Journal of Mechanics Engineering and Automation 5 (2015) 185-196
doi: 10.17265/2159-5275/2015.03.008

A Local Ensemble Transform Kalman Filter Data Assimilation System for the Global FSU Atmospheric Model

Rosangela Saher Cintra[1] and Steve Cocke[2]

1. Laboratory of Computation and Applied Mathematics, National Institute for Space Research, São José dos Campos 12227010, Brazil

2. Center of Ocean-Atmospheric Prediction Studies, Florida State University, Tallahassee, Florida, USA

Received: February 16, 2015 / Accepted: March 17, 2015 / Published: March 25, 2015.

Abstract: Data assimilation is the process by which measurements and model predictions are combined to obtain an accurate representation of the state of the modeled system. We implemented a data assimilation scheme called LETKF (local ensemble transform Kalman filter) with FSUGSM (Florida State University Global Spectral Model) and made an experiment to evaluate the initial condition generated to numerical weather prediction to FSUGSM model. The LETKF analysis carries out independently at each grid point with the use of "local" observations. An ensemble of estimates in state space represents uncertainty. The FSUGSM is a multilevel (27 vertical levels) spectral primitive equation model, where the variables are expanded horizontally in a truncated series of spherical harmonic functions (at resolution T63) and a transform technique is applied to calculate the physical processes in real space. The assimilation cycle runs on the period 01/01/2001 to 31/01/2001 at (00, 06, 12 and 18 GMT) for each day. We examined the atmospheric fields during the period and the OMF (observation-minus-forecast) and the OMA (observation-minus-analysis) statistics to verify the analysis quality comparing with forecasts and observations. The analyses present stability and show suitable to initiate the weather predictions.

Key words: Data assimilation, Kalman filter, numerical weather prediction, global atmospheric model.

1. Introduction

Predictions from computer models of the atmosphere integrating the Navier-Stokes equations for a three dimensional multi-constituent multi-phase rotation fluid, and coupled to representation of the ocean and land surface, are continually put to the test through the daily weather forecast [1]. However, the predictability is determined by the projection of uncertainties in both initial conditions and model formulation onto flow-dependent instabilities of the chaotic weather and climate attractor [2]. Uncertainty is a characteristic of the atmosphere, coupled with inevitable inadequacies in observations and computer models and increases errors in weather forecasts, seasonal climate and hydrological predictions.

Model forecasts have the limited predictability of the behavior of the atmosphere. It is because the atmosphere is an inherently chaotic fluid. The accuracy of weather forecast are influenced by the ability to represent computationally the full equation of that governs the climate, in addition to error in initial conditions. The NWP (numerical weather prediction) model is sensitive to the initial error in function of the initial state [3]. The initial condition is represented by an objective analysis of an atmospheric state. Some techniques are used to determine initial conditions for weather forecasts given a set of atmospheric and oceanic observations whose density is heterogeneous in both space and time. The analysis is useful in itself as a description of the physical system, and it can be used as an initial state for the

Corresponding author: Rosangela Saher Cintra, Ph.D., research fields: computational intelligence and numerical weather prediction. E-mail: rocintra@gmail.com.

further time evolution of the system [4].

Ensemble forecasting is a form of multiple NWP that the models are conducted using slight different initial conditions. Ensembles are used to capture forecasting uncertainties. The multiple simulations have multiple forecasts, which are often referred to as sensitive dependence on the initial conditions. Ensemble prediction systems provide the means to estimate the flow-dependent growth of uncertainty during a forecast.

Data assimilation is the process by which measurements and model predictions are combined to obtain an accurate representation of the state of the modeled system. A data assimilation scheme is recognized as essential in weather, climate analysis, and forecast activities. All data assimilation schemes require reasonable estimates of the initial condition to run AGCMs (atmospheric general circulation models) considering the errors of the model, the observations and the analysis. The KF (Kalman filter) [5] is one approach to estimate an appropriate analysis to atmospheric models. In this work, a data assimilation scheme, the LETKF (local ensemble transform Kalman filter) was implemented, where the uncertainty is represented not by a covariance matrix, but by an ensemble of estimates in state space. The ensemble is evolved in time through the full model, which eliminates any need for a linear hypothesis as to the temporal evolution.

The EnKF (ensemble Kalman filter) [6] and the particle filter [7] use a probability density function associated with the initial condition, characterizing the Bayesian approaches [8]. Doucet [9] proposed the local ensemble Kalman filter (LEKF) where the EnKF is restricted to small areas (local); followed by LETKF (local ensemble transform Kalman filter; [11]) which is a kind of EnSRF (serial ensemble square root filter; [10]), but the algorithm is designed to be particularly efficient in parallel computer architecture by taking an advantage of independent local analyses.

This paper presents the first results of evaluation os analysis from LETKF scheme implemented with the AGCM of FSU (Florida State University). This data assimilation scheme is part of the analysis research at Center for Atmospheric-Ocean Prediction Studies of FSU, USA and the data assimilation with artificial neural networks at Laboratory of Computational and Applied Mathematics in INPE, Brazil. This experiment was conducted with synthetic observations, simulating measurements from surface and upper-air, and provides the basis of future research on FSUGSM data assimilation.

The paper is organized as follows: Section 2 has a brief description of the LETKF data assimilation scheme related to software implemented; Section 3 describes briefly the global atmospheric model FSUGSM used; Section 4 presents the methodology of experiment, the observations and the results; finally, Section 5 provides the discussion and summary.

2. Local Ensemble Transform KF

DA (data assimilation) is the process of finding the model representation of the atmosphere, which is consistent with the observations. According Talagrand [19], the purpose of assimilation is reconstructed as accurately as possible of the atmospheric or oceanic flow, using all available appropriate information. The DA essentially consists of:

- Observation proper, which vary in nature, resolution and accuracy, and are distributed more or less regularly in space and time;
- Physical laws governing the evolution of the flow, available in practice in form of discretized, and necessarily approximate, numerical model.

