Spectral Transformation Using a Cubed-Sphere Grid for a Three-Dimensional Variational Data Assimilation System

HYO-JONG SONG* AND IN-HYUK KWON
Korea Institute of Atmospheric Prediction Systems, Seoul, South Korea

(Manuscript received 13 March 2014, in final form 12 March 2015)

ABSTRACT

Atmospheric numerical models using the spectral element method with cubed-sphere grids (CSGs) are highly scalable in terms of parallelization. However, there are no data assimilation systems for spectral element numerical models. The authors devised a spectral transformation method applicable to the model data on a CSG (STCS) for a three-dimensional variational data assimilation system (3DVAR). To evaluate the 3DVAR system based on the STCS, the authors conducted observing system simulation experiments (OSSEs) using Community Atmosphere Model with Spectral Element dynamical core (CAM-SE). They observed root-mean-squared error reductions: 24% and 34% for zonal and meridional winds (U and V), respectively; 20% for temperature (T); 4% for specific humidity (Q); and 57% for surface pressure (Ps) in analysis and 28% and 27% for U and V, respectively; 25% for T; 21% for Q; and 31% for Ps in 72-h forecast fields. In this paper, under the premise that the same number of grid points is set, the authors show that the use of a greater polynomial degree, np, produces better performance than use of a greater element count, ne, on equiangular coordinates in terms of the wave representation.

1. Introduction

In a massive computing environment, the parallel scalability that a numerical method can guarantee is a primary issue in developing the dynamical core for atmospheric modeling. If used effectively, a huge computing resource can allow us to achieve a resolution under which decoupling between global and regional modeling is not required, and a seamless approach to unified atmospheric modeling is possible. The spectral element method (SEM) using cubed-sphere grids (CSGs) is a highly scalable numerical method (Taylor et al. 1997; Fournier et al. 2004). Communication within each element is global; nonetheless, the elements need only boundary information from their neighboring elements for parallel computing. If each processor of a parallel system handles a minimal number of elements, the communication overhead becomes remarkably low (Dennis et al. 2005).

SEM has been applied primarily to climate simulations for atmospheric applications (Evans et al. 2013). Using an atmospheric numerical model for numerical weather prediction requires an adequate data assimilation system (Klinker et al. 2000; Rawlins et al. 2007). If a data assimilation system applicable to atmospheric numerical models using SEM is developed, we can expect SEM to be used also to build a high-resolution numerical weather prediction system.

Identification and modeling of the covariance matrix of background errors are critical tasks in data assimilation. Spectral transformations often function as horizontal filtering of background-error correlations (Courtier et al. 1998; Lorenc et al. 2000). To apply the Fourier and Legendre transformations to the fields on a CSG, however, we must interpolate the CSG variables onto a Gaussian grid (Temperton 1991). The interpolation yields numerical errors and data redistributions on multiple memory units. We thus devised a spectral transformation method for a CSG using equiangular coordinates (CSGEOA), abbreviated as STCS in this work.

* Current affiliation: Atmospheric Sciences Research Center, University at Albany, State University of New York, Albany, New York.

Corresponding author address: Hyo-Jong Song, Korea Institute of Atmospheric Prediction Systems, 4F, Hankuk Computer Building, 35 Boramae-ro 5-gil, Dongjak-gu, Seoul 156-849, South Korea.

E-mail: hsong2@albany.edu

DOI: 10.1175/MWR-D-14-00089.1

© 2015 American Meteorological Society
A formulation of STCS is derived in section 2, based on calculus on CSGEA, a nonorthogonal coordinate, given in the appendices. STCS is based on the representativeness of the spherical harmonic functions (SHFs) by CSG points. For the SHFs represented on the CSG points to be proper bases for a spectral transformation, they must be the eigenfunctions of the Laplace operator defined on the CSGEA. We verified the validity of the STCS methodology by comparing the estimated eigenvalues with analytical eigenvalues of the Laplace operator. The relevant results are delineated in section 3. We designed a cost function for a three-dimensional variational data assimilation system (3DVAR) using STCS as one of the components of a background-error covariance model for the 3DVAR. The formulations and the experimental results are shown in section 4. In section 5, we summarize the proposed methodology and the relevant results, and we address future works, including development possibilities of the STCS.

2. Formulation of STCS

To use a CSGEA for this work, we followed the explanation of calculus on CSGEA provided by Levy et al. (2009) as arranged in appendices A and B. In this study, we chose equiangular coordinates, because they provide a more uniformly spaced grid compared to an equidistant projection (Rančić et al. 1996).

We can define the inner product of two arbitrary real-valued functions, \( a(\lambda, \theta) \) and \( b(\lambda, \theta) \), in spherical coordinates as

\[
\langle a, b \rangle_{sp} = \frac{1}{4\pi R^2} \int_{-\pi/2}^{\pi/2} \int_{-\pi}^{\pi} a(\lambda, \theta) b(\lambda, \theta) R^2 \cos \theta \, d\lambda \, d\theta,
\]

where \( R \) is radius of a sphere, and \( \lambda \) and \( \theta \) are longitude and latitude, respectively. We obtain the spectral coefficients of an arbitrary function \( f_l^m \) using the inner product of the SHF \( Y_l^m \) and the function \( f \) on the sphere:

\[
f_l^m = \langle Y_l^m, f \rangle_{sp} = \frac{1}{4\pi R^2} \int_{-\pi/2}^{\pi/2} \int_{-\pi}^{\pi} Y_l^m(\lambda, \theta) f(\lambda, \theta) R^2 \cos \theta \, d\lambda \, d\theta.
\]

(3)

