Article ID: 1006-8775(2011) 01-0043-07

IMPROVED ANALYSES AND FORECASTS WITH AIRS TEMPERATURE RETRIEVALS USING THE LOCAL ENSEMBLE TRANSFORM KALMAN FILTER

LI Hong (李　泓)[1], LIU Jun-jie (柳俊杰)[2], Elana FERTIG (艾莲娜・费尔蒂歌)[3], Eugenia KALNAY (尤金尼亚・康奈)[4], Eric KOSTELICH (埃瑞克・考斯特里奇)[5], Istvan SZUNYOGH (伊斯万・苏纽)[6]

(1. Shanghai Typhoon Institute, and Laboratory of Typhoon Forecast Technique/CMA, Shanghai 200030 China; 2. University of California, Berkeley, CA, USA; 3. Johns Hopkins University, Baltimore, MD, USA; 4. University of Maryland, College Park, MD, USA; 5. Arizona State University, Tempe, AZ, USA; 6. Texas A&M University, College Station, TX, USA)

Abstract: In this paper we investigate the impact of the Atmospheric Infra-Red Sounder (AIRS) temperature retrievals on data assimilation and the resulting forecasts using the four-dimensional Local Ensemble Transform Kalman Filter (LETKF) data assimilation scheme and a reduced resolution version of the NCEP Global Forecast System (GFS). Our results indicate that the AIRS temperature retrievals have a significant and consistent positive impact in the Southern Hemispheric extratropics on both analyses and forecasts, which is found not only in the temperature field but also in other variables. In tropics and the Northern Hemispheric extratropics these impacts are smaller, but are still generally positive or neutral.

Key words: AIRS retrievals, data assimilation, LETKF, observation impact

CLC number: P412.27 **Document code:** A **doi:** 10.3969/j.issn.1006-8775.2011.01.006

1 INTRODUCTION

The Local Ensemble Transform Kalman Filter (LETKF) (Hunt et al.[1]) is an efficient data assimilation scheme of the square root ensemble Kalman filter family (Tippet et al.[2]). It is closely related to Ensemble Transform Kalman Filter (Bishop et al.[3]) and was developed from the Local Ensemble Kalman Filter (LEKF, see Ott et al.[4]), but with improved computational efficiency. Unlike previously proposed serial square-root schemes that assimilate the observations one by one (Anderson[5], Whitaker et al.[6]), both LEKF and LETKF assimilate all observations within a local region simultaneously. In this way, these schemes can utilize parallel computation and are efficient when assimilating satellite observations, the number of which can be much larger than the number of degrees of freedom in the model.

The LEKF has been implemented to assimilate simulated observations in the NCEP Global Forecast System (GFS) model (Szunyogh et al.[7]). Liu et al.[8] have applied the LETKF to the NASA fvGCM model. The results from both LEKF and LETKF are much better than those from 3D-Var in a perfect model scenario. Szunoygh et al.[9] further showed that the LETKF analysis has been shown to be superior to the NCEP SSI (operational 3D-Var of NCEP in 2004) analysis when assimilating real conventional (non-radiance) observations on the NCEP GFS model at T62L28 resolution. We expect that LETKF will show a similar advantage when assimilating real satellite data.

The Atmospheric Infra-Red Sounder (AIRS) was launched on EOS Aqua in 2002. Some positive impacts on global analysis and forecast have been found in 3D-Var when assimilating AIRS radiances (LeMarshall et al.[10]) and AIRS version 3 (v3) retrievals (Atlas[11]). Since the LETKF has advantage over 3D-Var mentioned above, it is important to assess the ability of the LETKF to assimilate AIRS retrievals.