A summarized history of the main data assimilation algorithms used in meteorology and oceanography, roughly classified according to their complexity (and cost) of implementation, and their applicability to real-time problems. The computational complexity involved in DA systems, has been presented in the literature [20]. An important problem in atmospheric data assimilation lies in the large number of degrees of freedom of NWP models. Very large numerical dimensions: 10^7-10^9 parameters to be estimated, 2.10^7 observations per 24 h period. The large number of degrees of freedom of covariance matrices involved can prohibit the implementation of the best assimilation method known that needs for the forecast to be ready in time. Many strategies can be adopted to fit the intensive computation with the operation period: the use of advanced computing, reduction of problem dimension to obtain a computer code feasible to run in real time; even with the use of parallel computing with thousands of processors. The algorithms are constantly updated and improved in its performance.

One common approach at present is EnKF where the uncertainty is represented, not by an error covariance matrix, but by an ensemble of point estimates in state space, which are meant to sample the conditional probability distribution for the state of the system. Ensemble is evolved in time through the full model, which eliminates any need for linear hypothesis as to the temporal evolution. The ensemble forecasts are used for evaluate the probability distribution. Based in ensemble forecasting, the probabilistic state space formulation and the requirement for updating information when observations are encountered, the Bayesian approach is used to get the "errors of the day" on the predictions. The Bayesian approach is a set of efficient and flexible Monte Carlo methods for solving the optimal filtering problem.

The KF, a sequential assimilation scheme, it is the best linear unbiased estimate of analysis, where the equations are obtained from an analytical solution from setting the gradient of the cost function to zero, considering recursive least square and the assumption of the Gaussian probability density functions (pdf). A brief description for KF algorithm could express below:

Analysis step: update the analysis covariance matrix

$$x_n^a = x_n^f + W_K\left[x_{n+1}^{obs} - H_{n+1}\left(x_{n+1}^f\right)\right] \quad (1)$$

$$W_{n+1} = P_{n+1}^f H_{n+1}^T\left(H_{n+1}P_{n+1}^f H_{n+1}^T + R_{n+1}\right)^{-1} \quad (2)$$

$$P_n^a = P_n^f - W_K H_n P_n^f \quad (3)$$

Forecast step:

$$x_{n+1}^f = M_n\left(x_n^a\right) \quad (4)$$

$$P_{n+1}^f = M_n P_n^a M_n^T + W_n^b \quad (5)$$

The analysis x_n^a in Eq. (1) updates the analysis covariance matrix for P_n^a in Eq. (3) at analysis step, by solving for K_n in Eq. (2), and we get the optimal weight (e.g. Kalman gain). The matrices M_n and $\boldsymbol{H}$ represent the dynamical system and observation operator, respectively. The covariance matrix $\boldsymbol{R}$ identifies the observation error. The covariance matrix P_{n+1}^f in Eq. (5) is associated to forecast model x_{n+1}^f is updated in Eq. (4) at forecast step and W_n^b is the modelling error.

On EnKF approach, an ensemble of estimates on state space represents the covariance matrix P_{n+1}^f in Eq. (5). The nonlinear evolution problem for the error covariance is calculated by sampling from the probability density function and propagating samples, the model states, forward in time with the fully nonlinear model equations. At any time, the samples can be used to calculate an approximate mean and error covariance. The best implementation is in localization, in which applying a cut-off radius of influence for each observation eliminates spurious correlations. This is the LEnKF (local ensemble Kalman filter) algorithm that captures the space of forecast uncertainties, formulated by ensemble-based Kalman filter scheme.

The LETKF algorithm is an EnKF-based scheme, in which the analysis ensemble members are constructed by a linear combination of the forecast ensemble member [21]. The ensemble transform matrix, composed of the weights of the linear combination, is computed for each local subset of the state vector independently. The local subset depends on the error covariance localization [15] with limited ensemble size [22]. By "local", we mean that the analysis can be carried out independently at each grid point with the use of only local observations. The ensemble transform matrix, composed of the weights of the linear combination, is computed for each local subset of the state vector independently, which allows parallel computations. The experiment LETKF-FSUGSM, a local subset of the state vector contains all variables at the region centered at given grid point.

Each member of the ensemble gets its forecast: $\{x_{n-1}^f\}^{(i)} : i = 1,2,3,...,k$, where k is the total member at time t_n, to estimate the state vector $\overline{x^f}$ of the reference model. The ensemble is used to calculate the forecasting by the average:

$$\overline{x^f} \equiv k^{-1} \sum_{i=1}^{k} \{x^f\}^{(i)} \tag{6}$$

and the model error matrix is given by:

$$P^f = (k-1)^{-1} \sum_{i=1}^{k} (\{x^f\}^{(i)} - \overline{x^f})(\{x^f\}^{(i)} - \overline{x^f})^T \tag{7}$$

The LETKF determines an analysis $\{x_{n-1}^a\}^{(i)} : i = 1,2,3,...,k$ to each member of ensemble and an appropriate sample mean state estimate $\overline{x^a}$ in Eq. (8) and covariance P^a by Eq. (9).

$$\overline{x^a} \equiv k^{-1} \sum_{i=1}^{k} \{x^a\}^{(i)} \tag{8}$$

$$P^a = (k-1)^{-1} \sum_{i=1}^{k} (\{x^a\}^{(i)} - \overline{x^a})(\{x^a\}^{(i)} - \overline{x^a})^T \tag{9}$$

Data assimilation problems are often limited by the high dimensionality of states created by special discretization over large high-resolution grids and the extensive spatial structure of observations. LETKF are suitable for such problems, promising computational efficiency and accuracy in localization method.

The code of the LETKF in this experiment is based on the system initially developed by Miyoshi [21] and has been continuously improve. The current version is a MPI (message passing interface)-parallelized Fortran90 code and includes spatial covariance localization with physical distance [15] four-dimensional EnKF (4D-EnKF) for appropriate treatment of asynchronous observations [11] and temporal covariance localization. The LETKF code has been applied to and assessed with the Lorenz 40-dimensional model [23] a low-dimensional AGCM known as the SPEEDY model [15, 16, 24], realistic atmospheric models such as the AGCM for the Earth Simulator (Miyoshi and Yamane, 2007) and the Japan Meteorological Agency operational global and mesoscale models [13, 15, 25], a global ocean model know as the GFDL (Geophysical Fluid Dynamics Laboratory) Mars AGCM [26, 27] and Center for Weather Forecast and Climate Studies (CPTEC) AGCM [28]. All applications showed successful data assimilation using the LETKF code. The core part of the LETKF code is shared and improvements from applications can benefit other applications directly. This research with FSUGSM is an important application to use LETKF system; this model has some different characteristics to improve the LETKF system after this evaluation.

