Ensemble singular vectors as additive inflation in the Local Ensemble Transform Kalman Filter (LETKF) framework with a global NWP model

Seoleun Shin | Ji-Sun Kang | Shu-Chih Yang | Eugenia Kalnay

We test an ensemble data assimilation system using the four-dimensional Local Ensemble Transform Kalman Filter (4D-LETKF) for a global numerical weather prediction (NWP) model with unstructured grids on the cubed-sphere. It is challenging to selectively represent structures of dynamically growing errors in background states under system uncertainties such as sampling and model errors. We compute Ensemble Singular Vectors (ESVs) in an attempt to capture fast-growing errors on the subspace spanned by ensemble perturbations, and use them as additive inflation to enlarge the covariance in the area where errors are flow-dependently growing. The performance of the 4D-LETKF system with ESVs is evaluated in real data assimilation, as well as Observing System Simulation Experiments (OSSEs). We find that leading ESVs help to capture fast-growing errors effectively, especially when model errors are present, and that the use of ESVs as additive inflation significantly improves the performance of the 4D-LETKF.

KEYWORDS
atmospheric global model, ensemble data assimilation, ensemble singular vectors, local ensemble transform Kalman filter, numerical weather prediction

1 | INTRODUCTION

The Ensemble Kalman Filter (EnKF) can provide useful information on the flow-dependent background and analysis errors for a numerical weather prediction (NWP) system. NWP systems typically have millions of degrees of freedom for state variables, while the EnKF uses many fewer ensemble members to estimate background state covariance. Efforts have been made to properly account for sampling errors associated with the limited number of ensemble members, and to well represent model uncertainties in cycling of forecasts and analyses. For example, an application of multiplicative inflation prevents the underestimation of background error covariance by computing inflation factors in consideration of the differences between observation errors and ensemble spread scales (e.g. Miyoshi, 2011). There are also relaxation methods such as relaxation to prior spread (Whitaker and Hamill, 2012) and relaxation to prior perturbation (Zhang et al., 2004) that change analysis only where observations are assimilated. Whitaker and Hamill (2012) recommended relaxation methods when observations are dense (e.g. Kotsuki et al., 2017), while suggesting that additive inflation better treats model errors which do not depend on the distribution of observation. Additive inflation algorithms are suggested and widely used in the community of EnKF to perturb the ensemble subspace spanned by original ensemble members, by capturing growing errors missed by original ensembles (Corazza et al., 2003; Houtekamer et al., 2005; Whitaker et al., 2008). For example, it has been suggested using a randomly sampled normalized difference between 6 and 12 h
forecasts for one valid time, for each ensemble perturbation, which could handle such problems as sampling and model errors in terms of atmospheric instability (e.g. Whitaker et al., 2008). However, Yang et al. (2015) questioned whether those approaches can selectively capture fast-growing error modes that would be essential information to estimate background error covariance. It has been a challenge to properly target the errors-of-the-day in background states when we use the traditional method of additive inflation such as Whitaker et al. (2008) (Houtekamer and Zhang, 2016). If additive inflations are added in areas of sparse observation network, unnecessary noise can be accumulated (Enomoto et al., 2015; Yang et al., 2015).

Bishop and Toth (1999) discussed the application of ensemble transform technique in finding “fast growing perturbation on the subspace of ensemble perturbations.” They derived the ensemble-based singular vectors (SVs), which are a set of vectors that maximize given initial and final norms during a given period. Leutbecher and Lang (2014) extensively investigated the role of SVs in assessing the ability of ensemble forecasts in capturing fast-growing error modes, and for the representation of initial uncertainties. Meanwhile, Enomoto et al. (2006) derived Ensemble Singular Vectors (ESVs) by using the concept of a Lagrange multiplier. Recently, Yang et al. (2015) first suggested that additive perturbations need to be flow-dependent and proposed the use of leading ESVs as additive inflation in the Local Ensemble Transform Kalman Filter (LETKF) framework. They used a quasi-geostrophic model in Observing System Simulation Experiment (OSSE) tests, and showed that the performance of LETKF was improved.

In this study we compute ESVs for a global NWP model, and examine if ESVs can capture fast-growing error modes selectively. We also investigate whether the use of ESVs as additive inflation can improve the capability of the LETKF analysis to produce more accurate initial conditions due to a better representation of background error covariance, in comparison with the use of the traditional additive inflation method that includes randomly sampled adjacent forecast differences, and with the use of multiplicative inflation only. Tests are done for real data assimilation as well as OSSE formulations. We have found that the use of ESVs as additive inflation consistently improves the performance of LETKF in all test sets, and its impact is especially significant when model errors are assumed to be present.

In the following sections we introduce the global NWP model used in our tests, and briefly summarize the derivation of ESVs (Enomoto et al., 2015; Yang et al., 2015), as well as the LETKF system configuration. In section 3, we present the list of tests done for the evaluation of the LETKF performance when ESVs are used as additive inflation perturbations. In section 4, results from those tests are explained in detail, and then we discuss the use of ESVs to improve the analysis accuracy for global NWP in the final section.

2 | METHODOLOGY

2.1 | Forecast model

The forecast model that is being developed at KIAPS (Korean Institute of Atmospheric Prediction Systems) is using the spectral element method for horizontal, and the finite difference method for vertical spatial discretization of the governing equations (Taylor et al., 1997), formulated on the cubed-sphere to avoid singularity problem at the Poles (Sadourny, 1972). We have used the hydrostatic version of the NWP model developed at KIAPS in this study. This hydrostatic model has indeed the same set of governing equations and discretization methodology as the spectral element hydrostatic dynamic core of High Order Method Modeling Environment (HOMME: Dennis et al., 2012), although there are some differences. A major difference between the standalone HOMME and the NWP model used in this study is an infrastructure such as modeling framework, input–output (IO) system, and a 50-level vertical coordinate extended up to about 60 km, as well as a physical parametrization package coupled to the dynamical core (Shin et al., 2014). We use the same spatial resolution as in Shin et al. (2016), and the average grid spacing at the Equator is 1° with the minimum grid spacing of 0.83° (Evans et al., 2013). The simulations were run using the initial atmospheric states and ancillary data including sea-surface temperature (SST), snow and ice obtained from the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) analysis (Environmental Modeling Center, 2003).