In this study, we use LETKF assimilation (Szunyogh et al.[9]) of all the real operational non-radiance observations on the NCEP GFS as a control run. Then, we add v5 emulation AIRS temperature retrievals, a more advanced version than

Received 2010-03-14; **Revised** 2010-11-18; **Accepted** 2011-01-18
Foundation item: National Natural Science Foundation of China (40975067); 973 Program (2009CB421500); CMA Grant GYHY200806029; NASA grant NNX07AM97G in U.S.A.
Biography: LI Hong, Ph. D., primarily undertaking research on atmospheric data assimilation.
Corresponding author: LI Hong, e-mail: lih@mail.typhoon.gov.cn

that used by Atlas[11], for the AIRS run. The impact of AIRS temperature retrievals are examined by comparing the accuracy of the analyses and forecasts of AIRS run with that of control run, which did not assimilate AIRS data. As in Szunyogh et al.[9], we use the 4D-LETKF to assimilate the observations within an analysis time window at their observation time.

The paper is organized as follows: In section 2, we briefly describe the 3D-LETKF scheme and its extension, 4D-LETKF. In sections 3 and 4, AIRS retrievals and experimental setup are described. In sections 5 and 6, AIRS impacts on both analyses and forecasts are given and discussed. Section 7 provides a conclusion and discussion.

2 3D- AND 4D-LETKF

Hunt et al.[1] provide a detailed description of the LETKF and explain how it differs from the other formulations of ensemble-based Kalman filters. Here, we emphasize its main characteristics. The LETKF uses the observations to update only the ensemble mean (shown in Eq. (1)), while it updates the ensemble perturbations by transforming the forecast perturbations $\mathbf{X}^b$ through a transform matrix $[(K-1)\widetilde{\mathbf{P}}^a]^{1/2}$ (Eq. (2):

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{X}^b\widetilde{\mathbf{P}}^a(\mathbf{H}\mathbf{X}^b)^T\mathbf{R}^{-1}[\mathbf{y}^o - h(\bar{\mathbf{x}}^b)], \quad (1)$$

$$\mathbf{X}^a = \mathbf{X}^b[(K-1)\widetilde{\mathbf{P}}^a]^{1/2}, \quad (2)$$

here K is the total number of ensemble members, h is the nonlinear observation operator and $\mathbf{H}$ its linear matrix. $\mathbf{X}^a, \mathbf{X}^b$ are the analysis and forecast ensemble perturbations, respectively. $\mathbf{X}^b$ is updated every analysis time step, therefore the forecast error covariance $\mathbf{P}^b = \frac{1}{K-1}\mathbf{X}^b\mathbf{X}^{bT}$ is flow-dependent, unlike the constant forecast error covariance used in 3D-Var. $\widetilde{\mathbf{P}}^a$, the analysis error covariance in ensemble space, is given by

$$\widetilde{\mathbf{P}}^a = \left[(K-1)\mathbf{I} + (\mathbf{H}\mathbf{X}^b)^T\mathbf{R}^{-1}(\mathbf{H}\mathbf{X}^b)\right]^{-1}, \quad (3)$$

which has dimension K by K, much smaller than either the dimension of the model or the number of observations. Thus, the LETKF performs the matrix inverse in the space spanned by the forecast ensemble members, which greatly reduces the computational cost.

The mean analysis state generated by this 3D-LETKF is the linear combination of the background ensemble states which best fits the available observations at analysis time. 4D-LETKF modifies 3D-LETKF by seeking the linear combination of the ensemble trajectories that best fits the observations within the assimilation window $[n-\frac{T}{2}, n+\frac{T}{2}]$ centered at the current analysis time n. Here T is the assimilation window length. Specifically, 4D-LETKF solves the following analysis equations.

$$\bar{\mathbf{x}}_n^a = \bar{\mathbf{x}}_n^b + \mathbf{X}_n^b\widetilde{\mathbf{P}}^a\left(\sum_{l=n-\frac{T}{2}}^{n+\frac{T}{2}}(\mathbf{H}_l\mathbf{X}_l^b)^T\mathbf{R}_l^{-1}[\mathbf{y}_l^o - h_l(\bar{\mathbf{x}}_l^b)]\right), \quad (4)$$

$$\mathbf{X}_n^a = \mathbf{X}_n^b[(K-1)\widetilde{\mathbf{P}}^a]^{1/2}, \quad (5)$$

$$\widetilde{\mathbf{P}}^a = \left[(K-1)\mathbf{I} + \sum_{l=1}^{n}(\mathbf{H}_l\mathbf{X}_l^b)^T\mathbf{R}_l^{-1}(\mathbf{H}_l\mathbf{X}_l^b)\right]^{-1}, \quad (6)$$

where the subscript l refers to the corresponding model state within the assimilation window. In this way, like 4D-Var, 4D-LETKF can assimilate observations at their right observed time, which benefits assimilating satellite data. The details of the 4D-LETKF can be found in Hunt et al.[12] and Harlim and Hunt [13].