The SHF \( Y_l^m \) was normalized for the spectral coefficient to be 1 when \( f = Y_l^m \). It has the triangular real form:

\[
Y_l^m = \begin{cases} 
N_l^m P_l^m(\cos \theta) \cos m \lambda & \text{if } m > 0, \\
N_l^m P_l^m(\cos \theta) & \text{if } m = 0, \\
N_l^m P_l^m(\cos \theta) \sin |m| \lambda & \text{if } m < 0,
\end{cases}
\]

(4)

where the indices \( m \) and \( l \) are the zonal and spherical wavenumbers, respectively. Terms \( P_l^m \) and \( N_l^m \) are an associated Legendre polynomial and a normalization factor, respectively. In accordance with convention, we use a Gaussian grid to perform the integration in Eq. (3) when we do not use an integration formulation defined on CSGEA [see appendix A, Eq. (A7)]. First, we take the longitudinal integration over a latitudinal band, the Fourier transformation. Second, we take the latitudinal integration with the aid of Gaussian weights, the Legendre transformation (Temperton 1991).

Making use of Eq. (A7), we can write the formulation of the spectral transformation using the cubed sphere as

\[
f_l^m = \frac{1}{4\pi R^2} \sum_{i,j,k} Y_l^m(i,j,k) \sqrt{g} \, d\alpha \, d\beta.
\]

(5)

To discretize this integration, we use an element-based Galerkin approach, the spectral element method. Based on the interpolation rule using CSGEA in appendix B, we can discretize Eq. (5) as

\[
f_l^m = \frac{1}{4\pi R^2} \sum_{i,j,k} \hat{Y}_l^m (i,j,k) \sqrt{g} w_i w_j.
\]

(6)

This is the discretized formulation of STCS based on the spectral element method.

3. Characteristics of STCS

a. Dependency of the wave representation of a CSGEA on the element and the polynomial numbers

The formulation of STCS, Eq. (6), normally works when the SHFs discretized on a CSGA \( Y_l^m \), represent the analytic SHF \( Y_l^m \) well. To check whether \( \hat{Y}_l^m \) represents \( Y_l^m \) well or not, we used the eigenvalue of the Laplace operator as a measure of the accuracy of discretized SHFs, since the SHFs are the eigenfunctions of the horizontal Laplace operator on the sphere. The gradient, the divergence, and the Laplace operators on a CSGEA are as follows (Levy et al. 2009):

\[
\nabla g = \left( \frac{\partial}{\partial \alpha}, \frac{\partial}{\partial \beta} \right)^T,
\]

(7)

\[
\nabla f = A^{-1} \nabla g f,
\]

(8)

\[
\nabla \cdot f = \frac{1}{\sqrt{g}} \nabla g \cdot (\sqrt{g} A^{-1} f).
\]

(9)

Based on these operators, we can derive a strong Laplace operator on the CSGA as follows:
\[ \nabla^2 f = \nabla \cdot (\nabla f) \]
\[ = \frac{1}{\sqrt{g}} \nabla_g \cdot (\sqrt{g} A^{-1} A^{-T} \nabla_g f). \]

(10)
(11)

Since the Lagrange polynomial, \( \phi \), is \( C^0 \)-continuous, we need a weak Laplace operator, which can be defined by making use of \( \phi \) to define the Laplace operator on the CSGEA, as follows:

\[ L_f = \nabla^2 f, \]
\[ \tilde{L}_{f_{ij},e} = M_{i,j,ie}^{-1} \int_{\Omega_e} \nabla^2 f \sqrt{g} \phi_{ij} \, da \, d\beta \]
\[ = M_{i,j,ie}^{-1} \int_{\Omega_e} \frac{1}{\sqrt{g}} \nabla_g \cdot (\sqrt{g} A^{-1} A^{-T} \nabla_g f) \phi_{ij} \sqrt{g} \, da \, d\beta \]
\[ = -M_{i,j,ie}^{-1} \int_{\Omega_e} (A^{-1} A^{-T} \nabla_g f) \cdot (\nabla_g \phi_{ij}) \sqrt{g} \, da \, d\beta \]
\[ = \nabla^2_{wk} f, \]

(12)
(13)
(14)
(15)
(16)

where the mass-matrix element is

\[ M_{i,j,ie} = \int_{\Omega_e} \phi_{ij} \sqrt{g} \, da \, d\beta, \]

(17)

and the basis function \( \phi_{ij} \) is

\[ \phi_{ij} = \phi_{ij}(x, y, z) = \phi_{ij}(\alpha, \beta, \gamma). \]

(18)

The term \( \tilde{L}_{f_{ij},e} \) is the value at grid point \( i \) and \( j \) in element \( e \) obtained by applying the weak Laplace operator, \( \nabla^2_{wk} \), defined on the CSGEA, to an arbitrary function \( f \) on the sphere. Similarly to the inner product in spherical coordinates, \( \langle a, b \rangle_{sp} \), we can define another inner product for \( a(\alpha, \beta) \) and \( b(\alpha, \beta) \) on a CSGEA, as

\[ \langle a, b \rangle_{cs} = \frac{1}{4\pi \rho^2} \sum_{i} \int_{\Omega_e} a(\alpha, \beta) b(\alpha, \beta) \sqrt{g} \, da \, d\beta. \]

(19)

We project the weak Laplacian of a discretized SHF \( \nabla^2_{wk} \tilde{Y}^m_l \) onto \( \tilde{Y}^m_l \) to obtain the estimated eigenvalue \( \Lambda^m_l \):

\[ \Lambda^m_l = \langle \tilde{Y}^m_l, \nabla^2_{wk} \tilde{Y}^m_l \rangle_{cs}. \]