2.2 | The LETKF data assimilation system at KIAPS

An ensemble data assimilation system in this study is based on the four-dimensional LETKF (thereafter 4D-LETKF) formulation by Hunt et al. (2007), and an implementation methodology by Miyoshi and Yamane (2007) is used for its development. In this section we briefly introduce the main idea of LETKF. More details on the LETKF algorithm and its implementation can be found in Hunt et al. (2007), and additional information on the LETKF data assimilation system at KIAPS in Shin et al. (2016).

Since the forecast model is formulated with fully unstructured quadrilateral meshes based on the cubed-sphere grid which is distributed irregularly when projected on the longitude and latitude grids, more careful examination is required to search for the closest background points surrounding the position of observation within an observation (forward) operator. The Jordan Curve Theorem is applied for a search algorithm to sample grid points enclosing the position of observation nearby in our LETKF system (Shin et al., 2016). As a result, remapping between two different grid structures is not required within an observation operator. Advantages are no loss of accuracy due to those repeated interpolations during the data assimilation (Terasaki et al., 2015), and reduced computational cost for remapping, which can
increase rapidly as the model resolution increases. We use the Fortran90 code developed by Miyoshi et al. (2010) where the parallelization is designed independently of geographical areas in consideration of load balancing. Recently, Yamazaki et al. (2017) suggested an efficient search algorithm for finding nearest-neighbour for regular Gaussian latitude–longitude grids. We may also use a search algorithm for finding neighbouring points on the cubed-sphere effectively (e.g. Kim et al., 2018) for error covariance localization in the future.

In this study we use the adaptive multiplicative covariance inflation suggested by Miyoshi (2011) as a tool to treat the problem of underestimation of error variance associated with sampling error. The adaptive multiplicative inflation is implemented within the LETKF framework, in such a way that the inflation parameters are updated at each grid point which needs less effort for manual tuning (Miyoshi, 2011). We choose a prior variance of the inflation parameter which allows a relatively strong temporal smoothing, since it results in a better performance of the LETKF (see more details in Shin et al. (2016)).

2.3 Ensemble singular vectors (ESVs)

Bishop and Toth (1999) show how to obtain SVs of the ensemble-based prediction and analysis error covariance matrices when the SVs are defined in terms of cost functions that SVs maximize. Enomoto et al. (2015) use the concept of Lagrange multiplier for the derivation of ESVs which are equivalent to SVs derived in Bishop and Toth (1999). Here we briefly summarize the definition and the derivation of ESVs following Nishii and Nakamura (2010), Yang et al. (2015) and Kawai et al. (2017).

Suppose that we have a set of initial perturbations denoted by \( Y=(y_1, y_2, \ldots, y_M) \), and their linear combination \( y \) defined by

\[
y = p_1 y_1 + p_2 y_2 + \cdots + p_M y_M = Yp,
\]

where \( p = (p_1, p_2, \ldots, p_M)^T \) is the vector of optimal coefficients. Likewise, the linear combination of final perturbation \( z \) is given by

\[
z = p_1 z_1 + p_2 z_2 + \cdots + p_M z_M = Zp,
\]

where \( Z=(z_1, z_2, \ldots, z_M) \) are final ensemble perturbations. Now the question is to find \( p \) that maximizes the norm \( \|y\| \) given \( \|y\| = 1 \) within an optimal time window. Nishii and Nakamura (2010) used the dry energy norm, but we choose the definition of a norm for a linear combination of perturbations \( \|x\| = p^T X F X p \) as moist energy given by

\[
\|x\| = \frac{1}{2} \int_A u'^2 + v'^2 + \frac{C_p}{T_r} T'^2 + R_T \left( \frac{p'}{p_r} \right)^2 + L_v q dA dp,
\]

where \( F \) is a matrix to define the norm of perturbations for a target area (Nishii and Nakamura, 2010). We denote \( u' \) as zonal wind, \( v' \) meridional wind, \( T' \) temperature, and \( p' \) surface pressure perturbations, with \( C_p \) specific heat at constant pressure, \( T_r \) reference temperature, \( p_r \) reference surface pressure, \( L_v \) latent heat for water vapour, \( q \) specific humidity, \( A \) the area of a column, and \( R \) the gas constant. Here, we take \( T_r = 300 \text{ K} \). Also, the norm of \( \|z\| \) given \( \|y\| = 1 \) can be expressed as

\[
y = p^T Y^T G Y \ p, \quad \text{and} \quad z = p^T Z^T H Z \ p.
\]

where \( G \) and \( H \) are matrices as \( F \) to define the norm of initial and final perturbations, respectively. Then a function \( f(p, \lambda) \) is defined using the idea of Lagrange multiplier (Enomoto et al., 2006) as

\[
f(p, \lambda) = p^T Z^T H Z \ p - \lambda (p^T Y^T G Y \ p - 1),
\]

and to find \( p \) that maximizes \( \|z\| \) given \( \|y\| = 1 \), we obtain the following equation:

\[
\frac{df(p, \lambda)}{dp} = 2(Z^T H Z \ p - \lambda Y^T G Y \ p) = 0.
\]

This leads to an eigenvalue problem:

\[
Z^T H Z \ p = \lambda Y^T G Y \ p.
\]

Yang et al. (2016) showed that the results can be sensitive to the choice of the norm when using either potential vorticity or stream function. In this study, we have chosen the moist static energy as the norm. With this choice, we avoid any transformation of variables and the computation is simple.