3. Florida State University Global Spectral Model

The FSUGSM is a general circulation model, it is a global spectral model based on primitive equations. The vertical coordinates are defined on sigma surfaces. The horizontal coordinates are latitude and longitude on a Gaussian grid in real space. The spectral model, used in this study, runs with T63 horizontal resolution (approximately 1.875°) and 27 unevenly spaced vertical levels in σ-coordinate. Details of this model can be found in Cocke and LaRow [29] and Krishnamurti [30].

The dynamical processes are the six primitive equations to forecast atmospheric motions: vorticity, divergence, thermodynamic, continuity, hydrostatic and moisture, which are expanded in their spectral form. The nonlinear terms are calculated on a Gaussian grid using a transform method. The vertical discretization of the FSUGSM, uses finite difference schemes and semi-implicit leapfrog scheme for time integration. The full physical packages include orography, planetary boundary layer, dry adjustment, large-scale precipitation, moist-convection, horizontal diffusion, and radiation processes. The horizontal diffusion term is usually incorporated in a numerical weather prediction model to parameterize the effects of motions on the unresolved scales and to inhibit spectral blocking, that is, the growth of small scales in the dynamic model variables due to the accumulation of energy at high wavenumbers. The presence of any dissipation, physical or computational, can attenuate the amplitude of the short wavelengths very significantly, cited by Zhuin and Navon [31].

According to Cocke [29], the global model has been developed to take advantage of scalable parallel architectures. The grid calculations are done use a domain decomposition approach. In this experiment, we use a small number of processor, and then the domain decomposition is simply a one-dimensional partitioning of latitude bands. Each latitude band maybe arbitrarily assigned to any processor to achieve optimal load balance. The vertical calculations for any given domain are done on the same processor.

4. Experimental LETKF-FSUGSM

Data assimilation is the process where observations are embedded into models, and to adjust them in real time as new data becomes available. The result of DA process is a consistent model with the observed data and itself forecasting, which is initial condition to next model prediction period, this run is called DA cycle. The LETKF-FSUGSM is tested with synthetic observations simulating surface and upper-air observations seeking the model grid point localization. LETKF analyses the prognostic variables: zonal component wind (u), meridional component wind (v), temperature (T), humidity (q), and surface pressure (ps), on period starting in 01/01/2001 until 31/01/2000. Observations are taken and analyzed every six hours (00, 06, 12, 18 UTC). The control or “nature run” model fields, to this assimilation experiment, are obtained from the integration of models without analysis, e.g., the initial condition for the next run is its previous model forecast. The first analysis, to run the control models, is taken from the National Centers for Environmental Prediction (NCEP). FSUGSM run with NCEP analysis from 31/12/2000-18UTC and the 6 hours forecast (01/01/2001-00 UTC) is the initial condition to run the control model to entire period of experiment.

4.1 Observations

The data assimilation experiments in this study are based on observational synthetic simulation experiments, where a control model fields assumed to be known, and observations are simulate by adding Gaussian random noise to that control or “nature run” model; this noise is according to observational errors.

The observational grid is a regularly distributed dense network; it has (45 × 96 × 27) grid points for latitude, longitude, and vertical directions, respectively for four upper-air variables (u, v, T and q) and (45 × 96) for surface variable (ps). This grid localization is one point at each two FSUGSM model grid of latitude/longitude with (96 × 192, 27). We exclude the extremes points of latitude, simulating no observations in poles. Large errors exist in polar region where no observations are available. Fig. 1 shows the observation grid to one level. Fig. 1 shows on example of the temperature observation grid to one level that is about 4,320 observations.

Firstly, we perform the model FSUGSM to collect the control or “nature run” fields and then we perform the observational routine to collect the synthetic observations based on model fields. Next step is to perform the analysis-forecast cycles. We run 236 analysis-forecast cycles. The first forecast to initiate the analysis cycle is the control field model from 01/01/2001 at 00 UTC (the result of the model performed with NCEP analysis for 31/12/2000-18UTC).

LETKF system is running with 20 members, the first 6 h forecast is the same for the members. The LETKF

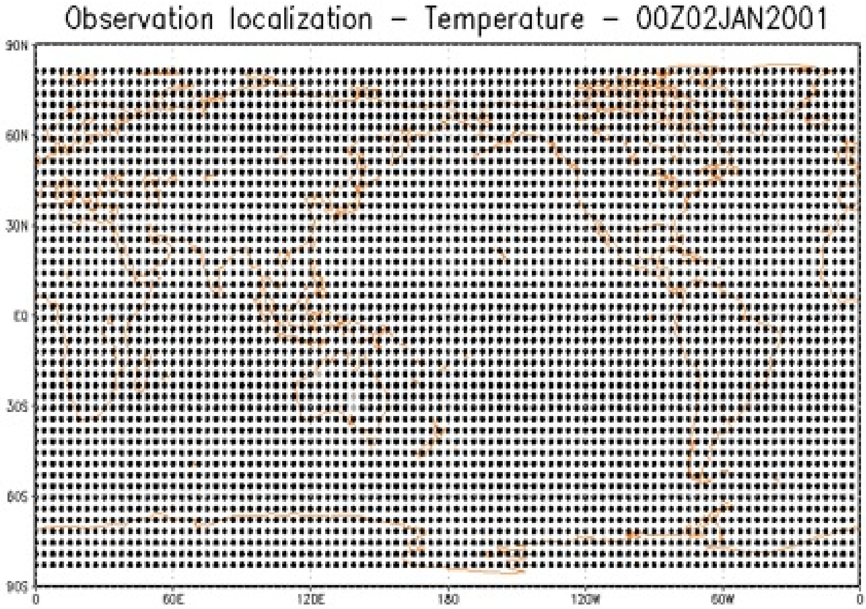

Fig. 1 The dot points shows the Temperature observation grid to 02/01/2001 at 00 UTC (each dot point is a localization observation).

program runs parallel with four processors. After LETKF perform, we obtain the error covariance matrix, the analysis for each member, the first-guess mean of the members and the analysis mean of the members. The LETKF program runs parallel with four processors. Then, we perform the FSUGSM model for each member, we run a single model simultaneously, submitting a job to one processor to each member at same time; each job has itself work area with itself analysis and its 6-hours forecast, these forecasts are the first-guess to the next assimilation cycle. Next assimilation cycle begins for the next time (6 h) as soon as the first-guess members and observations are ready. These tasks continue until 28/02/2001 at 18 UTC.

