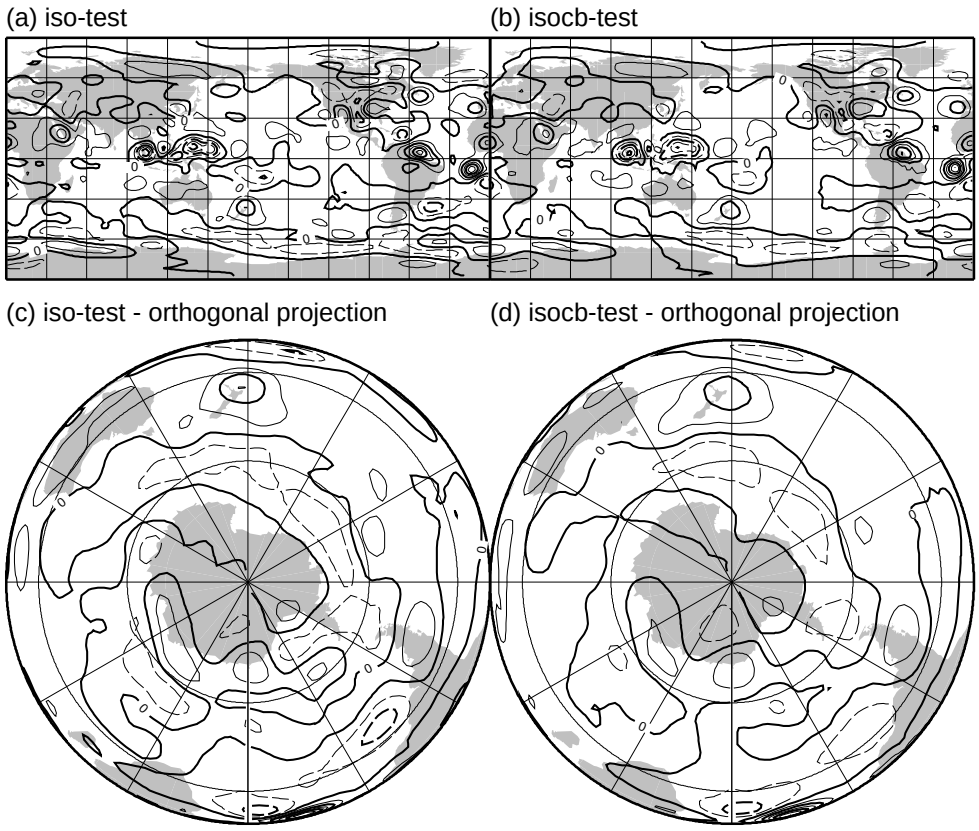

Figure 24. U-wind innovation fields in the full observation data assimilation test at $\log_2(\sigma) = -6.24$ (13.4mb). The upper panels show the global data in the Cartesian mapping and the bottom panels show the same one in the orthogonal projection with the center at the South Pole. Left panels show the result from standard GSI and the others are the ones from the isotropic GSI with the cubic grids. The contour interval is 1m/s and dashed contour lines show the negative values.

Fig. 24 shows the U-wind innovation fields at $\log_2(\sigma) = -6.24$ (13.4mb) in the full observation data assimilation test. Roughly speaking, the innovation fields show good agreement in both tests but, as mentioned in the context of the single observation test of Fig. 23, the innovation pattern around 60 °S is narrower in the iso-test than that in the aniso-test.

While the result of the isocb-test is reasonable over the mid-latitudes and polar regions, it shows noisy patterns over the Tropics at some levels. Fig. 25 shows the case. This problem is found in the range of -4.84 (35.4mb)$< \log_2(\sigma) < -3.07$ (121mb). As mentioned in section 5(a), the resolution deficit is ignored in this test, so the problem might just be a result of the deficient resolution.