2.4 Computations of ESVs and implementation of additive inflation

Here we explain procedures to compute ESVs according to the equations described in section 2.3. We use notation in section 2.3, and also adopt similar notation and regulations for global model states as those used in Hunt et al. (2007).

1. The inputs to the computation of ESVs are not total values but ensemble perturbations. Therefore, we first obtain the ensemble average for \( M_{[g]} \), the dimensional analysis, the subscript \( g \) denotes global states, and \( M \) the ensemble size (Hunt et al., 2007). They are subtracted from original states of analysis at the beginning of the assimilation window at \( t=6 \), to yield initial perturbations \( y_i \), \( i=1, \ldots, M \). Likewise, \( z_i \), \( i=1, \ldots, M \) are obtained from the background states at the current analysis time \( t \).

2. Those perturbations are used to compute the moist energy of a specific domain of interest. The operator matrices \( G \) and \( H \) are symbolic operators that represent the energy norm defined for a specific area. Here we define \( G_N \) and \( G_S \) separately as the energy norm of Northern and Southern Hemisphere respectively.

3. The next step is to compute \((Y^T G Y)^{-1}\) and \(Z^T H Z\) to form the eigenvalue problem (Equation 7) in section 2.3. The inverse matrix \((Y^T G Y)^{-1}\) is efficiently calculated by using its positive definite and symmetric property. Using
its eigenvalues, eigenvectors and its transpose, the inverse matrix can be easily obtained.

4. Then the matrix \((Y^\top G Y)^{-1}\) is multiplied by \(Z^\top H Z\)
   (Equation 7), and we obtain the ESVs which are derived using the moist energy norm defined for the chosen subdomains.

5. Between the solutions to the eigenvalue problem (Equation 7), initial and final ESVs, we use the initial ESVs (IESVs), which grow faster during a given time window (Yang et al., 2015), as additive inflation perturbations.

6. Interestingly, only about half of all the available IESVs can capture growing error modes, which was also observed in Yang et al. (2015). Therefore, we select only such IESVs of growing error modes, and multiply minus sign to those growing IESVs to prepare positive–negative pairs of ESVs as described in Yang et al. (2015). As they have suggested, we also randomize the order of IESVs before we add them to the analysis ensemble.

7. Finally, we begin data assimilation and produce analysis states at the current time \(t\) after IESVs are obtained using the analysis at \(t-6\) and background at the current time \(t\). Since we observed that the structure of IESV does not change significantly within the 6 h time window (Yang et al., 2015), those IESVs are multiplied by a scalar factor 0.2 chosen by experimentation and are added to the current analysis states at time \(t\). We found that, within our current system, the model blows up in time if we choose a scalar factor larger than 0.3. The orders of the magnitude of analysis increments and additive inflation are similar to each other when the scaling factor 0.2 is used. However, the magnitude of perturbation becomes stronger when the scaling factor larger than 0.3 is used. This implies that too large perturbations have been added to the analysis in dynamically active regions where errors grow fast and possibly lead to the model blow-ups. It suggests that the ensemble can be sensitive to what have been added with ESVs and also brings up the property of ESVs indicating the dominant growing modes.

3 | EXPERIMENTAL SET-UP

The model resolution is chosen to be 100 km, and the number of ensemble members is 30. The test period of every 6 h forecast–analysis cycle in this study is 1–28 February 2014. The initial states of 30 ensemble members are obtained by arbitrarily choosing 30 model states from long-term simulation of the model so that the initial error of the ensemble mean is supposed to be quite large. We attempt to examine if such a large initial error decreases as forecast–analysis cycles continue, and if the error becomes smaller when ESVs are used as additive inflation in the LETKF framework.

There are three sets of experiments (Table 1): two sets of experiments for OSSEs, and one experiment set for real data assimilation. In the first experiment set (PerfModel), we attempt to examine the performance of the LETKF system with three different inflation methods under the perfect model assumption. In this OSSE formulation, we assume a single model run using the forecast model as a true state (nature), and generate simulated observations by projecting the true state into the observational space where spatial and temporal information of observations are obtained from the real NCEP Prep-Buf data. Observation errors of realistic scale are assigned to simulate conventional data such as sondes, surface pressure observations, satwinds, etc. Since we simulate conventional observations, typically the volume of observation is larger at 0000 and 1200 UTC than at 0600 and 1800 UTC. Also, more observations are distributed over land than over ocean, and over the Northern Hemisphere than the Southern Hemisphere.

In the second experiment set (ImPerfModel), we plan another set of OSSEs with an imperfect model scenario, and thus use SST and soil moisture quantity information from a different season to the chosen test period. That is, the test period of forecast–analysis cycles in this study is 1–28 February 2014, while we use ancillary data containing SST and soil moisture information for July 2011. With this design to incorporate model errors, we investigate the impact of using ESVs as additive inflation perturbation on the performance of EnKF data assimilation. In the third experiment set (RealData), we examine if the use of ESVs leads to a more accurate analysis in real situations.

In each experiment set, we carry out a group of three tests, with details summarized in Table 1. In the first test of each group, we use only a multiplicative inflation updated every grid point adaptively (Miyoshi, 2011) (MultInfl), in the second test we use the adaptive multiplicative inflation plus additive inflation from randomly chosen differences between two adjacent model forecasts (Random), and in the third one, we use multiplicative inflation plus additive inflation using ESVs. We name those tests according to the experiment set to which they belong (see Table 1). For the additive inflation in the “Random” test, we obtain the additive perturbations from the state differences between two adjacent 6-hourly forecasts that are randomly chosen from the pool that contains every 6 h output from the model run for 3 months. Then, we normalize the 30 sampled state differences in such a way that their mean becomes zero. This random sampling for additive inflation is similar to Whitaker and Hamill (2012). Before those differences are added to the analysis ensembles, they are multiplied by a scaling factor. The scalar factor is chosen to be 0.2, and the same value is used for a scaling when ESVs are used as additive inflation in this study.