3 OBSERVATIONS

We assimilate two observational data sets. The first set used in the control run contains all operationally available data except radiances (including all conventional data, satellite cloud-tracked winds) in January 2004. This conventional data is the same as that used by Szunyogh et al.[9] The second set of observations (Fig. 1) is the first set together with the AIRS temperature retrievals in the same period, provided by NESDIS at every 3°×3° and 100 vertical levels. The AIRS retrieval algorithm is a “version 5 emulation” system, based on the AIRS operational version 4 (Susskind et al.[14], Susskind et al.[15]) but with some changes (Chris Barnet, personal communication). The quality control flag resembles v4 qual_temp_mid=0 flag, but is applied to the whole column.

4 EXPERIMENT SETUP

Using 4D-LETKF, we first assimilate the first set of observations in January 2004 with the same NECP GFS T62/L28 model as used in Szunyogh et al.[9] (control run). The observation error standard deviation is provided along with the observations by NCEP. 40 ensemble members are used. We update analysis every 6 hours at 0000 Coordinated Universal Time (UTC), 0600 UTC, 1200 UTC and 1800 UTC, using observations from a 6-hour window centered at the analysis time. The background ensemble trajectories are outputted every 3 hours to form the corresponding model states within the 6-hour analysis window. We verify the analyses and forecasts against the NCEP GFS T254/L64 (much higher resolution than the model we used here) operational analysis that used all available operational observations, including radiances from ATOVS, but no AIRS data. We then assimilate the second set of observations including the AIRS temperature retrievals for the same time period (AIRS run) and assess the AIRS impact by comparing the results from the AIRS run with those from the control run.

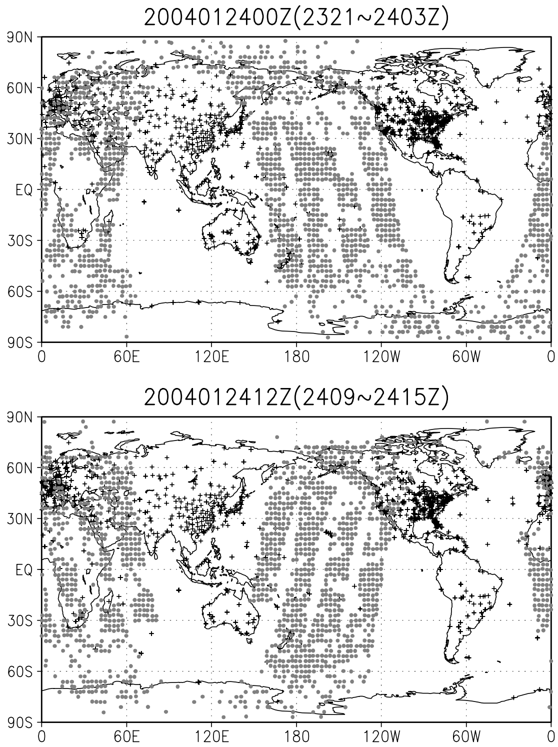

Fig. 1. Spatial distribution of the conventional temperature observations (plus sign) and AIRS temperature retrievals (closed circle) around 500 hPa in each window centered at analysis times on 24 January 2004

The observational errors are usually assumed to be independent with each other in the data assimilation process. However this assumption is not valid for retrievals due to the strong error correlations in the same vertical column that are the result of overlaps between the weighting functions of the different channels. Ideally, it is necessary to provide the error correlations when the retrievals are assimilated. However, these correlations are very difficult to determine (Joiner and Da Silva[16]). For simplicity, we ignore the error correlations, but increase the error standard deviations for AIRS temperature retrievals (provided by Chris Barnet, private communication) by a factor of 2 to compensate for the reduced magnitude of the observation error covariance matrix. The original error standard deviations at 100 vertical levels (Fig. 2) indicate that the AIRS retrievals have achieved an accuracy of 1 K between 300–900 hPa and 2 K near the surface and at upper levels.