(20)

If \( \tilde{Y}^m_l \) approximates \( Y^m_l \) well, \( \Lambda^m_l \) can be expected to be equal to the analytical eigenvalue \( -l(l+1)/R^2 \) \cite{Courant and Hilbert 1962}. We define the error of the estimated eigenvalue \( e \) as follows:

\[ e = R^2 \Lambda^m_l + l(l+1). \]

(21)

Figure 1a shows \( \log_{10}(|e|) \) as a function of \( ne \) with a fixed \( np = 4 \) and \( np \) with a fixed \( ne = 16 \) for the SHF of spherical wavenumber \( l = 32 \). The term \( ne \) varies from 16 to 80 with an interval of 8 for \( np = 4 \), while \( np \) varies from 4 to 16 with \( ne = 16 \). To get the logarithmic value, \( |e| < 10^{-20} \) has been discarded. The \( |e| \) values of all zonal wavenumbers for \( l = 32 \) are plotted with assignment to the same degrees of freedom (DOF). (b) Accuracy-saturated spherical wavenumber as a function of DOF. We get a spherical wavenumber \( l \) for which CSGEA has \( |e| > 10^{-11} \), then \( l - 1 \) is defined as the accuracy-saturated spherical wavenumber of the CSGEA.

\[ \text{DOF} = 6(np - 1)^2 ne^2 + 2. \]

(22)

A CSGEA with a higher \( np \) represents the SHF more effectively than that with a higher \( ne \) given a DOF. The accuracy saturation level appears to be approximately \( 10^{-12} \). An average of \( |e| \) over spherical wavenumbers is proportional to \( \text{DOF}^{-3.01} \) at \( np = 4 \) \( (R^2 = 1) \). The quantity \( |e| \) grows as \( \tilde{Y}^m_l \) loses ortho-normality (not shown).
Based on the results of Fig. 1a, we can set an accuracy saturation level of wave representation of a CSGEA. If for a CSGEA with a combination of ne and np we get a spherical wavenumber \( l \) for which a CSGEA has \( |e| > 10^{-11} \), then \( l - 1 \) is defined as the accuracy-saturated spherical wavenumber for the CSGEA. Figure 1b explicitly supports the conclusion that the use of a higher np and a lower ne given the same DOF improves the ability of the CSGEA to represent the SHF. While increasing np can cause the CSGEA to resolve more SHFs at an accuracy saturation level, increasing ne is weakly effective in improving a CSGEA up to the accuracy limit. The accuracy-saturated spherical wavenumber is approximately proportional to \( \text{DOF}^{0.9} \) at ne = 4 (\( R^2 = 0.86 \)).

b. Dependency of the wave representation of a CSGEA on the zonal and spherical wavenumbers

When ne = 16 and np = 4, Fig. 2a shows the errors of estimated eigenvalues of the CSGEA for various zonal wavenumbers at fixed spherical wavenumbers and also for various spherical wavenumbers at fixed zonal wavenumbers. Wave representativeness of a CSGEA tends to depend on the spherical wavenumber, in other words how well it resolves the associated Legendre polynomials. It is possible that
this result is related to the irregularity in the positions of the roots of the associated Legendre polynomials.

The cases corresponding to \( m = 0 \) and \( l = 64 \) are shown in Fig. 2b to look further into the error distribution according to zonal and spherical wavenumbers. We have \( |e| \) for \( np = 4, 7, \) and \( 10 \) with \( ne = 16 \) and \( ne = 16, \) \( 32, \) and \( 48 \) with \( np = 4, \) for which the corresponding DOFs are 13 826, 55 298, and 124 418, respectively. The error in wave representation of a CSGEA increases as the spherical wavenumber grows at a fixed \( m = 0. \) This clear dependency of \( |e| \) on the spherical wavenumber supports the statement that the wave representativeness of a CSGEA is determined by how well the associated Legendre polynomials are resolved under a CSGEA configuration. On the other hand, configurations having higher \( ne \) values tend to have a relatively large surplus of \( |e| \) for a smaller \( l. \)

When the spherical wavenumber is fixed at 64 with a variable zonal wavenumber, the SHF with \( m = 45 \) is represented well by the CSGEA relative to \( m = 0. \) Eigenvalue estimation for zonal wavenumbers greater and less than 45 and \(-45, \) respectively, tend to have greater errors. This error distribution is related to geometric positions of CSGEA grids. For example, when \( ne = 16 \) and \( np = 4 \) for the SHF with \( m = 45 \) and \( l = 64, \) the number of waves covered by an element of the CSGEA is about one-half, while the number of waves covered by an element of the CSGEA is about one, for \( m = 64. \) The CSGEA represents the SHF better as the number of waves covered by an element of the CSGEA decreases. The error distribution over zonal wavenumber is similar regardless of the change of configurations. Improving the estimation by increasing \( np \) is much more effective than by increasing \( ne \) also in the error distribution according to zonal wavenumber.

In Fig. 2b, the values of \( |e| \) for \( m \) values of different signs are asymmetric around \( m = 0. \) To look into this phenomenon, we investigated the horizontal maps of a CSGEA with \( ne = 32 \) and \( np = 4 \) for \( m = 2 \) and \( l = 64 \) (Fig. 3). While the parts of the SHF with the biggest amplitude with \( m = 2 \) are located on the centers of the four cube faces passing through the equator, those of the SHF with \( m = -2 \) are on the edges (not shown) because a different sign of \( m \) implies a \( \pi/2 \) shift in \( \lambda. \) While the edge has the least value of the determinant of the metric tensor \( \sqrt{g}, \) the center of a cube face has the greatest value of \( \sqrt{g}. \) This indicates that the face center experiences a larger surface expansion occurring when a vector field on the equiangular coordinates is transformed into spherical coordinates. There are, thus, a relatively small number of grid points over the center of the cube face in which the largest expansion of
the space occurs, while there are a relatively large number of grid points over the edge of the cube face in which the smallest expansion of the space occurs. The asymmetry in the values of $|e|$ according to different signs of $m$ in Fig. 2b can be explained also by the reasoning that the CSGEA represents the SHF better as the number of waves covered by an element of the CSGEA decreases.