According to Laroz [32], the OMF (observation minus forecast) increment gives a raw estimation of agreement of the forecast information (i.e., the first guess) with the observation information prior to assimilation. Usually, a small OMF increment indicates a high quality forecast, and OMF increments are used as a primary measure of the quality of the assimilation. The OMA (observation minus analysis) increment represents the changes to the model forecast that are derived from the analysis algorithm. If the assimilation system weighs the observations heavily relative to the forecast, then the OMA increments will have significant differences relative to the OMF increments. If the model information is weighed more heavily than the observational information then there will be little change represented by the OMF increments.

The computer used to this experiment is the HPC (high performance computing) Cluster that provides 403 computer nodes and 6,464 CPU cores. Jobs are managed by scheduling software (e.g., batch processing). All programs are developed in Fortran90 codes and bash scripts are developed to implement the operation of data assimilation cycle. Although, the parallelization employed, the focus of this LETKF implementation is the evaluation of the analysis to FSUGSM, and then time results will not be commented.

4.2 Results

In this section we show the results of the LETKF-FSUGSM analysis-forecast cycles experiment. We compare the behavior of the model with analysis comparing with its first-guess, observations and control model fields.

Figs. 2 and 3 show that the model fields of LETKF analysis, first-guess (6 h forecast), control model and the differences between analysis and control model. Fig. 2 presents temperature in Celsius degree (C) on second vertical level (bottom to top), generated from assimilation cycle 08/01/2001-00UTC. Fig. 3 presents surface pressure in hector-Pascal (hPa) generated from assimilation cycle 13/01/2001-18UTC.

We noted that the analysis field is similar to forecast field, and the differences from control model are only some points over Australia and Asia (temperature), these differences for surface pressure are around 5 hPa to -5 hPa, and in temperature are about 4 °C or -6 °C.

These results show the effectiveness of data assimilation, the combination of forecast and some points of observation based on control model. The analysis follows the forecast bias.

Fig. 4 shows truncated fields over United States of America and Fig. 5 shows truncated fields over Brazil, these truncated fields are from LETKF analysis, control field (base of observations), and differences between those. Fig. 4 shows surface pressure generated in 15/01/2001 at 12UTC assimilation cycle.

Fig. 5 shows temperature at level 3 in 03/01/2001 at 06 UTC. The differences present between analysis field and control field of FSUGSM are small. Figs. 6 and 7 present the zonal global means from three fixed points at latitude: 30° Norte (30 N), EQ (equator) and 30° S (30 S) during the 124 assimilations cycles of January, 2001, the means are from control model, 6 h forecast and LETKF analysis.

Fig. 6 shows surface pressure: means between 960 and 976 hPa to 30 N; mean between 992 and 1,008 hPa to equator; means between 1,007 and 1,016 to 30 S. Fig. 7 shows surface temperature: means between 10 and 16 °C to 30 N; means between 27 and 29 °C to Equator; means between 22 and 24 °C to 30 S, considering the summer season to south hemisphere and winter season to north hemisphere, the results are coherent and the forecast and analysis are stable.

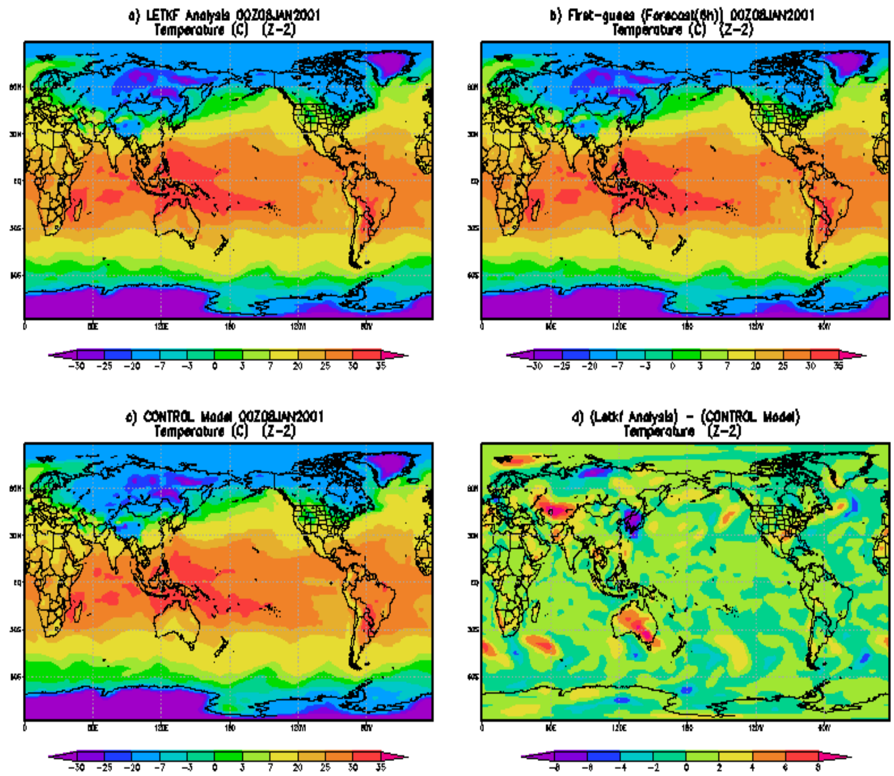

Fig. 2 Comparison of FSUGSM fields of temperature (°C) on sigma level 0.976 at 08/Jan/2001-00UTC, where (a) Letkf analysis field; (b) the forecast field; (c) control model field and (d) the differences field between analysis and control model.

