National Weather Association, *Electronic Journal of Operational Meteorology*, 2011-EJ05

# An Example of Vertical Resolution Impact on WRF-Var Analysis

Shih-Hung Chou

Earth Science Office, NASA Marshall Space Flight Center, Huntsville, Alabama

(Manuscript received 15 June 2010; in final form 13 May 2011)

ABSTRACT

Two different configurations of the Weather Research and Forecasting (WRF) model are used to examine the effect of vertical resolution on its 3D variational data assimilation system (WRF-Var). These configurations use the same horizontal spacing, physical parameterizations, input dataset for initial and boundary conditions, and observation data. The only difference is in their vertical resolution: the first set uses 37 vertical levels with 50 hPa spacing near the tropopause, while the second uses 50 vertical levels with 20 hPa spacing near the tropopause.

When initialing the WRF model with the North American Mesoscale (NAM) analysis, the low-resolution case exhibits a large vertical interpolation error near the tropopause compared to the high-resolution case due to the large gradient in lapse rate at the tropopause. The difference in interpolation error persists through the data assimilation time. When the profiles of Atmospheric InfrRed Sounder (AIRS) are assimilated, the larger upper-level interpolation error in the low-resolution case produces an analysis with warmer low-level temperatures and larger surface pressure increments compared to the high-resolution case. The difference is most pronounced over the eastern Gulf of Mexico where the largest upper-level interpolation errors occur. The surface pressure increase, together with the low-level temperature increase, contributes to a geopotential height bias throughout the troposphere. Increasing the number of model levels to 50 with a vertical spacing of 20 hPa near the tropopause reduces the spurious increases in surface pressure and geopotential height. The interpolation error and its impact on the WRF-Var analysis are presented in both a single test case and a 37-day time series.

## 1. Introduction

The terrain-following hydrostatic-pressure coordinate or its variants are commonly used in numerical weather prediction (NWP) models. A stretched vertical coordinate is typically used such that finer spacing is assigned to the lower atmosphere while coarser vertical spacing is applied at higher levels. The finer vertical grid is used to accommodate the rapid change of atmospheric variables in the boundary layer and to resolve the small-scale features near the Earth’s surface, while the coarser grid at higher levels is used to reduce computational cost.

*Corresponding author address: Dr. Shih-Hung Chou, Short-term Prediction Research and Transition (SPoRT) Center, NASA/MSFC, 320 Sparkman Drive, Huntsville, AL 35805.*
*E-Mail: shih.h.chou@nasa.gov*

Several studies have been conducted to investigate the impact of vertical resolution on NWP, focused mainly on interpolation error and the optimal resolution to produce a converged solution (Houghton et al. 1993; Manobianco et al. 1994; Williamson et al. 1998). However, very little work has been done to demonstrate the consequences of applying a coarser resolution to the upper troposphere of NWP models

An additional impact of model vertical resolution occurs when initializing a NWP model that will be used as the background (i.e. first-guess) field for an analysis system. If the vertical levels defined in the NWP model and the analysis system do not adequately capture the vertical variation of the initializing model, interpolation errors can occur and lead to improper assignment of initialization information. These interpolation errors can have far-reaching impacts including misrepresenting the first-guess field for a data assimilation system and producing large spurious increments in the analyzed fields used for model simulation. This paper presents a study where profiles from NASA's Atmospheric Infrared Sounder (AIRS; Aumann et al. 2003) are assimilated into the Weather Research and Forecasting (WRF) model using its three-dimension variational analysis component (WRF-Var; Barker et al. 2004). In this case, a coarse assignment of vertical levels near the tropopause results in larger than expected temperature innovations (observation minus background) which lead to spurious analysis increments in lower-tropospheric temperature, surface pressure, and tropospheric geopotential height. Increasing vertical resolution in the upper troposphere and lower stratosphere reduces the interpolation error and produces a more representative analysis. Section 2 outlines the AIRS data and experiment setup. In Section 3, the impact of vertical resolution on model initialization and assimilation analysis is evaluated. A summary and conclusions are presented in Section 4.