4 | RESULTS

4.1 | OSSEs with a perfect model scenario (PerfModel)

We first evaluate the performance of the LETKF data assimilation system in OSSEs under perfect model assumption
TABLE 1  Three sets of experiments including two sets of OSSEs (PerfModel and ImPerfModel) and one of real data assimilation experiments (RealData): each set of experiments includes three experiments with different inflation methods (MultInfl, Random and ESVs).

| EXP sets | Inflation methods | [PerfModel] OSSEs with perfect model assumption | [ImPerfModel] OSSEs with imperfect model assumption | [RealData] Real data assimilation experiment |
|----------|-------------------|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|
| [MultInfl] | Multiplicative inflation only | PerfModel-MultInfl | ImPerfModel-MultInfl | RealData-MultInfl |
| [Random] | Multiplicative inflation + additive inflation with randomly sampled difference between adjacent model forecasts (6 h) | PerfModel-Random | ImPerfModel-Random | RealData-Random |
| [ESV] | Multiplicative inflation + additive inflation using growing ESVs with 6 h optimization period | PerfModel-ESV | ImPerfModel-ESV | RealData-ESV |

FIGURE 1  Time series of root mean square error (RMSE) for analysis temperature during the forecast–analysis cycles from the PerfModel. The RMSE of temperature analysis is averaged spatially in both horizontal and vertical direction and denoted in blue for the MultInfl, green for Random, and red for ESV.

TABLE 2  Time-averaged RMSE of temperature $T$ (K) and zonal wind $U$ (m/s) analysis for PerfModel, ImPerfModel, and RMSD for RealData. Numbers in () are standard deviation of RMSE over time for each variable

| Experiment set | Variables | MultInfl | Random | ESV |
|----------------|-----------|----------|--------|-----|
| PerfModel      | $T$       | 1.34 (0.27) | 1.29 (0.29) | 1.29 (0.29) |
|                | $U$       | 3.27 (0.69) | 3.11 (0.57) | 3.10 (0.58) |
| ImPerfModel    | $T$       | 1.36 (0.26) | 1.32 (0.28) | 1.28 (0.31) |
|                | $U$       | 3.27 (0.52) | 3.20 (0.54) | 3.09 (0.61) |
| RealData       | $T$       | 3.09 (0.23) | 3.10 (0.18) | 2.93 (0.17) |
|                | $U$       | 5.35 (0.31) | 5.55 (0.48) | 5.15 (0.28) |

(PerfModel). From the singular values, we find that the number of growing error modes is 16 for the first 6 cycles and 15 for the rest of the cycles. Figure 1 shows the time series of root mean square error (RMSE) of temperature analysis averaged spatially in both horizontal and vertical directions. At the beginning of the cycles, there are little differences between time series from those three tests. As the forecast–analysis cycles continue, RMSEs of both PerfModel-Random and PerfModel-ESV become smaller than that of PerfModel-MultInfl. This indicates that the use of additive inflation in addition to the multiplicative inflation leads to a further reduction of errors in analysis. However, the difference between PerfModel-ESV and PerfModel-Random is minimal, although RMSE of PerfModel-ESV is slightly smaller than that of PerfModel-Random during most of the forecast–analysis cycle. The time-averaged RMSE values of temperature ($T$), and zonal wind ($U$) analysis from these experiments are listed in Table 2.

4.2  OSSEs with an imperfect model (ImPerfModel)

In the experiment set ImPerfModel, we simulate some model deficiency by using the ancillary data of a different season from the test period, and enhance errors of background states (Figure 2). RMSEs of temperature analyses from both ImPerfModel-Random and ImPerfModel-ESV are smaller than those in ImPerfModel-MultInfl during most of the cycles, as in the first experiment set. However, the RMSE in ImPerfModel-ESV becomes significantly smaller than that in ImPerfModel-Random after 50 cycles (about 12 days). This result suggests that the positive impact from using ESVs as additive inflation on the analysis accuracy becomes greater in the presence of model errors. The time-averaged RMSE values of temperature and zonal wind during the whole forecast–analysis cycle are lowest when we use ESVs as additive inflation (see Table 2). When the differences between background state and the true state become larger due to the model deficiencies, ESVs may help to capture those growing errors effectively, and this leads to a better representation of background uncertainties. In reality, models have errors which cannot be sufficiently represented with a limited number of ensemble members. Therefore, the result from our experiments with an imperfect model scenario is promising and encouraging in that ESVs can add perturbations to ensemble subspace by capturing fast-growing errors to better deal with such sampling errors.

We examined the kinetic energy spectrum analysis of errors, and it confirms that errors in spatial scales of interest for data assimilation can be effectively handled by the use of ESVs. Figure 3 shows the kinetic energy spectrum of time-mean analysis errors at wave numbers between 1 and
Errors are defined to be the temporally averaged RMSE over the last 14 days, between 0000 UTC 15 and 1800 UTC 28 February. Here, we show the vertically averaged values of the kinetic energy spectrum of time-mean analysis errors. The range of the wave numbers from 6 to 20 corresponds to synoptic-scale flows of 2,000~6,000 km scale, and corrections by data assimilation are mainly characterized by the structure of those systems. Figure 3a shows that the error is smallest in PerfModel-ESV in most scales of interest, but the difference of energy distribution between PerfModel-Random and PerfModel-ESV is rather minor, consistent with the relatively small difference between time series of RMSE in those experiments (Figure 1).