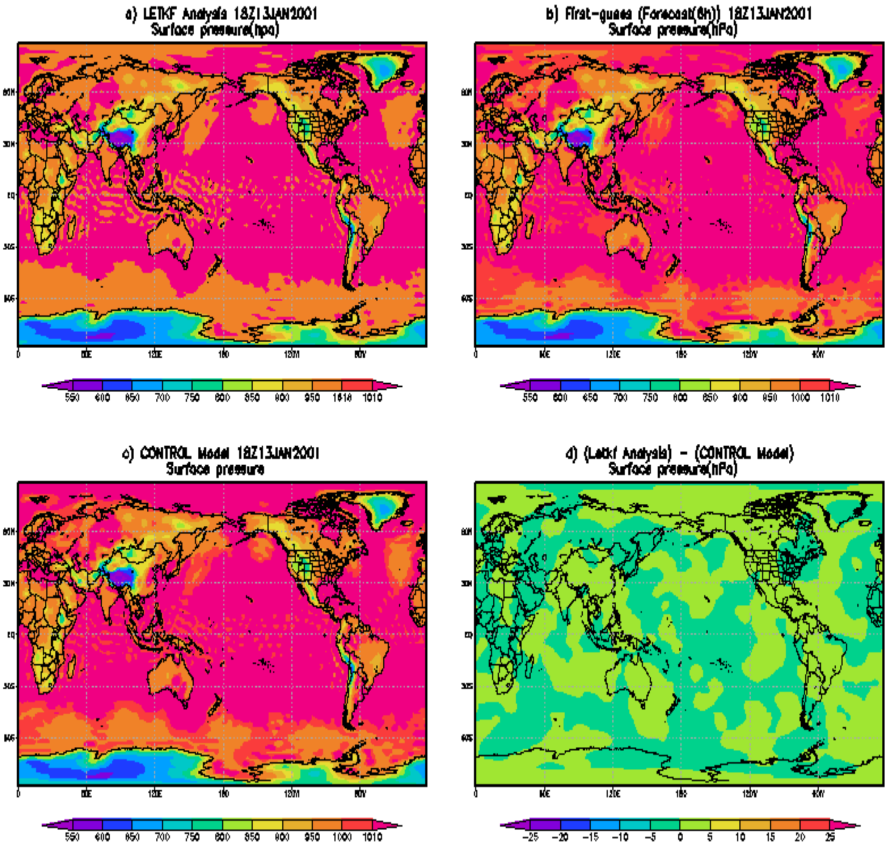

Fig. 3 Comparison of FSUGSM fields of surface pressure (hPa) at 13/Jan/2001-18UTC, where (a) Letkf analysis field; (b) the forecast field; (c) control model field and (d) the differences field between analysis and control model.

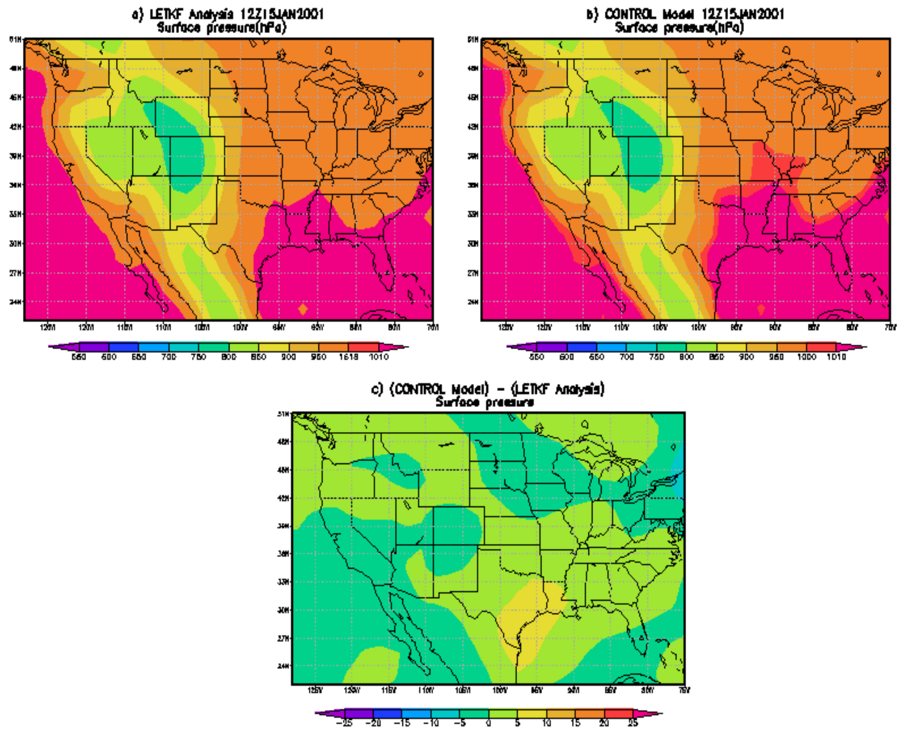

Fig. 4 Comparison of FSUGSM fields of surface pressure (hPa) at 13/Jan/2001-18UTC, over USA region, where (a) Letkf analysis field; (b) control model field and (c) the differences field between analysis and control model.