## 2. Experimental design

Version 2.2.1 of Advanced Research WRF (ARW, Skamarock et al. 2005) and its 3D variational data assimilation component, WRF-Var, are used in this study. The physical options used in this study are summarized in Chou et al. (2009). The model domain consists of a 450 x 360 horizontal grid with 12-km spacing that covers the contiguous United States, Western Atlantic Ocean, and Gulf of Mexico. For this numerical experiment, AIRS Level-2 Version 5 retrieved temperature and moisture profiles are assimilated. Because of its hyperspectral nature, AIRS can provide near rawinsonde-quality atmospheric temperature profiles (Aumann et al. 2003). More details on the retrieved AIRS profiles and their implementations in the WRF-Var can be found in Chou et al. (2009).

Two otherwise identical WRF configurations are run in parallel with different vertical resolutions: one uses 37 vertical levels (L37) and the other uses 50 vertical levels (L50). The model top-level pressure for both cases is set at 50 hPa. The vertical distributions of sigma levels in L37 and L50 are illustrated in Fig. 1, in which the corresponding pressure level is calculated assuming a surface pressure of 1000 hPa. While both cases have the same vertical resolution below $\sigma$=0.8, a coarser resolution is assigned above for L37—resulting in a vertical spacing of 50 hPa for L37 and 20 hPa for L50 near the tropopause. The vertical structure of the WRF-Var analysis grid for each resolution is identical to that of their respective WRF-ARW configurations.

The WRF-Var estimates the true state of the atmosphere by minimizing a cost function that combines a previous forecast (background), observations, and their respective errors (Barker et al. 2004). At each grid point, these errors define the weighting of the background and observations such that larger background error for a given variable will result in an analysis more closely resembling the observations (and vice versa). To reduce the computational requirements, the cost function is computed in control variable space instead of model variable space. The WRF-Var control variable transform is implemented through a series of operations. The multivariate transform converts the model variables to analysis control variables; for regional application, the control variables are the streamfunction, unbalanced velocity potential, unbalanced temperature, relative humidity, and unbalanced surface pressure (Barker et al. 2004). The balanced components are calculated from regression coefficients using the streamfunction as a predictor. The horizontal transform is performed using recursive filters, while the vertical transform is applied via an EOF decomposition of the vertical component of the background error on model levels. Unlike recursive filters or other successive correction techniques whose influence is local and limited to its neighboring grid points (depending on the length scale of the filter); the EOF eigenmodes describe the background error covariance of the whole atmospheric column. Therefore, the observation information can propagate vertically throughout the entire atmospheric column. For this study, a separate background error covariance matrix is created for L37 and L50 from a 37-day series of daily simulations (17 January to 22 February 2007). Hence, each background error represents the estimate of the model performances at different vertical resolution such that the analysis difference between L37 and L50 mainly comes from the difference in vertical level assignment, and not from the misapplication of the background error.

Retrieved profiles from AIRS over North America are valid at asynoptic times, so assimilation of these observations does not occur at a reported analysis or forecast time of large-scale models traditionally used for initialization. As a result, a short WRF forecast initialized at 0000 UTC is used as the background field for the WRF-Var analysis. This short WRF forecast uses cold-start initialization at 0000 UTC from the 40-km North American Mesoscale (NAM) model analysis with boundary conditions that are updated every 3 hours using NAM forecasts. The analysis time is chosen to fall between two successive AIRS swaths over central and eastern North America. Even though AIRS retrieved profiles are available at 100 vertical levels, a 28-level standard AIRS Level-2 retrieval product is used as a compromise between the superior vertical resolution of AIRS over other instruments (e.g. GOES Sounder, HIRS, etc.), while simultaneously minimizing the correlation between successive vertical levels and overlapping weighting functions. No other in situ observations, satellite radiances, or satellite profiles are assimilated in this experiment. The following section will first describe a single case study to illustrate the effects of interpolation error on the WRF-Var analysis when AIRS profiles are assimilated. To further demonstrate the presence of interpolation error in a coarse resolution, a 37-day time series will also be presented.