With the imperfect model scenario in ImPerfModel, the energy spectrum of error is clearly smallest when ESVs are used as additive inflation, in comparison with the errors in ImPerfModel-Random and ImPerfModel-MultInfl (Figure 3b). This significant error reduction is consistent with the lowest RMSE in ImPerfModel-ESV, as shown in Figure 2. Whitaker and Hamill (2012) also suggest that a part of model errors which does not depend on the observing network can be more effectively treated by additive inflation than multiplicative inflation. In this study, results from ImPerfModel consistently suggest that when we include systematic model errors, the use of ESVs as additive inflation improves the analysis accuracy further.

Now, we examine further the representation of ensemble covariance and accuracy of analysis in experiment set ImPerfModel, to investigate how ESVs can help a better estimation of background uncertainties and a removal of errors. We compare the horizontal distribution of time-averaged background RMSE, and ensemble spread multiplied by multiplicative inflation from the three tests ImPerfModel. The multiplication of inflation is attempted to show how the data assimilation system actually sees the background uncertainties. Figure 4 shows that temporally averaged RMSE and ensemble spread of temperature (T) multiplied by the inflation factor, at model level 40 (around 700 hPa), during the last 5 days of the analysis period from ImPerfModel-MultInfl (top), ImPerfModel-Random (middle) and ImPerfModel-ESV (bottom). The standard deviation of their differences over the last 5 days is also shown for each experiment. Also, we present globally averaged time-mean ratio of RMSE/spread on top of each map for information. The globally averaged values of the standard deviation are 0.756, 0.766 and 0.692, while their RMSE/spread ratios are 1.164, 0.768 and 1.048 in ImPerfModel-MultInfl, ImPerfModel-Random and ImPerfModel-ESV, respectively. These numbers indicate that the difference between RMSE and spread is smallest in ImPerfModel-ESV among the experiments so that the system performs best when ESVs are used for additive inflation. Areas of large RMSE are over the Indian Ocean, Indonesia, New Guinea, eastern Pacific Ocean near Central America, Madagascar, and the Antarctic Ocean next to Chile. The magnitude of RMSE from ImPerfModel-ESV is smaller than those from ImPerfModel-MultInfl and ImPerfModel-Random in most of those regions.

Ensemble spreads in general represent areas of large RMSE well, but are significantly smaller than RMSEs in tropical regions in all three tests. Among the three tests, the spread in ImPerfModel-Random is generally larger than that in the other two tests. Interestingly, ensemble spread is even larger...
FIGURE 4  RMSE (left), background ensemble spread multiplied by inflation factors (centre), and the standard deviation of their differences over the last 5 days of experiment period (right) of 700 hPa temperature from the experiment set ImPerfModel (units: K). The upper panels show results from MultInfl, middle panels from Random, and bottom panels from ESV. RMSE and spread multiplied by inflation factor are averaged values over the last 5 days of experiment period. For the column of the standard deviation of (RMSE−spread), its horizontal mean is given on the top middle of each figure and the averaged ratio of RMSE/spread is also given on the top right of each figure [Colour figure can be viewed at wileyonlinelibrary.com].
than RMSE in some broad areas in ImPerfModel-Random such as some areas over the Northern Pacific and most areas over Eurasia. Also, the globally averaged values of RMSE/spread are less than 1 in ImPerfModel-Random as shown in Figures 4 (0.768) and 5 (0.844). This shows that ImPerfModel-Random possibly gives unnecessary inflation over those areas. Those perturbations can lead to the accumulation of errors which cannot be fully removed by data assimilation, and subsequently results in the degradation of performance.

The background error reduction in ImPerfModel-ESV is greater than those in ImPerfModel-MultInfl and ImPerfModel-Random, while the values of ensemble spread are not significantly reduced in comparison to those in the other two tests. This positive impact of using ESVs on the performance of LETKF is also shown in the temporally averaged RMSEs and ensemble spread of zonal wind background states (Figure 5). The RMSE values of zonal wind in ImPerfModel-ESV are smaller than those in the other two tests in general, but the ensemble spreads are similar to those found in ImPerfModel-MultInfl. The overestimated ensemble spread in some areas over the Northern Pacific Ocean and over the Eurasian area in ImPerfModel-Random is also seen in the zonal wind background states. The globally averaged values of ratio (RMSE/spread) is closer to 1 in ImPerfModel-ESV (1.077) than in the other two tests MultInfl (1.225) and Random (0.844). Also, the standard deviation of their difference over time is also smaller in ESV (1.690) than in MultInfl (1.811) and Random (1.820). This indicates that the gap between the ensemble spread and RMSE has been maintained smaller in ImPerfModel-ESV.

A question that can be raised is whether the exaggerated ensemble spread in space in ImPerfModel-Random is directly related to the accumulation of noise due to inflation. We further examine whether such a problem can be avoided by using ESVs, and hence fast-growing errors can be selectively corrected by data assimilation. Figure 6 shows the spatial distribution of analysis increments and background errors of zonal wind in ImPerfModel-Random and ImPerfModel-ESV. The analysis increment is defined to be the difference between the background and analysis states (analysis − background), and the background error is defined as the difference between background and true state (background − truth) in OSSEs. Figure 6a shows a snapshot of the background errors and analysis increments at the model level 45 (around 850 hPa) over a selected region in the Northern Hemisphere at 0600 UTC 20 February from the test ImPerfModel-Random. Over the Northern Pacific Ocean, there is in general good agreement between the analysis increments and errors, but there are also large increments in areas of non-significant errors over North America and East Asia. Analysis increments have too much projection in “quiet areas” where it is not dynamically unstable. These unnecessary increments in areas of small errors can lead to obvious degradation in analysis accuracy overall. Since 6 h differences of atmospheric states sampled randomly from a pool of model forecasts are added to ensemble perturbations, we cannot guarantee that the resulting ensemble perturbations will capture the spatial and temporal position of growing errors, (the “errors of the day”). This means that random perturbations can be redundantly added even over the well-constrained area with plenty of observations. By contrast, the analysis increments and errors from ImPerfModel-ESV are in better agreement, and these suggest that background errors are better captured by ESVs (Figure 6b). In other words, ESVs let the analysis focus on correcting major errors effectively, and lead to more accurate analysis.