Therefore, the representativeness of a CSGEA can be expected to depend somewhat on the kind of map projection. The evaluation and comparison of a CSG with projections other than equiangular projection, such as equidistance central projection, might be an interesting research topic (Nair et al. 2005).

4. Application of STCS to 3DVAR development

a. Formulation of a 3DVAR using STCS (STCS-3DVAR)

1) A FORMULATION OF A BACKGROUND-ERROR COVARIANCE MODEL USING STCS

To construct a 3DVAR system based on STCS, we designed a background-error covariance model making use of STCS. Suppose that the forecast model is a hydrostatic model, and that the ns background-error samples $\delta X$ on the CSGEA are given as the differences of the forecasts at the same time issued from different times, or the ensemble deviations from the mean of an ensemble forecast as follows (Parrish and Derber 1992; Fisher 2003):

$$
\delta X = \begin{pmatrix}
\delta U(1,1,1,1,1) & \ldots & \delta U(1,1,1,1,ns) \\
\vdots & & \vdots \\
\delta U(i,j,k,ie,1) & \ldots & \delta U(i,j,k,ie,ns) \\
\vdots & & \vdots \\
\delta U(np,np,nk,6ne^2,1) & \ldots & \delta U(np,np,nk,6ne^2,ns) \\
\vdots & & \vdots \\
\delta V(np,np,nk,6ne^2,1) & \ldots & \delta V(np,np,nk,6ne^2,ns) \\
\vdots & & \vdots \\
\delta T(np,np,nk,6ne^2,1) & \ldots & \delta T(np,np,nk,6ne^2,ns) \\
\vdots & & \vdots \\
\delta Q(np,np,nk,6ne^2,1) & \ldots & \delta Q(np,np,nk,6ne^2,ns) \\
\vdots & & \vdots \\
\delta Ps(np,np,6ne^2,1) & \ldots & \delta Ps(np,np,6ne^2,ns)
\end{pmatrix}.
$$

(23)

Here $U, V, T, Q,$ and $Ps$ are the zonal and the meridional winds, the air temperature, the specific humidity, and the surface pressure, respectively. The indices $nk$ and $ns$ are the number of vertical levels and the number of samples to generate a background-error covariance $B$, respectively.

First, we calculated the error variance on the grid points of the cubed sphere, formulated as a diagonal matrix $D$. Using the error variance $D$, we are able to formulate $B$ as follows:

$$
B = D^{1/2}CD^{1/2}
$$

(24)

$$
= D^{1/2}S^{-1}CS^{T}S^{-T}D^{1/2}
$$

(25)

$$
= D^{1/2}S^{-1}CS^{T}S^{-1}D^{1/2},
$$

(26)

where $C$ is an error correlation matrix. A background-error correlation matrix in spectral space $C_s$ is defined as

$$
C_s = SCS^{T}.
$$

(27)

The linear operator $S$ converts the total model state from a model space to a spectral space using STCS as defined in Eq. (6). We can describe the inverse spectral transformation using the CSGEA (inverse STCS), $S^{-1}$, simply using the following synthesis process:

$$
f_{\text{syn}}(\alpha, \beta) = \sum_{lm} f_{lm}^{m} e^{i m} \left[ \lambda(\alpha, \beta), \theta(\alpha, \beta) \right].
$$

(28)

The matrices $S^{T}$ and $S^{-T}$ are the adjoint operators of the forward and inverse STCS.

To estimate $C_s$, the given samples are normalized by the error standard deviation $D^{1/2}$, and transformed into triangular spectral space using the STCS ($S$):

$$
\delta W = SD^{-1/2}\delta X.
$$

(29)

The matrix $\delta W$ is a matrix of the samples generating $C_s$. Assuming that the model state is derived from a hydrostatic model, $\delta W$ contains vertical profiles for each wavenumber in the following form:
The measurement

\[ J_b \quad \text{with} \quad \delta \mathbf{w} = \begin{pmatrix} \delta W_0^m \ 
\delta W_{-1}^m \\
\vdots \\
\delta W_L^m \\
\end{pmatrix} \] (30)

where

\[ \delta W_m^m = \begin{pmatrix} \delta U_1^m(1,1) & \ldots & \delta U_1^m(1,ns) \\
\vdots & \ddots & \vdots \\
\vdots & \ddots & \vdots \\
\delta U_{ns}^m(1,1) & \ldots & \delta U_{ns}^m(1,ns) \\
\end{pmatrix} \] (31)

and the indices \( M \) and \( L \) are the largest zonal and spherical wavenumbers, respectively. Now we can write a background-error correlation matrix in the spectral space as follows:

\[ \mathbf{C}_s = \frac{1}{ns-1} \delta \mathbf{w} \delta \mathbf{w}^T \] (32)

This is an error covariance of the transformed state \( \mathbf{w} \), the control variable for the STCS-3DVAR:

\[ \mathbf{w} = \mathbf{S} D^{-1/2} \mathbf{x} \] (33)

In Eq. (32), we used as a denominator \( ns - 1 \) because we assumed an unbiased estimate of the forecast distribution (Leith 1974). The control variable \( \mathbf{w} \) can be inverted as follows:

\[ \mathbf{x} = \mathbf{D}^{1/2} \mathbf{S}^{-1} \mathbf{w} \] (34)

Note that the correlation matrix \( \mathbf{C}_s \) allows multivariate correlations among different model variables that have the same wavenumber.