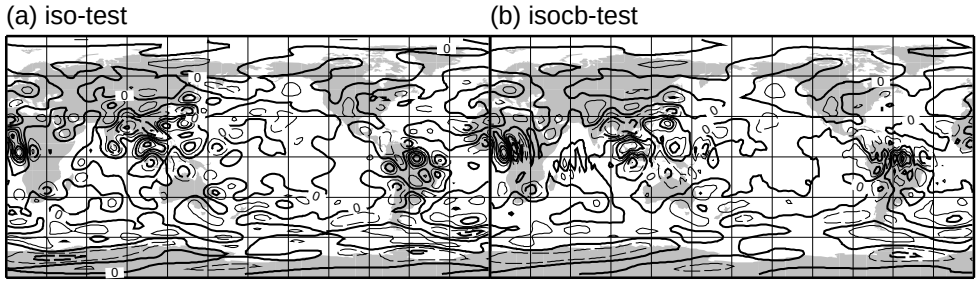

Figure 25. U-wind innovation fields in the full observation data assimilation test at $\log_2(\sigma) = -3.25$ (107mb). Contour interval is 1m/s and dashed contour lines show the negative values.

To check this point, a different configuration of the cubic grids is tested. The following two cubes are used for this sensitivity test:

$$\begin{cases} R = 80\sigma^{0.35} & \text{:Cube1x} \quad (-7.10\ (7.4\text{mb}) < \log_2(\sigma) < -1.92\ (268\text{mb}), & R = 12 - 51) \\ R = 160\sigma^{0.35} & \text{:Cube2x} \quad (\log_2(\sigma) < -4.84\ (35.4\text{mb}), & R = 12 - 51) \end{cases} \tag{5.7}$$

The result is shown in Fig. 26. The noisy pattern could not be found in this result, but the computational cost for the cubic grids is twice that of the setting introduced above. Further sensitivity tests are required to find the optimum settings.

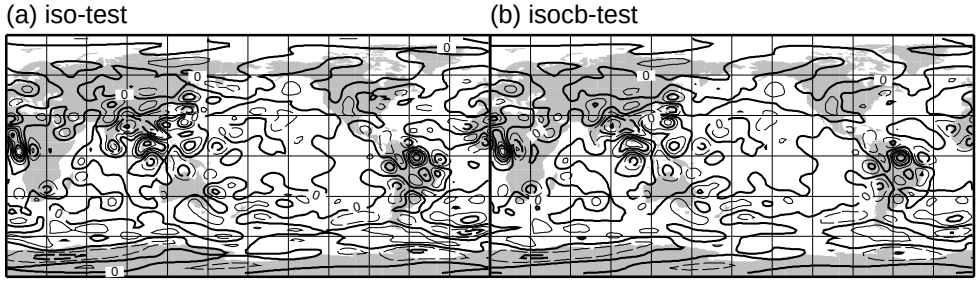

Figure 26. Same as Fig. 25 but uses the cubic grids configuration described in Eq. (5.3) for isocb-test.

(*f*) *Cycled test in the T62-GSI*

Finally, a cycled data assimilation test is performed with a resolution of T62. The cycle period is from 10/21/07 to 11/11/07 and the assimilation is performed every six hours. Fig. 27 shows the zonal-mean (left) U-wind and (right) temperature in the analyzed field, averaged over the whole test period. The contour lines show the iso-test result and the color scales show the difference between the iso-test result and the isocb-test result. As expected, the U-wind is reduced around the South Polar regions, but the difference is small. However, the difference over the Tropics is relatively large, but might come from the noise problem mentioned in the last section. A temperature difference is found above $\log_2(\sigma) = -7.0$ (7.9mb), where the isocb-test shows a higher temperature than the iso-test. Fig. 28 is the same as Fig. 27, but for the RMSE fields at the 120 hour forecast time calculated with each 'error' computed by reference to its

'own' analysis, as the cycling experiment allows. The figure shows that the slight positive impact (blue color, which means smaller RMSE in the isocb-test) dominates in the U-wind forecast, but the impact on the temperature field is pretty small. While the RMSE of temperature shows very small positive impact at -4.0 (63mb)$< \log_2(\sigma) < -7.0$ (7.9mb), it shows negative impact over $\log_2(\sigma) = -8.0$ (4.0mb). As mentioned in section 5(b), the interpolation accuracy is worse at the top level, which might explain the negative impact.