## 3. Results and discussion

As mentioned in the introduction, adequate vertical resolution is critical to the proper designation of information at each vertical model level. The results of interpolating the NAM analysis to the WRF model grids are illustrated in Fig. 2. It shows a comparison between temperature soundings from the NAM analysis and the WRF initial state for the L37 and L50 WRF grids at Key West, FL at 0000 UTC 17 January 2007. For reference, the 0000 UTC rawinsonde at Key West is also included. Because the Key West rawinsonde has already been assimilated in the NAM analysis, the temperatures of the NAM analysis should be quite close to those of the collocated rawinsonde, especially at the mandatory levels. The main differences between the NAM analysis and rawinsonde are a result of the former being plotted at every 25 hPa level and the latter being plotted on mandatory and significant levels. The WRF temperatures above 75 hPa are not plotted because mass variables (temperature, pressure, moisture, etc) are computed only on half-sigma levels in WRF-ARW. Sharp changes in the vertical gradient appear in the lower troposphere but are interpolated accurately due to the fine model resolution in the boundary layer. However, near the tropopause—where a large change in lapse rate occurs—the 50 hPa resolution in the L37 domain appears too coarse to resolve the vertical temperature variation from the NAM analysis resulting in an L37 initial state that is approximately 2$^{o}$C warmer than the NAM analysis at 100 hPa (Fig. 2a). This discrepancy occurs because the designation of half-sigma-levels in the WRF grid does not include the 100 hPa level, so the NAM temperatures at 125 and 75 hPa are interpolated linearly (in log pressure) to the 100 hPa level. Increasing the number of model levels to 50 with a vertical spacing of 20 hPa near the tropopause sufficiently reduces the temperature difference between WRF model and NAM analysis such that there is almost no difference between the NAM analysis and the interpolated WRF initial state at 100 hPa (Fig. 2b).

The 100 hPa warm bias of L37 persists during the short-term model forecast through the AIRS assimilation time at 0800 UTC. Figure 3 shows the sounding of the background field used for the WRF-Var assimilation (i.e. 8-h WRF forecast), the nearest AIRS profile at Key West, FL at 0800 UTC, and the resulting WRF-Var analysis (also valid at 0800 UTC). Although only one AIRS profile is shown in Fig.3, a number of nearby profiles are also assimilated. The AIRS profile does not reach the ground due to the presence of low cloud near 850 hPa. Note that the background temperature at 100 hPa of L37 is still 3-4$^{o}$C warmer than that of L50 after 8 hours of model integration. This results in an approximately -7$^{o}$C temperature innovation in L37, while only an approximately -3°C temperature innovation occurs in L50. With a warmer background temperature and, thus, a larger innovation at100 hPa, L37 produces an analysis which is not only warmer near tropopause, but also warmer in the surface layer compared to L50.

To further illustrate the effect of 100 hPa temperature innovation on the WRF-Var analysis, an experiment has been conducted to assimilate only the 100 hPa AIRS data in each vertical resolution. Figure 4a shows that the L37 analysis increment displays a dipolar structure with cooling above 270 hPa and warming below, indicating a strong negative correlation between the upper troposphere/lower stratosphere and the lower troposphere. This vertical correlation, exemplified in Fig. 3 of Lee and Barker (2005), is primarily due to the use of the EOF decomposition in WRF-Var's vertical transform, in which the observational information is spread vertically throughout the air column. With a 100 hPa innovation of -7°C, L37 can produce up to 4°C warming in the lower troposphere. With higher vertical resolution and smaller innovation, L50 shows a similar correlation, but with a much reduced increment. This inverse vertical correlation can also have physical meaning. The upper-tropospheric cooling due to the AIRS assimilation would displace the 100 hPa pressure level in data assimilation analysis to a higher altitude because of the use of hydrostatic balance in the WRF-Var. That means the air column in the analysis field will be stretched, which is attributed to the warming on the low level atmosphere in the analysis.