We now consider whether the adaptive multiplicative inflation and additive inflation are complementary in order to better represent background uncertainties. We have examined the fields of multiplicative inflation factors together with background error at around 850 hPa from all experiments of ImPerfModel (Figure 7). These fields show that errors have been removed effectively over land where observations are dense and a larger multiplicative inflation factor is estimated. Meanwhile, the use of additive inflation leads to a smaller multiplicative inflation factor over land. The magnitude of the multiplicative inflation factor is smallest when randomly sampled perturbations are used for additive inflation. Over the ocean a smaller magnitude of multiplicative inflation factor is estimated broadly where observations are sparse and it seems to be less changed by the use of additive inflation than over land. In contrast with the situation over land, additive inflation may play an important role in correcting errors in those areas, so that errors are removed more effectively in both ImPerfModel-Random and ESV, but they are best corrected in the ESV experiment. This comparison indicates that the use of additive inflation can affect the estimation of the adaptive multiplicative inflation significantly in regions of dense observations. This suggests that additive inflation either from randomly sampled perturbations or ESVs may complement the multiplicative inflation, and the added perturbations can partially fill the gap between observed minus background (O−B) and ensemble spread. Meanwhile, additive inflation removes errors over the ocean effectively where observations are sparse and the multiplicative inflation is relatively underestimated. Therefore, we believe that additive inflation is partially complementing multiplicative inflation with a role especially important where observations are sparse. Importantly, ESVs can capture fast-growing errors more selectively and avoid the accumulation of noise by adding perturbations only in areas of dynamical instability.

### 4.3 Assimilation of real data (RealData)

We now experiment with real data assimilation using NCEP PrepBufr data containing conventional observations such as sondes, surface pressure, and satwind observations. As in the OSSEs, we use 30 ensemble members, and the same initial ensemble members in the experiment RealData. The
FIGURE 5  Same as Figure 4, except of zonal wind (units: m/s) [Colour figure can be viewed at wileyonlinelibrary.com].
Figure 6 shows the time series of globally averaged root mean square differences (RMSDs) of zonal wind analysis from RealData-MultInfl, RealData-Random and RealData-ESV. Soon after the forecast–analysis cycles start, the RMSDs from the test RealData-Random show the largest values (Table 2) with the strongest fluctuations overall in the test RealData-Random. Meanwhile, the RMSD from the test RealData-ESV fluctuates with much smaller amplitude, and after 10 days of spin-up, it remains always lowest among RMSDs from the three tests. This result indicates that the use of ESVs as additive inflation perturbations can also improve the performance of LETKF in real data assimilation, especially when the observations are not enough to constrain the states (0600 and 1800 UTC). The time-averaged RMSD of temperature and zonal wind analysis is also lowest in the test RealData-ESV (Table 2).

Figure 9 shows the time series of ensemble spread of zonal wind from each test in the experiment set RealData. The ensemble spread decreases rapidly during a couple of forecast–analysis cycles after the cycles start. Then the value...
FIGURE 7  Background error of zonal wind $U$ (m/s; contours, solid: positive, dash: negative) and multiplicative inflation (shading) at the model level 45, around 850 hPa at 0600 UTC 20 February 2014 from the tests: (top) ImPerfModel-MulInfl, (middle) ImPerfModel-Random, and (bottom) ImPerfModel-ESV [Colour figure can be viewed at wileyonlinelibrary.com].
As for the representation of background uncertainties in space, we examine how well areas of significant analysis increments match the areas of background errors in a real data assimilation experiment, as discussed for ImPerfModel in section 4.2. The analysis increment here is defined to be the difference between the analysis and the background states (analysis – background) as in section 4.2, but error here defined to be the difference between background and ECMWF reanalysis data (background – reference). Figure 10 shows a snapshot of the spatial distribution of analysis increments (shading), and the background errors (contours) at the pressure-level 500 hPa over a selected region in the Northern Hemisphere at 1200 UTC 7 February 2014 from the test RealData-Random and RealData-ESV. First, RealData-Random shows that areas of strong increments are placed nearly everywhere, even in regions where no significant errors are present (Figure 10a). This resembles the situation discussed in ImPerfModel experiments. Also, in real data assimilation, unnecessary analysis increments add noise in those areas that can consequently induce a degradation of the performance overall. Meanwhile, in the test RealData-ESV, there is in general good agreement between the analysis increments and background errors (Figure 10b). It is clearly confirmed that fast-growing errors are selectively captured by ESVs and that this leads to more effective corrections in a real data assimilation system. This suggests that ESVs have strong potential to improve the performance of operational EnKF data assimilation in practice.

Lastly, we note that the Linux cluster that we use at KIAPS has the Central Processing Unit (CPU) from INTEL Xeon 2.9GHz RHEL 6.3. The computational time of LETKF is on average 14 min, and of 9 h model forecast is 45 min for 30 members, when we use 20 computing nodes (16 processors per node). The overhead due to the computation of ESVs is about 10% of the total computational time at the moment. Since the computational time is not significantly increased and it mainly depends on the number of ensemble members, it would be feasible to use ESVs for global data assimilation in operational mode.