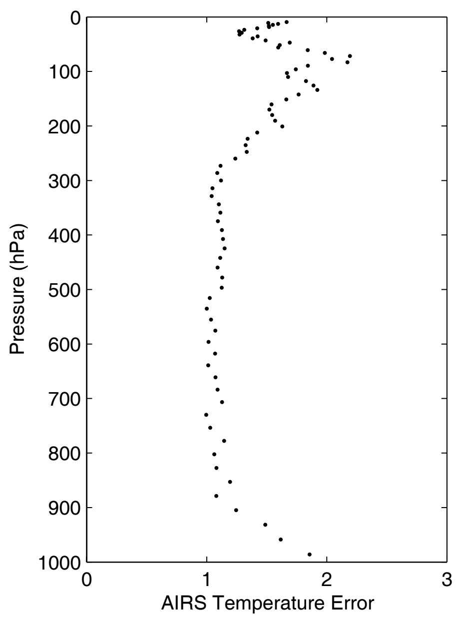

Fig. 2. Error standard deviations for AIRS temperature retrievals

5 ANALYSIS IMPACTS

Figure 3 shows the domain averaged 500-hPa temperature analysis RMS error verified against the

NCEP T254/L64 analysis for the control run (solid curve) and the AIRS run (dotted curve). There is a consistent reduction of errors when AIRS retrievals are assimilated in the global average point of view. It is remarkable that on days such as January 3rd, 13th and 27th, when the AIRS retrievals were missing, the AIRS run has almost the same RMS error as the control run, confirming the significant impact of AIRS retrievals on the analysis. As expected, the AIRS data improved the analysis more over the SH extratropics, by about 30%. The results for tropics and the Northern Hemisphere (NH) extratropics show a smaller but still consistent positive impact in the analysis.