Large 100 hPa temperature innovations occur over a broad region in the Gulf of Mexico as seen in Fig. 5a, where L37 shows an area of -7°C temperature innovations over Cuba and southern Florida. In contrast, L50 shows a maximum negative innovation of only -3°C (Fig. 5b). Thus, the interpolation error introduced near the tropopause during the model initialization has a widespread impact on the low-level temperature analysis. The 850 hPa temperatures produced by assimilating only the 100 hPa AIRS are shown in Figs. 5c and 5d for L37 and L50, respectively. Note that L37 shows a large area of 850 hPa warming extended from the Gulf of Mexico to New England, while L50 shows very little warming. This single-layer assimilation experiment clearly shows that the artificial warming of interpolated background temperature near the tropopause can cause significant warming in the lower levels. The 100 hPa interpolation error also leads to a surface pressure increase as a result of the cost function minimization. The surface pressure increments are shown in Figs. 5e and 5f for L37 and L50, respectively. As expected, L37 shows a larger surface pressure increment than L50 (3 hPa vs. 1.5 hPa).

When the full AIRS profiles are assimilated, the low-level warming is still evident as shown in Fig. 6a and 6b. The surface pressure increase is also more pronounced in L37 than in L50 in the eastern Gulf of Mexico, where L37 shows a maximum surface-pressure increment of 5 hPa, while L50 exhibits only a 3 hPa maximum increment (Fig. 6c and 6d). Since geopotential height is an integrated quantity calculated upward from the surface, the combination of surface pressure increase and the low-level warming contributes to a systematic geopotential height increase throughout the troposphere in L37. An example of geopotential height increase in L37 is shown in Fig. 6e for the 850 hPa pressure level. Conversely, increasing the vertical resolution near the tropopause reduces the 100 hPa temperature innovation and produces an analysis with a cooler lower troposphere (Fig. 6b), smaller surface pressure increments (Fig. 6d), and smaller geopotential height increments (Fig. 6f). At 850 hPa, increasing the vertical resolution reduced the maximum height increment from 50 m to 30 m.

Although only one case study is presented here, several days from the winter of 2007 have also been examined. These experiments consistently show similar results: the L37 exhibits a warm bias at 100 hPa for the interpolated WRF initial state compared to the NAM analysis and leads to a warmer low-level temperature and high surface pressure resulting in higher geopotential height bias in the analysis. Meanwhile, the L50 reduces the 100 hPa interpolation error leading to improvements in surface pressure and geopotential height. Figure 7 shows the interpolation errors and surface pressure analysis errors during a 37-day test case period (January 17 to February 22 2007) time series at Key West, FL for the L37 and L50 cases. At model initialization, Fig. 7a shows that for all the case study days the 100 hPa temperatures in the L37 initial state is warmer than those in the L50 (averagely 3°C and 0.5°C, respectively, compared to the NAM analysis). This demonstrates that the problem is fairly systematic and the previous test case is neither an isolated case nor an extreme case. Since the AIRS assimilation occurs at asynoptic times, there are no upper-air observations for analysis validation. However, the hourly METAR observation can provide the surface pressure data for validation. Fig. 7b shows the surface pressure differences between the WRF-Var analysis and the METAR at Key West, FL. For almost all of the case study days, the L50 analysis consistently shows a smaller bias toward the METAR than the L37 analysis, which shows that the L50 configuration produces a better analysis. In this experiment, the AIRS profiles are used to illustrate the impact of vertical resolution on WRF-Var analysis; however, any observation assimilated near the tropopause would yield similar results in this model configuration. This is especially true in an assimilation system that uses EOF decomposition for the vertical transform.