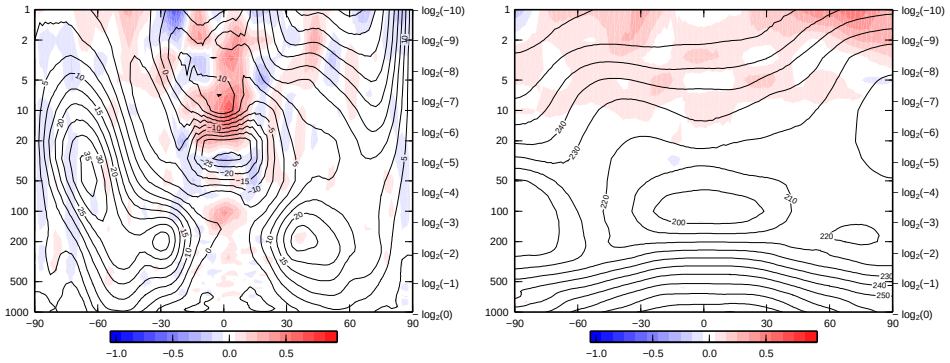

Figure 27. Contour lines show the averaged zonal-mean U-wind (left, the unit is m/s) and temperature (right, the unit is K) in the analyzed fields of the iso-test over the whole test period, and color scales shows the difference from the one in the isocb-test.

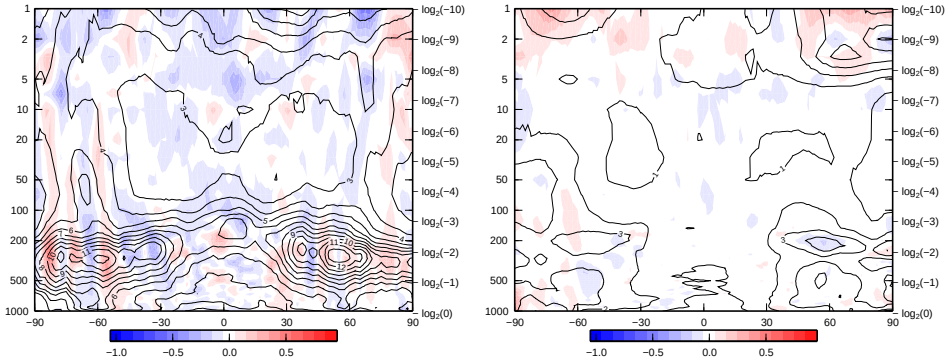

Figure 28. Same as Fig. 27 but for the RMSE field in the 120 hour forecast from the own analyses.

(*g*) *Summary of the test implementation of the cubic grid on the GSI*

In this section, a configuration of the cubic grids was designed and implemented within the GSI. Mostly, it produced reasonable increment patterns although noisy patterns were found at some levels. The noise can be avoided by using higher resolution cubic grids, although that makes the computational cost larger. Finally, a cycled assimilation test was performed, which showed a positive impact on the U-wind field forecast, but negative impact near the model top in the temperature forecast. Overall, the results look very promising, but further investigation and more careful design of the grid configuration is needed. Although it has not yet been mentioned in this work, flow-dependent covariances can be applied easily in this cubic grid

system since it already uses the anisotropic recursive filter regardless of the degree of intended horizontal anisotropy in the spherical coordinate domain. The planetary scale flow-dependent covariances can be tested in this system, though greater care must be taken in the design of the grid configuration to ensure that the grid resolution provided is everywhere adequate.

6. Discussion

The global GSI uses a three-patch system of overlapping grids. If the correlation scales are of modest extent, the synthesized background error covariance shows an adequate degree of continuity. However, when the correlation scale length is very long, it suffers from the horizontal boundary problem, which distorts the innovation patterns around the boundary and makes the horizontal gradients steeper. Since the horizontal gradient of the stream function is used to construct the wind increments, these wind components also suffer badly from this distortion.