2) DESIGNING AND MINIMIZING A COST FUNCTION DEFINED BY STCS

A cost function designed by the background-error covariance model described above is

\[ J(\delta \mathbf{w}) = \frac{1}{2} \mathbf{w}^T \mathbf{C}_s^{-1} \delta \mathbf{w} + \frac{1}{2} \left[ \mathbf{y}^o - \mathbf{H}(\mathbf{x}^b + \mathbf{D}^{1/2} \mathbf{S}^{-1} \delta \mathbf{w}) \right]^T \mathbf{R}^{-1} \left[ \mathbf{y}^o - \mathbf{H}(\mathbf{x}^b + \mathbf{D}^{1/2} \mathbf{S}^{-1} \delta \mathbf{w}) \right] \] (35)

\[ \delta \mathbf{w} = \mathbf{w} - \mathbf{w}^b \] (36)

\[ \mathbf{w}^b = \mathbf{S} D^{-1/2} \mathbf{x}^b \] (37)

where \( \mathbf{x}^b \) is a background model state on the CSGEA, \( \mathbf{y}^o \) is an observation vector, \( \mathbf{R} \) is an observation error covariance, and \( \mathbf{H} \) is a vector-valued observation operator with \( \mathbf{x} \)-gradient \( \mathbf{H} \). The cost function \( J \) consists of the measurement \( J_s \) of a distance from a background and the measurement \( J_o \), of a distance from an observation. The \( \delta \mathbf{w} \) gradient of the cost function is given by

\[ \nabla J = \mathbf{C}_s^{-1} \delta \mathbf{w} - \mathbf{S}^T \mathbf{D}^{1/2} \mathbf{H}^T \mathbf{R}^{-1} [\mathbf{y}^o - \mathbf{H}(\mathbf{x}^b + \mathbf{D}^{1/2} \mathbf{S}^{-1} \delta \mathbf{w})]. \] (38)

Assuming the convexity of \( J \), the minimization process for \( J \) is performed by solving the following linear system:

\[ \nabla J = \mathbf{C}_s^{-1} \delta \mathbf{w} - \mathbf{S}^T \mathbf{D}^{1/2} \mathbf{H}^T \mathbf{R}^{-1} \delta \mathbf{w} \]

To derive this linear system, we used a tangent linear assumption for the observation operator:

\[ \mathbf{H}(\mathbf{x}^b + \mathbf{D}^{1/2} \mathbf{S}^{-1} \delta \mathbf{w}) = \mathbf{H}(\mathbf{x}^b) + \mathbf{H} \mathbf{D}^{1/2} \mathbf{S}^{-1} \delta \mathbf{w}. \] (40)

A conjugate gradient method is implemented to solve Eq. (39) (Hestenes and Stiefel 1952). To improve the convergence rate of the conjugate gradient solver, we performed a preconditioning of the cost function using a square root matrix of \( \mathbf{C}_s \), \( \mathbf{C}_s^{1/2} \).
where $C^{1/2}$ is obtained by using an eigen-decomposition module of LAPACK, DSYEV (available online at http://www.netlib.org/lapack). As a result, the STCS-3DVAR obtains the analysis field through the following final step:

$$x^a = x^b + D^{1/2}S^{-1}\delta w^a.$$  \hfill (43)

We used 122 forecast difference samples every 18 h from March to May for the static background-error covariance. The forecast difference is the subtraction of the nature run from a 6-h forecast. To add an error to the initial condition, we used the nature runs averaged at $-24, 0,$ and $24$ h from initial time with weights of 0.13, 0.74, and 0.13, respectively.

We have assimilated OSSE radiosonde and surface pressure observations using STCS-3DVAR with the static background-error covariance previously obtained from CAM-SE.

b. Experimental setting to test the STCS-3DVAR

To evaluate the STCS-3DVAR, observing system simulation experiments (OSSEs) were conducted using the Community Atmosphere Model with Spectral Element dynamical core (CAM-SE) with $n_e = 16$ and $n_p = 4$ (Evans et al. 2013). We assumed that the model run of CAM-SE with a year spinup using a default initial condition obtained from the website of the Community Earth System Model (available online at http://www. cesm.ucar.edu/models/cesm1.0/cam) is true and designated as “nature.” The number of vertical levels of the CAM-SE is 30, with the model top equaling approximately 3.6 hPa. The rank of the spherical background-error correlation matrix $C_s$ is 121; hence, the 122 samples are sufficient for spanning the vector space defined by the eigenvectors of each spectral background-error correlation matrix $C^{m^m}_s = [1/(n_s - 1)]\delta W^m_s\delta W^{m^m}_s$. A matrix $C^{m^m}_s$ represents a specific block in the block-diagonal matrix $C_s$, corresponding to the spherical wavenumber $l$ and the zonal wavenumber $m$.