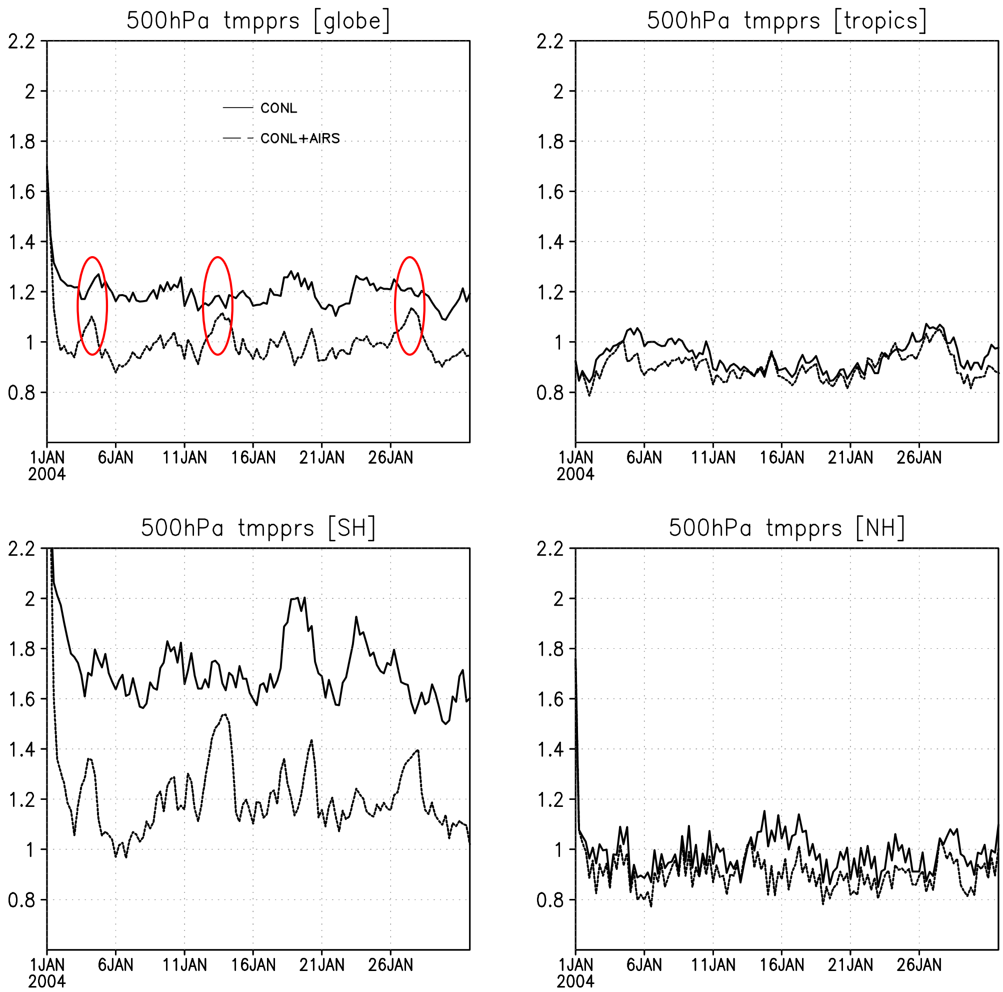

Fig. 3. Time series of 500-hPa temperature analysis RMS error in January 2004 for the control run (solid) and the AIRS run (dotted), averaged over globe, tropics (30°S–30°N), the Southern Hemisphere (SH) extratropics (90°S–30°S) and the NH extratropics (30°N–90°N). The ovals indicate days in which AIRS retrievals were missing.

Though only AIRS temperature retrievals were assimilated, the improvements are also found in other variables. The temperature information benefits other variables through the cross correlation between the observed temperature variable and other variables within the evolving dynamical forecast error covariance $\mathbf{P}^b$ and through the forecast process. This is seen in Fig. 4 with an improved AIRS analysis for the 500-hPa zonal wind field.

Figure 5 shows the AIRS impact at all vertical levels. In general, the analysis of the AIRS run is more accurate than that of control run for most of tropospheric levels, shown with blue color. However, a negative impact is observed in high levels (yellow color) which may be due to relatively big error of AIRS temperature retrievals at these levels (Fig. 2). The largest positive impact is found between 500 hPa and 700 hPa. Near the surface, the positive impact is degraded and turns to negative in some areas.

6 FORECAST IMPACTS

We have shown a significant beneficial impact of AIRS retrievals on the analysis, at least verified against the higher resolution operational NCEP analysis. We now test their impact on forecast skill. Global averaged 500-hPa temperature 48-hour forecasts RMS error are sought for the control run and the AIRS run (figure omitted); forecasts are started only from 0000 UTC and 1200 UTC since the conventional data are sparser at 0600 UTC and 1800 UTC in the control run. It is shown (figure omitted) that the AIRS run analyses lead to more accurate 48-hour forecasts than the control run analyses in all the domains. Though the improvement in tropics or NH extratropics is smaller than that in SH extratropics, we can see that the positive impact from the assimilation of AIRS temperature retrievals is consistent with time. This improvement is also seen at other levels and for other variables (Fig. 6). In general, the positive impact of AIRS retrievals is larger in the SH but less apparent in tropics and the NH.