On the other hand, the cubic grid system does not have a horizontal boundary and thus does not incur horizontal boundary problems. Although it has vertical boundaries of the nested cubic grids, the impact tends to be smaller than the horizontal boundaries because the construction of the cubic nesting grids and the associated choice of weighting functions ensures that even relatively coarse-scale contributions to the covariance are blended smoothly from one nest to the next at a sufficiently gradual pace as the boundary is approached. Also, owing to the more vertical orientation of the nested grid boundaries, the taking of horizontal derivatives of the control variables has much less of an aggravating effect in this geometry. However, the cubic grid system incurs errors produced by interpolations needed to accommodate its map projection. The noisy patterns shown in Fig. 25 suggest the desirability of a more careful cubic grid design and improved interpolation than we have applied in these preliminary experiments.

Since the current cubic grid configuration is "added on" to the existing standard (isotropic) or anisotropic GSI, it is an encumbrance that can only increase the computational cost. In the cycled data assimilation test, while the average computation time for the iso-test was 9.1 min, the time for the isocb-test was 12.2min (an increase of 34%). A mitigation of the computational cost might result from improvements such as:

(i) skipping the redundant filtering in the three patch system at altitudes sufficiently immersed in the nested cubic configuration;

(ii) employing the new, more economical normalization method for the anisotropic recursive filter.

There is also a need for some sensitivity tests at the operational resolution (T382) GSI since the horizontal resolution of our experiments (T62) is very coarse compared with present operations. Given the additional degrees of control available in designing the nests, the proportionate computational burden of the supplementary cubic grids configuration at a resolution more commensurate with the T382 standard might be found more favorable than the percentage quoted above.

Although this note does not discuss in detail the effects of employing nontrivially anisotropic covariances, such a formulation could apply either an appropriate variant of the Riishøjgaard (1998) type of flow-dependent covariance, or the ensemble-based covariance proposed in Sato et al. (2009). By transforming the flow-dependent aspect tensor to the Cartesian coordinates of the nested cubic grids, it should not be hard to generalize the procedure we have described

here to produce fully adaptive and flow-dependent covariances; future tests should certainly include this more flexible system.

Acknowledgments

The authors would like to thank Drs. Manuel Pondeca, David Parrish, Wan-Shu Wu, John Derber and Henry Juang for the helpful discussions and suggestions. This work was carried out under the collaboration of NCEP and JMA, supported by UCAR. The authors also would like to thank those concerned with this program.

References

Kageyama, A., and T. Sato	2004	The "Yin-Yang grid": An overset grid in spherical coordinates. *oai:arXiv.org:physics/0403123*
Parrish, D. F., and J. C. Derber	1992	The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. *Mon. Wea. Rev.,* **120,** 1747–1763.
Peng, X., F. Xiao, K. Takahashi	2004	Application of the CIP-CSLR on the "Yin-Yang" grid in spherical geometry. *The 2004 Workshop on the Solution of Partial Differential Equations of the Sphere*, 21 pp.
Purser, R. J.	2005	A geometrical approach to the synthesis of smooth anisotropic covariance operators for data assimilation. NOAA/NCEP Office Note 447, 60 pp.
Purser, R. J.	2007	Diamond interpolation: a class of accurate compact-stencil grid interpolation methods. NOAA/NCEP Office Note 454, 13 pp.
Purser, R. J.	2008a	Normalization of the diffusive filters that represent the inhomogeneous covariance operators of variational assimilation, using asymptotic expansions and techniques of non-Euclidean geometry; Part I: Analytic solutions for symmetrical configurations and the validation of practical algorithms. NOAA/NCEP Office Note 456, 48 pp.
Purser, R. J.	2008b	Normalization of the diffusive filters that represent the inhomogeneous covariance operators of variational assimilation, using asymptotic expansions and techniques of non-Euclidean geometry; Part II: Riemannian geometry and the generic parametrix expansion method. NOAA/NCEP Office Note 457, 55 pp.
Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts	2003	Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.,* **131,** 1536–1548.
Riishøjgaard, L.-P.	1998	A direct way of specifying flow-dependent background error correlations for meteorological analysis systems. *Tellus,* **50A,** 42–57.
Sato, Y., and M. S. F. V. De Pondeca, R. J. Purser, and D. F. Parrish	2009	Ensemble-based background error covariance implementations using spatial recursive filters in NCEP's Grid-point Statistical Interpolation system. NOAA/NCEP Office Note 459.
Ziman, J. M.	1979	*Principles of the Theory of Solids* Cambridge, 448 pp.