We have assimilated OSSE radiosonde and surface pressure observations using STCS-3DVAR with the static background-error covariance previously obtained from CAM-SE.
FIG. 6. Background-error standard deviations (BESDs) of the: horizontal winds (m s\(^{-1}\)) (a),(b) zonal and (c),(d) meridional; (e),(f) air temperature (K); (g),(h) specific humidity (kg kg\(^{-1}\)), at model levels (left) 24 and (right) 15; and surface pressure (hPa) estimated by the STCS-3DVAR. Model levels 15 and 24 correspond approximately to 274 and 860 hPa, respectively.
obtained. Figure 4 represents 671 radiosonde stations and 4871 surface observation stations chosen for gathering data in the National Centers for Environmental Prediction (NCEP) binary universal form (BUFR) data. We assume that every radiosonde has 30 vertical levels that are the same as those of the nature run. The zonal wind (U), meridional wind (V), temperature (T), and specific humidity (Q) of the nature run have been horizontally interpolated into radiosonde location at each model level. Similarly, surface pressure (Ps) has been horizontally interpolated into surface observation locations. To conduct the horizontal interpolation from the CSGEA to the observation positions, we first projected the model state into a cube, which is rectangular, and then conducted the bilinear interpolation using the distance in the Cartesian coordinates (Nair et al. 2005). After the interpolation, Gaussian random errors with zero means, and standard deviations, 0.8 m s\(^{-1}\), 0.8 m s\(^{-1}\), 0.5 K, 0.05 kg kg\(^{-1}\), and 0.8 hPa, were added into the variables U, V, T, Q, and Ps, respectively.

To see a performance of the STCS-3DVAR and its effect in forecasts, eight cases from 0000 UTC 10 March to 1800 UTC 11 March were selected, and the 6-h forecasts at those times were used as the background. We diagnosed the results of the analysis at each time and ran the CAM-SE model for 72 h with the analyses as initial conditions.

To set the spherical wavenumber \( l \) for the STCS, we examined the difference between an original variable of a CAM-SE forecast \( x_{\text{org}} \) and a synthesized variable \( x_{\text{syn}} \) obtained by applying the STCS \( S \) and the inverse STCS \( S^{-1} \):

\[
x_{\text{syn}} = S^{-1} x_{\text{org}}.
\]

Figure 5 shows the normalized RMSD of the first sample among the samples for generating the background-error
FIG. 10. (left) The background error and (right) the error reduction that is the subtraction of the absolute value of the analysis error from that of the background error for (top to bottom) the zonal and meridional winds, air temperature, and specific humidity at model level 24 at 0000 UTC 10 Mar.
covariance. Consistent with the results of the eigenvalue estimation in Fig. 2a, the errors in the synthesis by the STCS and inverse STCS for all variables increase sharply after \( l = 63 \). The errors in synthesis processes of \( U, V, \) and \( Q \) are larger than those of \( T \) and \( P_s \). Based on the result of this synthesis process and the previous test of the eigenvalue estimation, for a CSGEA with \( n_e = 16 \) and \( n_p = 4 \) we used the SHFs of up to \( l = 63 \) to examine the STCS-3DVAR.

c. Test results of the STCS-3DVAR

1) Structure of a background-error covariance

We first checked the background-error standard deviations (BESDs) on the CSG points \( D^{1/2} \) to determine the structure of background-error covariance. The BESDs of \( U, V, \) and \( T \) are relatively large over the midlatitudes between 30° and 50°N and the ocean near Antarctica, while those of \( Q \) are large in the tropics for both the upper and lower troposphere (Figs. 6a–h). Meanwhile, the BESDs of \( P_s \) have the most similar spatial distribution with lower-tropospheric temperature BESDs (Fig. 6i). The horizontal scale difference between upper and lower levels is not conspicuously large, but the horizontal scale of \( T \) at level 15 is greater than that at level 24 (Figs. 6e,f). In Fig. 6, model levels 15 and 24 correspond to approximately 274 and 860 hPa, respectively.

The diagonal entries of the background-error correlation matrix, \( \text{diag}(C) \), express autocorrelations represented

![Fig. 11. The cross section of zonal mean RMSE reduction by the STCS-3DVAR analysis at 0000 UTC 10 Mar. for each (a) U, (b) V, (c) T, and (d) Q.](image-url)
3) ERROR REDUCTION IN ANALYSIS

Figure 10 shows the background error that deviates from nature and the error reduction due to data assimilation at model level 24 for $U$, $V$, $T$, and $Q$ at 0000 UTC 10 March. Background errors of $U$, $V$, and $T$ are relatively large over the North Pacific and the ocean near Antarctica, which are similar to the BESD (Fig. 6). While the background error of $U$ has a dipole structure with positive (negative) deviation over north (south) Kamchatka Peninsula (Fig. 10a), the error reduction has the same structure, but with both cores being positive. That indicates that the analysis increment is negatively correlated with the background error, hence, the analysis error is less than the background error (Fig. 10b). The error reduction of $V$ is even bigger than $U$ (Fig. 10d). While the greater parts of the temperature error and the horizontal wind reduction are positive, a distinct negative background error of temperature was found around Tasmania in Australia (Fig. 10e). It was captured well as an analysis increment with the opposite sign, thus, the error reduction around this area is significantly larger than others (Fig. 10f). As the background of specific humidity is distributed over tropical and low-latitudinal regions, the error reduction at high latitude was not obvious. Dense observation regions such as Europe, East Asia, and North America have positive error reductions (Fig. 10g) but the error reduction is mostly neutral over the Pacific Ocean because of a lack of $Q$ observation (Fig. 10h).

The error reduction is mostly positive for not only low levels but also mid- and upper levels (Fig. 11). The error reductions of horizontal wind are apparent in zonal bands 30°–80°N and 40°–60°S and concentrated on about level 15 (Figs. 11a,b). The error reduction of $V$ is evidently correlated to that of temperature. In Fig. 11b, the levels of error reduction peak in the Southern and Northern Hemispheres are 14 and 16, respectively. Meanwhile, temperature error reduction has double cores on the same latitude, and levels 14 and 16 cut through upper and lower cores of the temperature error reduction in the Southern and Northern Hemispheres, respectively (Fig. 11c). The error reduction of $Q$ did not reach the upper levels, because there is little humidity above model level 20 (Fig. 11d).