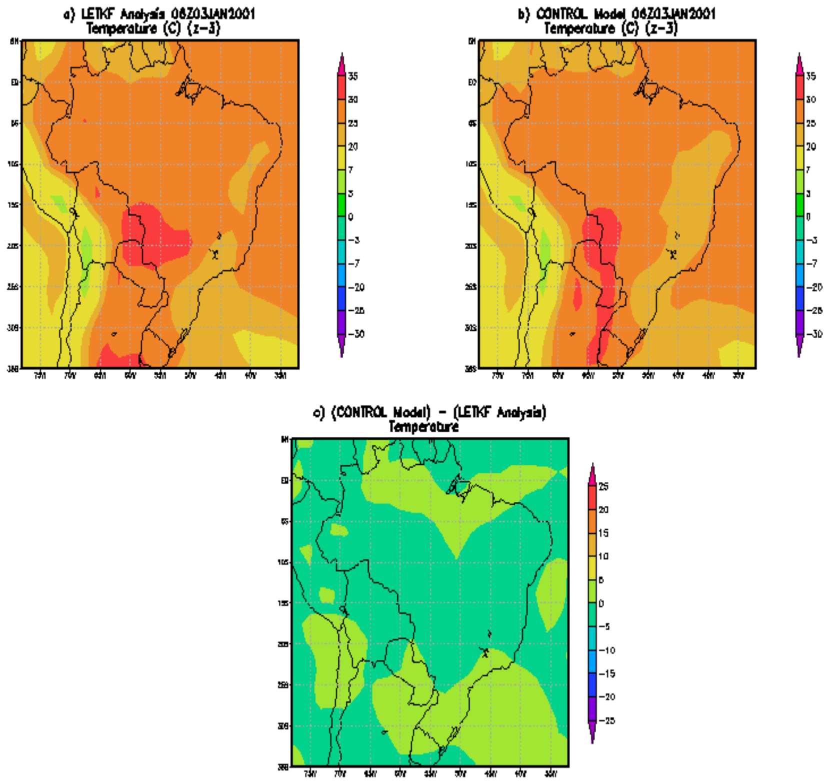

Fig. 5 Comparison of FSUGSM fields of temperature (°C) on 0.953 sigma level, at 13/Jan/2001-18UTC, over Brazil region, where (a) Letkf analysis field; (b) control model field and (c) the differences field between analysis and control model.

Fig. 8 presents the monthly average of humidity differences between control model and LETKF analysis during January 2001. The result shows that the analysis is consistent with control model, the major differences are between 0.05 and 0.2 Kg/Kg.

Since observations are constructed from the control model, OMF and OMA are expected to match each other. Furthermore, good performance means that the RMSE (root mean square error) in O-A is smaller than RMSE in O-F. The quality of the analysis is evaluated in terms of bias (observation minus analysis), and RMS (root mean square) of the bias, then these increments are shown at Fig. 9 and its RMSE at Fig. 10 where the OMA in red line is smaller than OMF in blue marks, and confirm that LETKF-FSUGSM works well.

5. Summaries and Discussions

We summarized the experiment of LETKF data assimilation on FSUGSM as follow:

(1) The LETKF data assimilation cycle experiment using a dense synthetic observational network, which are not in all points of model grid and the observation is located on the model point, the all observations are available to each 6 h and shows good performance. No observations are available in polar region;

(2) The ensemble size is chosen to be 20 members. A smaller ensemble size requires a smaller length scale and the data assimilation is more stable;

(3) The FSUGSM is an operational model to weather prediction with resolution T63L27 and also efficient computationally;

(4) The LETKF data assimilation perform analysis at each grid point simultaneously using the state variables and all observations in the region centered at given grid point, considering the dynamical nature, in which case a local subset contains only a part of the variables at a grid point. The ensemble is used to calculate the average of forecasting which is used to get the model error covariance matrix.

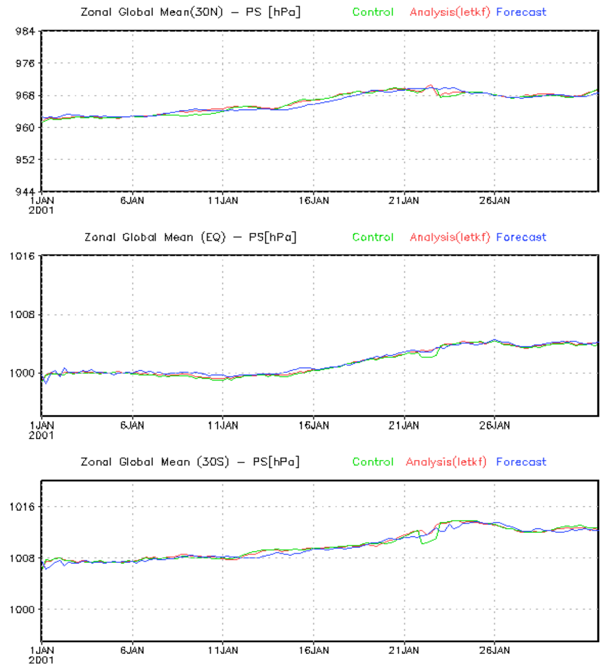

Fig. 6 Zonal global mean of surface pressure (hPa) of FSUGSM trajectory during January, 2001 at fixed latitude point: 30° Norte (30 S); EQ; 30° Sul (30 S). The blue lines are 6 h forecast means, the red lines are LETKF analysis and green lines are control model.

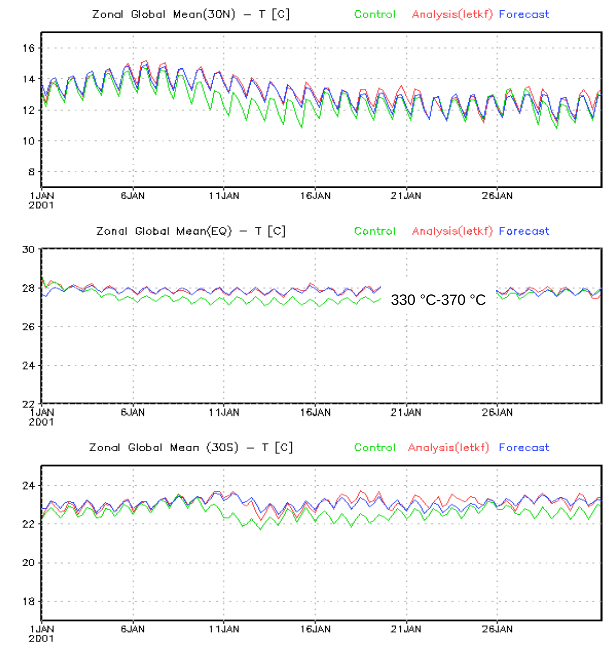

Fig. 7 Zonal global mean of surface temperature (°C) of FSUGSM trajectory during January, 2001 at fixed latitude point: 30° North (30 N); EQ; 30° South (30 S). The blue lines are 6 h forecast means, the red lines are LETKF analysis and green lines are control model.