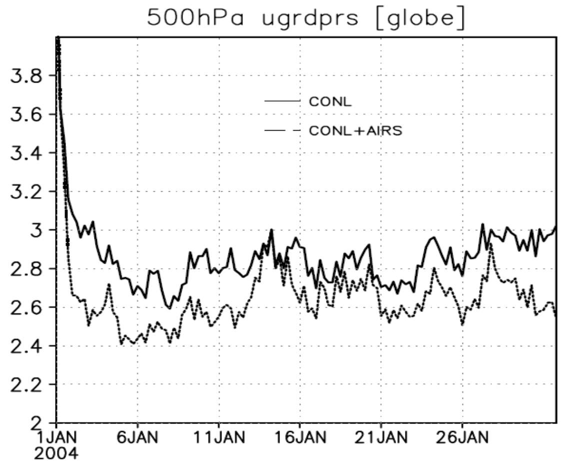

Fig. 4. Same as the top-left panel in Figure 3, except for the zonal wind field

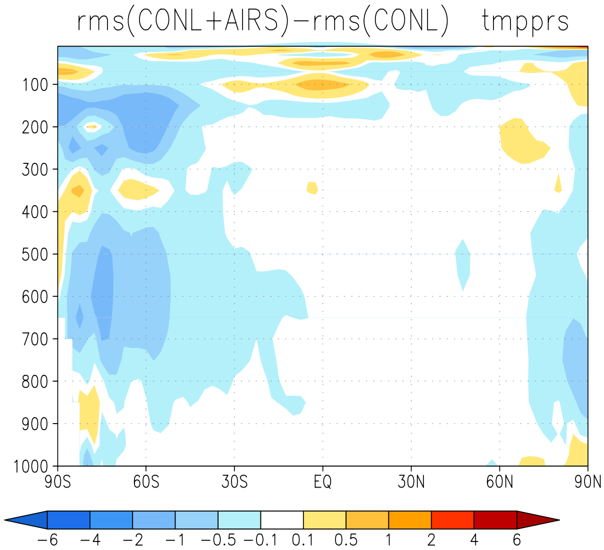

Fig. 5. Time-longitude averaged vertical cross-section of the RMS error of the AIRS run analysis minus the RMS error of the control run analysis for the temperature field, averaged over the last 10 days in January 2004. Blue colors indicate better agreement of the AIRS run analysis with the operational analysis.

7 CONCLUSION AND DISCUSSION

We investigated the AIRS retrievals impact by comparing the control run assimilating all the NCEP operational non-radiance data and the AIRS run added AIRS temperature retrievals. The AIRS temperature retrievals have a consistent positive impact on both analyses and forecasts, which is found not only in the temperature field but also in other variables. This positive impact is biggest in SH extratropics but still significant in tropics and the NH extratropics.

In our earlier tests, we did another experiment in which we arbitrarily assumed a constant 2-K error for AIRS temperature retrievals at all vertical levels in the LETKF assimilation. The result is worse than the current AIRS run, especially near the surface and at the upper levels. This is due to the different qualities of the AIRS temperature retrievals. As indicated in Fig. 2, the retrievals near the surface and in the upper levels are worse than those in the middle. Therefore, it is not a good assumption to use a constant observation error in the data assimilation. These indicate that there is a significant impact for having relatively accurate vertical error information. We also neglected the correlation of retrieval errors from different locations, which are difficult to estimate. We are exploring the possibility to estimate the observation error covariance of the retrievals based on the covariances of analysis and observational increments (Desroziers et al.[17]). This can be done online within the LETKF, as indicated by Li et al.[18].

It should be noticed that the control run in the current study does not include the assimilation of satellite radiance, resulting in a big impact of AIRS retrievals in the SH. In an operational assimilation system with satellite radiance, the impact of AIRS retrievals could be reduced, especially in the SH.

We plan to perform assimilation of both clear and cloud-cleared AIRS radiances and compare their impact with that of retrievals. Clear radiances are easier to assimilate because they are close to uncorrelated observational errors. The LETKF has the advantage that it does not require either the Jacobian or the adjoint of the radiative transfer model. Cloud-cleared radiances are much more abundant than clear radiances currently used in operations, but like the retrievals, they have correlated errors.