5 | SUMMARY

We test the use of ESVs as additive inflation using the LETKF data assimilation system implemented in the global NWP model at KIAPS (Shin et al., 2016). The use of ESVs as an additive inflation was first suggested by Yang et al. (2015) using a quasi-geostrophic model. They showed a positive impact on the performance of the LETKF system for that idealized quasi-geostrophic model. In this study, the positive impact from using ESVs as additive inflation is investigated and confirmed for real data assimilation, in addition to OSSEs with a global NWP model. We have tested three sets of experiments: OSSEs with perfect model scenario (PerfModel), OSSEs with imperfect model scenario (ImPerfModel), and
real data assimilation experiments (RealData). In each set of experiments, we performed every 6 h forecast–analysis cycle-runs of the LETKF with multiplicative inflation only (MultInfl), with multiplicative inflation plus additive inflation using 6 h model state differences sampled randomly from the pool of long-run model outputs (Random), and with multiplicative inflation plus additive inflation using ESVs (ESV), respectively.

Under the perfect model scenario, the analysis errors in PerfModel-ESV are smallest after a spin-up of the cycles among the three test sets, but the difference between PerfModel-ESV and PerfModel-Random is minimal (Figure 1). The positive impact from using ESVs as additive inflation becomes significantly larger under the imperfect model scenario. In this experiment named as ImPerfModel, we mimic model deficiencies by using SST and soil moisture data from a different season. With the presence of model deficiencies, the analysis accuracy was significantly improved in the ImPerfModel-ESV compared to the other two tests (Figures 2 and 3). In the test ImPerfModel-Random, random additive inflation could help a better representation of background uncertainties than multiplicative inflation (ImPerfModel-MultInfl), but also could lead to an accumulation of unnecessary perturbations in areas where errors were not actually growing significantly. By contrast, errors associated with flow dynamics were better captured by ESVs, and this led to an addition of perturbations selectively in areas of dynamical instability. This has been demonstrated by much better correspondence between analysis increments and background errors in ImPerfModel-ESV than in ImPerfModel-Random (Figure 6). This means that the improved representation of background error covariance by ESVs could result in a more effective correction by data assimilation, and subsequently an analysis with a reduced error (Figure 3 and Table 2). These results from the second experiment set ImPerfModel suggest that the role of ESVs can become more important when systematic model errors are present, in contrast to the PerfModel experiment where only initial errors are assumed to exist. This encouraged us to perform real data assimilation (the third experiment set, RealData) for the further investigation of the impact of ESVs on the performance of LETKF.

For the quantitative evaluation of the performance in real data assimilation experiment RealData, ECMWF Interim Reanalysis was used as a reference. There was a significant RMSD reduction in the RealData-ESV compared to the other inflation settings RealData-MultInfl and RealData-Random. As in the ImPerfModel-ESV, we found good agreement between background error and analysis increments.

A clear advantage of using ESVs was also shown in the time series of global ensemble spread. When no additive inflation was used, ensemble spread dropped rapidly as forecast–analysis cycle repeated and then remained nearly unchanged during the later stages. In RealData-Random,
ensemble spread decreased at first, but remained at a quite large value. An interesting point is that those periodically large RMSD values can be related to unnecessary analysis increments on areas of non-growing errors which might have led to an accumulation of noise in RealData-Random. Meanwhile, the ensemble spread decreased during a spin-up of the cycles, but grew again continuously narrowing the gap between the RMSD and spread in RealData-ESV.

In the experiment set PerfModel, only initial errors of ensemble existed, since the forecast model was assumed to be perfect. Thus, the simulated observations might be enough to constrain the system independently of the additive inflation used. In the experiment set ImPerfModel and RealData, however, initial and model errors were present, and hence the positive impact of using ESVs became outstanding, since ESVs could help capture fast-growing error modes associated with flow dynamics more effectively. With the limited ensemble size, the analysis needs to have more efficient corrections than in the case under the perfect model assumption. Since the handling of underestimation of background error covariance is more important in real data assimilation, we think that the use of ESVs would be beneficial to the performance of ensemble data assimilation system with an operational purpose. We note that the performance of the LETKF system is better when both adaptive multiplicative and additive inflation based on ESVs are used than when only additive inflation is used. This indicates that multiplicative and additive inflation cannot fully overlap and thus they need to be complemented by each other. The adaptive multiplicative inflation effectively corrects errors in regions of dense observations, and additive inflation can help to better remove errors over the ocean where observations are sparse and the multiplicative inflation is relatively underestimated. Thus, additive inflation based on ESVs complements the multiplicative inflation especially in dynamically active areas where observations are sparse. It would be interesting to examine which variables or combinations of variables are more benefitted by one or another type of inflation, and this might be worth investigating further in the future.

The choice of a scaling factor for additive inflation perturbation may have been made through experimentation to produce the lowest errors in the literature (e.g. Hamill and Whitaker, 2005; Whitaker et al., 2008; Miyoshi et al., 2010). In this study, we use the scaling factor 0.2 chosen also through experimentation for stable cycle runs in both OSSE and real data assimilation setting, and this value is not much different from the scaling factor 0.33 in Whitaker et al. (2008) and Miyoshi et al. (2010). With that choice, the order of magnitude of analysis increments and additive inflation are similar to each other, but the magnitude of added perturbation becomes stronger when the scaling factor larger than 0.3 is used in our system. It would be interesting to investigate a method to determine a scaling factor that varies adaptively for the optimal performance of an LETKF system in the future.