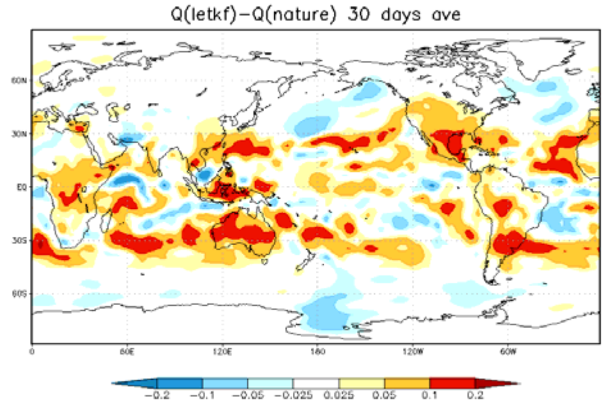

Fig. 8 Differences of zonal global mean of humidity (Kg/Kg) field of FSUGSM control model and LETKF-FSUGSM analysis during January 2001.

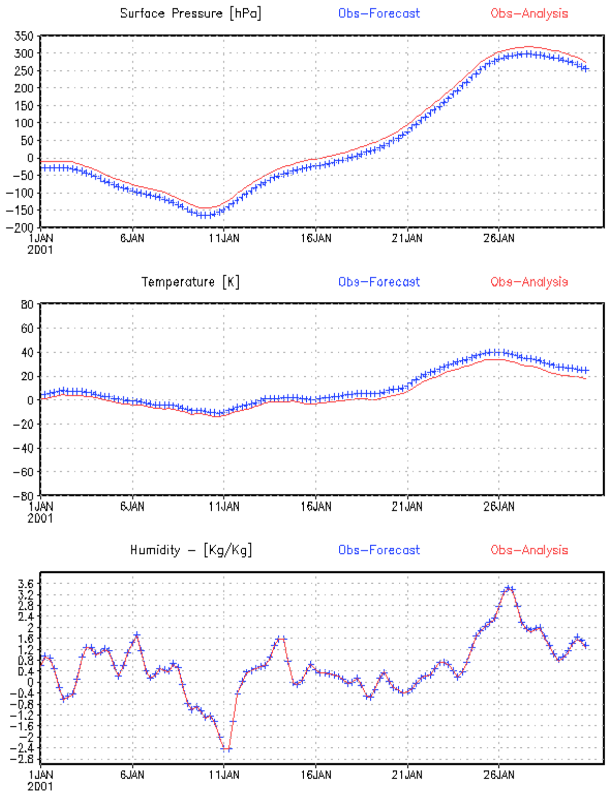

Fig. 9 OMF increment in blue marks comparing with OMA in red line to January 2001.

The results present by LETKF-FSUGSM data assimilation show stability to obtain analysis to initiate the FSU model and the analyses are suitable to get weather predictions.

The investigations continue, to get more results performing for a long period to evaluate errors and forecasts. Furthermore, we are implementing another

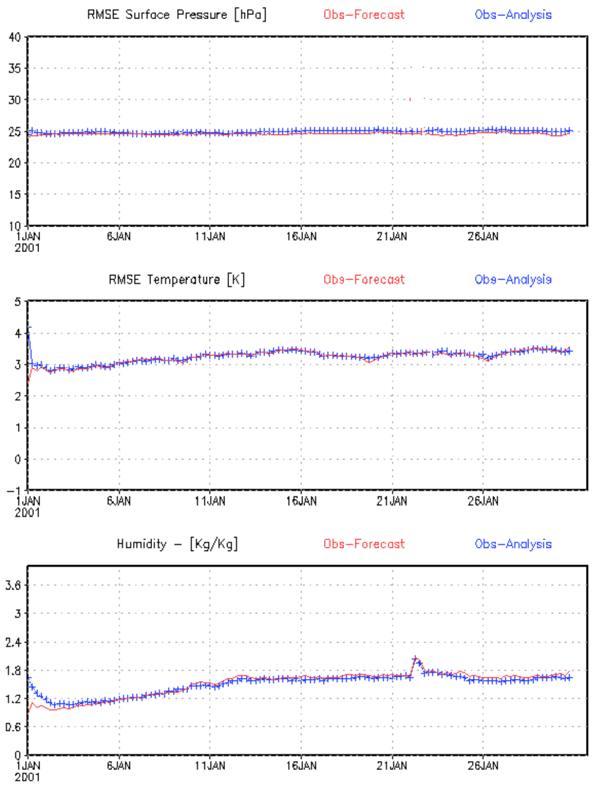

Fig. 10 Root Mean Square errors of OMF in blue marks and OMA in red line to January 2001.

strategy that use NCEP reanalysis to generate synthetic observations.

Acknowledgments

The authors thank Dr. Takemasa Miyoshi for providing computer routines of the LETKF code and Dr. Haroldo de Campos Velho for fruitful discussions. This paper is a contribution of the Brazilian National Institute of Science and Technology (INCT) for Climate Change funded by CNPq Grant Number 573797/2008-0 e FAPESP Grant Number 2008/57719-9.

References

[1] Bengtsson, L. 1999. **“**From Short Range Barotropic Modeling to Extended Range Global Weather Prediction: A 40-year Perspective.” *Tellus (A-B)* 51 (1): 13-32.

[2] Palmer, T. N. 1999. “A Nonlinear Dynamical Perspective on Climate Prediction.” *J. Climate* 12: 575-91.

[3] Lorenz, E. N. 1963. “Deterministic Nonperiodic Flow.” *J. Atmos. Sci.* 20 : 130-41.

[4] Hólm, E. V. 2008. *Lectures Notes on Assimilation Algorithms.* Technical report, European Centre for Medium-Rnage Weather Forecasts Reading, UK.

[5] Kalman, E. R. 1960. “A New Approach to Linear Filtering and Prediction Problems.” *Journal of Basic Engineering* 82: 35-45.

[6] Evensen, G. 1994. “Sequential Data Assimilation with a Nonlinear Quasi-Geostrophic Model Using Monte Carlo Methods to Forecast Error Statistics.” *Journal of Geophysics Research* 99 (10): 143-162.