The analysis increments of $U$ and $V$ at level 15 are greater than those at level 24 (Fig. 12). This fact accounts for the greater error reduction at level 15 in Fig. 11 as well. The analysis increments of $U$, $V$, and $T$ are dominant at scales of spherical wavenumbers 10–20 (Fig. 12), which is identical to autocorrelation of the background-error covariance (Fig. 7). The majority of analysis increment scales for $Q$ are higher spherical wavenumbers greater than 10.

![Spherical wavenumber vs. Analysis increments](image)
In Fig. 13, we can see that the most severe error of Ps changes to a positive error reduction even on the Southern Ocean (Figs. 13a,b), because the surface observations are distributed well on both hemispheres (Fig. 4).

Figure 14 presents background and analysis RMSEs of $U$, $V$, $T$, $Q$, and Ps for eight experimental cases. All analysis RMSEs are less than background RMSEs for whole variables and cases. Through the STCS-3DVAR, Ps definitely shows the best improvement owing to the dense observations. While the error of $V$ is greater than that of $U$, the error reduction is larger in $V$. The STCS-3DVAR shows a fairly good effect on $T$, but minimal effect on $Q$. On the average, RMSE reductions in analysis are 24% and 34% for $U$ and $V$, respectively; 20% for $T$; 4% for specific humidity $Q$; and 57% for Ps in the OSSE.
4) ERROR REDUCTION IN FORECAST

We conducted CAM-SE for 72 h using the analyses as the initial condition. Overall, RMSEs of forecasts are greater than those in Fig. 14, and the RMSEs of forecasts with analyses persist in being less than those of forecasts without analysis (Fig. 15). The RMSE error of the Ps forecast is not dramatically reduced compared to Fig. 14d, but it is still less than the forecast from background. It is noticeable that the difference between errors of Q forecasts has been greater than the initial (Fig. 15c). In summary, RMSE reductions in the 72-h forecast fields are 28% and 27% for U and V, respectively; 25% for T; 21% for Q; and 31% for Ps in the OSSE. It seems that the STCS-3DVAR has been assimilating observation data with the background well, and also maintaining the model’s balance.

5. Summary and discussion

Atmospheric numerical models using the spectral element method with a CSG are very scalable in terms of parallelization. However, for the spectral element numerical models to be the next-generation numerical weather prediction mode systems, corresponding data assimilation systems must be developed. As part of meeting such a need, we developed STCS. To devise STCS, we used SHFs represented on the CSG points and the spectral element method using equiangular coordinates, which give a more uniformly spaced grid compared to an equidistant projection (Rané et al. 1996).

To examine the accuracy of the method, we used the eigenvalues of the Laplace operator defined in a CSGEA. Given that DOF, Eq. (22), is the same, a CSGEA with small ne and large np can represent SHFs better than a CSGEA with large ne and small np. The wave representativeness of a CSGEA tends to depend on the spherical wavenumber, in other words, how well it resolves the associated Legendre polynomials. Probably this result is related to the irregularity in the positions of the roots of the associated Legendre polynomials. One thing to notice is that the STCS occurs in one time step and can be expected to incur a minor cost compared with model integrations. The choice of np...
and ne is primarily dictated by model integration efficiency. Using large np and small ne would be suboptimal for parallelization and increase of time step size. Thus comparison experiments for variable np have been performed with small sizes of np, even less than four (Lauritzen et al. 2014). This paper raises the question of which ne and np are the most efficient choice for weather prediction modeling using CSGs including a data assimilation step.

In the OSSE using STCS-3DVAR, we observed root-mean-squared error reductions in analysis, 24% and 34% for U and V, respectively; 20% for T; 4% for Q; and 57% for Ps. The Ps definitely shows the most improvement owing to the dense observations. While the error of V is greater than that of U, the error reduction is greater in V. The STCS-3DVAR shows fairly good effect on T but minimal effect on Q. We conducted CAM-SE for 72 h using the analyses as the initial conditions. Root-mean-squared error reductions in 72-h forecast fields were 28% and 27% for U and V, respectively; 25% for T; 21% for Q; and 31% for Ps. The RMSE error of the Ps forecast is not dramatically reduced compared to Fig. 14d, but it was still less than the forecast from the background. It is noticeable that the difference between errors of Q forecasts was greater than the initial (Fig. 15e). It seems the STCS-3DVAR has been assimilating observation data with the background well, and also maintains the model’s balance.

Provided that the number of grid points over one side of a rectangular model domain is N, the proposed spectral transformation method requires approximately $O(N^4)$ multiplications and additions. The method is $O(N^2)$ times costlier than the operation of the conventional spectral transformation method, $O(N^3)$. Using STCS, however, we can avoid interpolation errors and redistribution of model data on multiple central processing units (CPUs) to perform spectral transformations. When considering the calculation in model grid space, a load balancing of the proposed method will be better than the conventional spectral transformation, which consists of the Fourier and Legendre transformations. Following the development of fast Legendre transforms, applying the butterfly

![Fig. 15. RMSEs of 72-h forecasts from background and analysis for the (a) zonal and meridional winds, (b) air temperature, (c) specific humidity, and (d) surface pressure. The numbers on the x axis denote initial dates for model forecasts.](image)
algorithm to STCS might be expected to improve STCS efficiency (Wedi et al. 2013).