We think that Bred Vector (BV) can also provide reasonable perturbations useful for additive inflation perturbations. However, the derivation of BV is based on the concept that dynamical instabilities in a complex system are characterized by different growth rate and the breeding parameters are chosen with a physical meaning to select the type of instability. We should also note that the structure of BV and ensemble perturbations are alike due to the procedure of generation (i.e. the breeding cycle is like a data assimilation cycle), BV derived with a breeding interval the same as the assimilation interval may result in similar structures as in ensemble perturbations. In comparison, ESVs are derived with a focus of dynamically growing errors. Therefore, it is expected that BV may be less effective than ESV. It could be interesting to compare the performance of LETKF using BV with that using ESV in the future. We think that ESV can be advanced methodology more useful for the additive inflation perturbations than BV. There is also a review of BV and its relationship to final Singular Vector in Section 1 of Yang et al. (2015).

The computation of ESVs is nearly cost-free, since we can simply use products of ensemble data assimilation. The small computational overhead allows the on-line computations of ESVs, and it is feasible to use them in operational mode. In future, we will further examine the impact of ESV additive perturbations in the LETKF system using more types of observations including radiance data from satellite observations, and a forecast model with a higher resolution. Recently, Ota and Hotta (2016) used ESVs for improving the dynamical balance of the perturbations from LETKF for ensemble predictions. We also plan to investigate the impact of ESV additive perturbations on the performance of 5 to 10 day forecasts by the global NWP model.

ACKNOWLEDGEMENTS

We appreciate two anonymous reviewers for comments that helped us improve this manuscript. This work has been carried out through the R&D project on the development of global numerical weather prediction systems of Korea Institute of Atmospheric Prediction Systems (KIAPS) funded by Korea Meteorological Administration (KMA). Besides, this study was partly supported by “Development of big data technology for disaster management (K-18-L03-C03)” project of Korea Institute of Science and Technology Information (KISTI).

ORCID

Seoleun Shin https://orcid.org/0000-0001-7677-9661
Eugenia Kalnay https://orcid.org/0000-0002-9984-9906

REFERENCES

Bishop, C.H. and Toth, Z. (1999) Ensemble transformation and adaptive observations. Journal of the Atmospheric Sciences, 56, 1748–1765.
Corazza, M., Kalnay, E., Patil, D.J., Yang, S.-C., Morss, R., Cai, M., Sznayogh, I., Hunt, B.R. and Yorke, J.A. (2003) Use of the breeding technique to
estimate the structure of the analysis “error of the day”. Nonlinear Processes in Geophysics, 10, 233–243.

Dennis, J.M., Edwards, J., Evans, K.J., Guba, O., Lauritzen, P.H., Mirin, A.A., St-Cyr, A., Taylor, M.A. and Worley, P.H. (2012) CAM-SE: a scalable spectral element dynamical core for the Community Atmosphere Model. International Journal of High Performance Computing Applications, 26, 74–89.

Enomoto, T., Yamane, S. and Ohfuchi, W. (2006) Simple sensitivity analysis using ensemble forecast. In Proceedings of Third Workshop on Mechanisms of Climate Variation and its Predictability, Kyoto University, Kyoto: Disaster Prevention Research Institute, pp. 40–43 (in Japanese).

Enomoto, T., Yamane, S. and Ohfuchi, W. (2015) Simple sensitivity analysis using ensemble forecasts. Journal of the Meteorological Society of Japan, 93, 199–213. https://doi.org/10.2151/jmsj.2015-011.

Environmental Modeling Center. (2003) The GFS atmospheric model. NCEP Office Note 442. Camp Springs, MD: Global Climate and Weather Modeling Branch, EMC.

Evans, K.J., Lauritzen, P.H., Misra, S.K., Neale, R.B., Taylor, M.A. and Tribbia, J.J. (2013) AMIP simulation with the CAM4 spectral element dynamical core. Journal of Climate, 26, 689–709.

Hamill, T.M. and Whitaker, J.S. (2005) Accounting for the error due to unresolved scales in ensemble data assimilation. Monthly Weather Review, 133, 3132–3147.

Houtekamer, P.L. and Zhang, F. (2016) Review of the ensemble Kalman filter for atmospheric data assimilation. Monthly Weather Review, 144, 4489–4532.

Houtekamer, P.L., Mitchell, H.L., Pellerin, G., Buehner, M., Charron, M., Spacec, L. and Hansen, B. (2005) Atmospheric data assimilation with an ensemble Kalman filter: results with real observations. Monthly Weather Review, 133, 604–620.

Hunt, B.R., Kostelich, E.J. and Szunyogh, I. (2007) Efficient data assimilation for spatial-temporal chaos: a local ensemble transform Kalman filter. Physica D: Nonlinear Phenomena, 230, 112–126.

Kawai, Y., Moteki, O., Kuwano-Yoshida, A., Enomoto, T., Manda, A. and Nakamura, H. (2017) Impact propagation of radiosonde data assimilation over the Kuroshio and Kuroshio extension: case study on the early summer (Baiu) in 2012. Journal of the Meteorological Society of Japan, 95, 71–90. https://doi.org/10.2151/jmsj.2017-004.

Kim, K.-H., Shin, P.-S., Shin, S. and Kim, J. (2018) A simple method to find a neighboring grid point on the cubed-sphere. Asia-Pacific Journal of Atmospheric Sciences, 54(S1), 413–419.

Kotsuki, S., Ota, Y. and Miyoshi, T. (2017) Adaptive covariance relaxation methods for ensemble data assimilation: experiments in the real atmosphere. Quarterly Journal of the Royal Meteorological Society, 143, 2001–2015.

Leutbecher, M. and Lang, S.K.T. (2014) On the reliability of ensemble variance in subspaces defined by singular vectors. Quarterly Journal of the Royal Meteorological Society, 140, 1453–1466.

Miyoshi, T. (2011) The Gaussian approach to adaptive covariance inflation and its implementation with the local ensemble transform Kalman filter. Monthly Weather Review, 139, 1519–1535.