[7] Doucet, A., Godsill, S., and Andrieu, C. 2000. “On Sequential Monte Carlo Sampling Methods for Bayesian Filtering.” *Statistics and Computing* 10 (3): 197-208.

[8] Daley, R. 19991. *Atmospherics Data Analysis.* New York: Cambridge University Press.

[9] Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D. J., and Yorke, J. A. 2004. “A Local Ensemble Kalman Filter for Atmospheric Data Assimilation.” *Tellus A* 56 (5): 415-28.

[10] Whitaker, J. S., and Hamill, T. M. 2000. “Ensemble Data Assimilation without Perturbed Observations.” *Mon. Wea. Rev.* 130: 1913-24.

[11] Hunt, B. R., Kostelichi, E. J., and Szunyogh, I. 2007. “Efficient Data Assimilation for Spatiotemporal Chaos: A Local Ensemble Transform Kalman Filter.” *Physica D* 230: 112-26.

[12] Szunyogh, I., Kostelich, E. J., Gyarmati, G., Patil, D. J., Hunt, B. R., Kalnay, E., Ott, E., and Yorke, J. A. 2005. “Assesing a Local Ensemble Kalman Filter: Perfect Model Experiments with the National Centers for Environmental Prediction Global Model.” *Tellus A* 57: 528-45.

[13] Miyoshi, T., and Aranami, K. 2006. “Applying a 4D-LETKF (Four-Dimensional Local Ensemble Transform Kalman Filter) to the JMA NHM (nonhydrostatic model). *SOLA* 2: 128-31.

[14] Miyoshi, T., and Yamane, S. 2007. “Local Ensemble Transform Kalman Filtering with an AGCM at a T159/L48 Resolution.” *Mon. Wea. Rev*. 135: 3841-61.

[15] Miyoshi, T., Sato, Y., and Kadowaki, T. 2010. “Ensemble Kalman Filter and 4D-Var Inter-Comparison with the Japanese Operational Global Analysis and Prediction System.” *Mon. Wea. Rev*. 138: 2846-66.

[16] Cintra, R. S. 2010. “Assimilacao de dados com redes neurais artificiais em modelo de circulacao geral da atmosfera.” Ph.D. thesis, National Institute for Space Research.

[17] Cintra, R. S., and Velho, H. F. C. 2014. “Data Assimilation by Artificial Neural Networks for an Atmospheric General Circulation Model: Conventional Observation.” *arXiv preprint* arXiv: 1407.4360.

[18] Hoffman, M. J., Greybush, S. J., Wilson, R. J., Gyarmati, G., Hoffman, R. N., Kalnay, E., Ide, K., Kostelich, E. J., Miyoshi, T., and Szunyogh, I. 2010. "An Ensemble Kalman Filter Data Assimilation System for the Martian Atmosphere: Implementation and Simulation Experiments." *Icarus* 209: 470-81.

[19] Talagrand, O. 2008. *Data Assimilation in Meteorology and Oceanography.* San Diego: Academic Press.

[20] Wilson, R. J., and Hamilton, K. P. 2008. "Comprehensive Model Simulation of Thermal Tides in the Martian Atmosphere." *J. Atmos. Sci.* 53: 1290-326.

[21] Lyster, P. M., Guo, J., Clune, T., and Larson, J. W. 2004. "The Computational Complexity and Parallel Scalability of Atmospheric Data Assimilation Algorithms." *J. Atmos. Oceanic Technol.* 21: 1689-700.

[22] Miyoshi, T. 2005. "Ensemble Kalman Filter Experiments with a Primitive-Equation Global Model." Ph.D. thesis, University of Maryland.

[23] Miyoshi, T. 2011. "The Gaussian Approach to Adaptive Covariance Inflation and Its Implementation with the Local Ensemble Transform Kalman Filter." *Mon. Wea. Rev*. 139: 1519-35.

[24] Lorenz, E. 1996. "The Evolution of Dynamic Meteorology." In *Historical Essays on Meteorology 1919-1995,* edited by Fleming, J. R. Boston: Amer. Meteor. Society.

[25] Molteni, F. 2003. "Atmospheric Simulations Using a GCM with Simplified Physical Parameterizations. I: Model Climatology and Variability in Multi-decadal Experiments." *Climate Dynamics* 20: 175-91.

[26] Miyoshi, T., and Sato, Y. 2007. "Assimilating Satellite Radiances with a LETKF (Local Ensemble Transform Kalman Filter) Applied to the JMA Global Model (GSM)." *SOLA* 3: 37-40.

[27] Wilson, R. J., and Hamilton, K. 1996. "Comprehensive Model Simulations of Thermal Tides in the Martian Atmosphere." *J. Atmos. Sci.* 53: 1290-326.

[28] Greybush, S. 2011. "Mars Weather and Predictability: Modeling and Ensemble Data Assimilation of Spacecraft Observations." Ph.D. thesis, University of Maryland.

[29] Medeiros, M. S., Herdies, D. L., Aravequia, J. A., and Souza, S. S. 2010. "The Impact of Assimilation with the Inclusion of AMSU—A Radiances in 4D-LETKF/AGCM System." In *Proceedings of American Geophysical Union, Fall Meeting,* abstract #A13B-0197.

[30] Cocke, S., and LaRow, T. E. 2000. "Seasonal Predictions Using a Regional Spectral Model Embedded within a Coupled Ocean-Atmosphere Model." *Mon. Wea. Rev*. 128: 689-708.

[31] Krishnamurti, T. N., Kanamitsu, M., Ceselski, B., and Mathur, M. B. 1973. "Florida State University's Tropical Prediction Model." *Tellus* 25 (6): 523-35.

[32] Zhijin, I. Li, Navon, M., and Zhu, Y. 2000. "Performance of 4D-Var with Different Strategies for the Use of Adjoint Phisycs with FSU Global Spectral Model." *Monthly Weather Review* 128 (3): 668-88.

[33] Lahoz, W. A., Khattatov, B., and Ménard, R. 2009. *Data Assimilation: Making Sense of Observations,* edited by Laroz, W. A., Khattatov, B., and Ménard, R. NewYork: Springer.