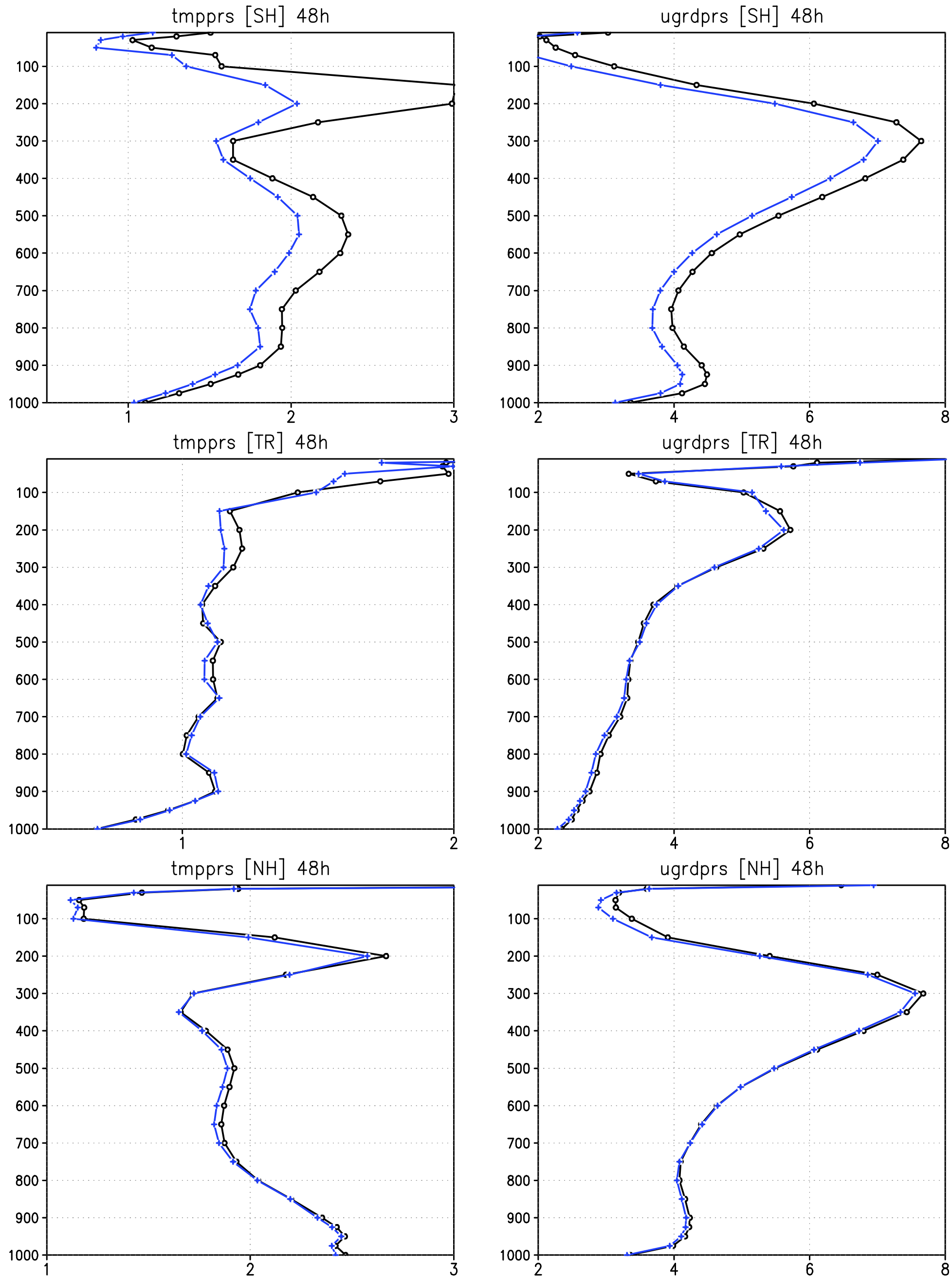

Fig. 6. RMS error of 48-hour temperature (left) and zonal wind (right) forecast starting from the control run analyses (black) and AIRS run analyses (blue), both verified against the operational NCEP analyses. The averages are taken over all grid points in SH extratropics (top), tropics (middle) and in NH extratropics (bottom), and over a period between 0000 UTC 15 January 2004 and 1200 UTC 31 January 2004.

Acknowledgments: The authors would like to thank Chris Barnet, Xingpin Liu and Lihang Zhou (NESDIS, NOAA) for providing the AIRS retrievals data and the codes to read the data, and the members of the AIRS science team for their generous guidance.