Acknowledgments. The authors are very grateful to Dr. Aimé Fournier and two anonymous reviewers for invaluable comments for improving this paper. Dr. Ji-Sun Kang supplied the CAM-SE run results, and Dr. In-Sun Song provided helpful comments about the characteristics of spherical harmonic functions represented on a CSG. Dr. Sang-Yoon Jun provided the parallel I/O modules for the STCS-3DVAR, and Dr. Mark Taylor provided a reference code that helped the authors familiarize themselves with the spectral element method using a CSG. The first author received supplemental support from the Basic Science Research Program of the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2014R1A6A3A03059247).

APPENDIX A

Definition of Areal Integration on the Cubed Sphere Using Equiangular Coordinates

A cubed sphere consists of six curved surfaces covering a sphere, with approximately equal spherical quadrilateral elements on each cube face (Fig. A1a). We call the number of elements on one side of a cube face the index ne and the number of grid points on one side of an element the index np. In Fig. A1, the index ne is equal to 4, and the index np is equal to 6.

We identify each grid point on the cubed sphere using equiangular coordinates (Fig. A1b). When the radius of a sphere is \( R \), and the longitude and latitude are \( \lambda \) and \( \theta \), respectively, a small displacement on the sphere \( dr \) is defined as

\[
dr = R \cos \theta d\lambda \hat{e}_\lambda + R d\theta \hat{e}_\theta,
\]

where \( \hat{e}_\lambda \) and \( \hat{e}_\theta \) are unit vectors along the longitudinal and latitudinal directions, respectively. The equiangular coordinates, \( \alpha \) and \( \beta \), are defined according to the position of each cube face. The coordinate \( \alpha \) is equal to \(-\pi/4 \) at the left end and \( \pi/4 \) at the right end of a cube face. Another coordinate \( \beta \) varies along the line of a fixed \( \alpha \) and is equal to \(-\pi/4 \) at the bottom of a cube face and \( \pi/4 \) at the top. The two unit vectors of the equiangular coordinates are, therefore, nonorthogonal. This coordinate system is, in other words, a curvilinear coordinate system, where the unit vectors, \( a_1 \) and \( a_2 \), are written as

\[
a_1 = \frac{\partial r}{\partial \alpha}, \quad a_2 = \frac{\partial r}{\partial \beta}.
\]

A vector on the cubed sphere \( \mathbf{v} \) can be expressed using the contravariant components, \( v^1 \) and \( v^2 \), and the covariant unit vectors \( a_1 \) and \( a_2 \) as follows:

\[
\mathbf{v} = v^1 a_1 + v^2 a_2.
\]

Transforming the contravariant components in equiangular coordinates to the orthogonal components in spherical coordinates uses the following matrix:

\[
A = (a_1 \ a_2) = R \begin{pmatrix}
\cos \theta \partial \lambda/\partial \alpha & \cos \theta \partial \lambda/\partial \beta \\
\partial \theta/\partial \alpha & \partial \theta/\partial \beta
\end{pmatrix}.
\]

Using \( A \), we can define a metric tensor \( g_{ij} \),

\[
g_{ij} = a_i \cdot a_j = (A^T A)_{ij}.
\]

On all six faces of the cubed sphere, \( g_{ij} \) is

\[
(g_{ij}) = \frac{R^2}{\rho^3 \cos^2 \alpha \cos^2 \beta} \begin{pmatrix}
1 + \tan^2 \alpha & -\tan \alpha \tan \beta \\
-\tan \alpha \tan \beta & 1 + \tan^2 \beta
\end{pmatrix},
\]

where \( \rho = (1 + \tan^2 \alpha + \tan^2 \beta)^{1/2} \). When \( g = \det(g_{ij}) \), \( \sqrt{g} = R^2/(\rho^3 \cos^2 \alpha \cos^2 \beta) \). Using the determinant of \( g_{ij} \), we can define the integration of a function \( f \) on the CSGEA as

\[\int \cdots \cdots \]
\[
\int_{-\pi/2}^{\pi/2} \int_{0}^{2\pi} f(\lambda, \theta) R^2 \cos \theta \, d\lambda \, d\theta = \sum_{ie} \int_{\Omega_{ie}} f(\lambda(\alpha, \beta), \theta(\alpha, \beta)) \sqrt{g} \, d\alpha \, d\beta, \quad (A7)
\]
where \(\Omega_{ie}\) specifies element number \(ie\) on the cubed sphere.

**APPENDIX B**

Element-Based Galerkin Approach to Discretization of an Integration

To discretize the integration, Eq. (5), on a CSGEA, we can use the Lagrange polynomials as basis functions. Provided that \(\psi_n\) is the \(n\)th-order Legendre polynomial, the Lagrange polynomial in one dimension \(\phi_i\) is defined as

\[
f(\alpha, \beta) \approx \tilde{f}(\alpha, \beta) = \sum_{ie=1}^{6n_e} \sum_{j=1}^{np} \sum_{i=1}^{np} f_{ijkl} \phi_i(\alpha, \beta) \phi_j(\alpha, \beta) \phi_k(\alpha, \beta) \phi_l(\alpha, \beta), \quad (B3)
\]
where the subscript \(j\) indicates the \(j\)th grid point over one side of the \(\beta\) direction of an element. The tilde accent indicates that a function has been discretized on the CSGEA using this element-based Galerkin method, the spectral element method (Taylor et al. 1997). Because the Lagrange polynomial \(\phi_i\) is equal to 1 only on the grid point \(\xi_i\), we can obtain \(\tilde{f}_{ijkl}\) by simply selecting the value of \(f\) on a grid point \((i, j)\) of an element \(\Omega_{ie}\).

REFERENCES