REFERENCES:

[1] HUNT B R, KOSTELICH E J, SZUNYOGH I. Efficient Data Assimilation for Spatiotemporal Chaos: a Local Ensemble Transform Kalman Filter [J]. Physica D, 2007, 230: 112-126.

[2] TIPPETT M K, ANDERSON J L, BISHOP C H, HAMILL T M, WHITAKER J S. Ensemble square root filters [J]. Mon. Wea. Rev., 2003, 131: 1 485-1 490.

[3] BISHOP, C H, ETHERTON B J, MAJUMDAR S J. Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects [J]. Mon. Wea. Rev., 2001, 129: 420-436.

[4] OTT E, HUNT B R, SZUNYOGH I, ZIMIN A V, et al. A local ensemble Kalman filter for atmospheric data assimilation [J]. Tellus, 2004, 56A: 415-428.

[5] ANDERSON, J L. An ensemble adjustment Kalman filter for data assimilation [J]. Mon. Wea. Rev., 2001, 129: 2 884-2 903.

[6] WHITAKER J S, COMPO G P, WEI X, et al. Reanalysis without radiosondes using ensemble data assimilation [J]. Mon.

Wea. Rev., 2004, 132: 1 190-1 200.
[7] SZUNYOGH I, KOSTELICH E J, GYARMATI G, et al. Assessing a Local Ensemble Kalman Filter: Perfect model experiments with the NCEP global model [J]. Tellus, 2005, 57A: 528-545.
[8] LIU J, FERTIG E, LI H, et al. Comparison between Local Ensemble Transform Kalman Filter and PSAS in NASA finite volume GCM-perfect model experiments [J]. Nonlinear Processes Geophy., 2008, 15: 645-659.
[9] SZUNYOGH I, KOSTELICH E J, GYARMATI G, et al. A local ensemble transform Kalman data assimilation system for the NCEP global model [J]. Tellus, 2008, 60A: 113-130.
[10] LE MARSHALL J, JUNG J, DERBER J, et al. Improving global analysis and forecasting with AIRS [J]. Bull. Amer. Meteor. Soc., 2006, 87: 891-894.
[11] ATLAS R. The impact of AIRS data on weather prediction [M]// Proc. SPIE Conf. Algorithms Technol. Multispectral Hyperspectral Ultraspectral Imagery XI, 2005, 5806: 599-606.
[12] HUNT B R, KALNAY E, KOSTELICH E J, et al. Four-dimensional ensemble Kalman filtering [J]. Tellus, 2004, 56A: 273-277.
[13] HARLIM J, HUNT B R. Four-dimensional local ensemble transform Kalman filter: numerical experiments with a global circulation model [J]. Tellus, 2007, 59A: 731-748.
[14] SUSSKIND J, BARNET C D, BLAISDELl J M. Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds [J]. IEEE Trans. Geosci. Remote Sens. 2003, 41: 390-409.
[15] SUSSKIND J, BARNET C D, BLAISDELL J M, et al. Accuracy of geophysical parameters derived from Atmospheric Infrared Sounder/Advanced Microwave Sounding Unit as a function of fractional cloud cover [J]. J. Geophys. Res. 2006, 111: doi:10.1029/2005JD006272.
[16] JOINER J, DA SILVA A M. Efficient methods to assimilate remotely sensed data based on information content [J]. Quart. J. Roy. Meteor. Soc., 1998, 124: 1 669-1 694.
[17] DESROZIERS G, BERRE L, CHAPNIK B, et al. Diagnosis of observation, background and analysis error statistics in observation space [J]. Quart. J. Roy. Meteor. Soc., 2005, 131: 3 385-3 396.
[18] LI H, KALNAY E, MIYOSHI T. Simultaneous estimation of covariance inflation and observation errors within ensemble Kalman filter [J]. Quart. J. Roy. Meteor. Soc., 2009, 135: 523-533.

Citation: LI Hong, LIU Jun-jie, Elana FERTIG et al. Improved analyses and forecasts with AIRS temperature retrievals using the Local Ensemble Transform Kalman Filter. *J. Trop. Meteor.*, 2011, 17(1): 43-49.