## 4. Summary and concluding remarks

This study demonstrates that vertical resolution plays an important role in regional variational analysis systems when assimilating observations near the tropopause. Increasing the vertical resolution can improve the accuracy of regional numerical models by decreasing interpolation errors. For data assimilation systems that use EOF decomposition (instead of filter or successive correction techniques) for the vertical transform such as WRF-Var, the impact of interpolation error can affect the whole atmospheric column by producing a spurious analysis that is not consistent with observations. However, using finer vertical resolution near the tropopause reduces the interpolation error; thereby, producing a background field that produces a smaller spurious warming near the tropopause and an analysis that is more physically representative of the real atmosphere. The results presented in this paper can serve as a guideline for setting up the vertical grid of a regional forecast/assimilation system. Ideally, one would like to have as many vertical levels as possible. In reality, a practical compromise has to be made between the model accuracy, timeliness of forecasts, and computational cost. With limited computational resources, it would be prudent to maximize the model vertical resolution at the layers where lapse rate changes rapidly with height such as the boundary layer and tropopause. Increasing the number of vertical levels adds only a linear increase in compute time. However, increasing the horizontal resolution requires an exponential increase in compute time. Furthermore, increasing horizontal resolution cannot alleviate the interpolation error at 100 hPa. The height of the tropopause varies with the season, longitude and, especially, latitude. For forecasters interested in assimilating data in tropical or subtropical regions such as the present study, it would be beneficial to increase the resolution near 100 hPa. However, for Polar Regions, it may be more practical to use finer vertical resolution near 300 hPa.

*Acknowledgements.* The author would like to thank Tsengdar Lee of NASA Headquarter for providing the funding for this work. The author would also like to thank Jonathan Case, Bradley

Zavodsky, Gary Jedlovec, and Tim Miller for the discussions and comments during this research. Thanks also to the anonymous reviewers for their constructive comments that helped improve this paper.  Computational resources for this work were provided by the NASA Center for Computational Sciences at the Goddard Space Flight Center.

REFERENCES

Aumann, H. H., M. T. Chahine, C. Gautier, M. D. Goldberg, E. Kalnay, L. M. McMillin, H. Revercomb, P. W. Rosenkranz, W. L. Smith, D. H. Staelin, L. L. Strow, and J. Susskind, 2003: AIRS/AMSU/HSB on the Aqua Mission: Design, Science Objectives, Data Products, and Processing Systems. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, 253-264.

Barker, D. M., W. Huang, Y.-R. Guo, A. J. Bourgeois, and Q. N. Xiao, 2004: A three-dimensional variational data assimilation system for MM5: Implementation and initial results. *Mon. Wea. Rev.*, **132**, 897-914.

Chou, S.-H., B. Zavodsky, and G. J. Jedlovec, 2009: Data assimilation and regional forecasts using Atmospheric InfraRed Sounder (AIRS) profiles. Preprints, *16$^{th}$ Conf. on Satellite Meteorology & Oceanography Conference*, Phoenix, Arizona, Amer. Meteor. Soc., P.6.11. [Available online at: http://ams.confex.com/ams/pdfpapers/147745.pdf.]

Houghton, D. D., R. A. Petersen, and R. L. Wobus, 1993: Spatial resolution impacts on National Center Nested Grid Model simulations. *Mon. Wea. Rev.*, **121**, 1450-1466.

Lee, M.-S., and D. M. Barker, 2005: Background error statistics in 3DVAR derived from WRF ensembles. Preprints, 6th WRF/15th MM5 Users' Workshop, Boulder, Colorado, National Center for Atmospheric Research. [Available online at: http://www.mmm.ucar.edu/wrf/users/workshops/WS2005/abstracts/Session3/20-Lee.pdf.]

Manobianco, J., S. Koch, V. M. Karyampudi, and A. J. Negri, 1994: The impact of assimilating satellite-derived precipitation rates on numerical simulations of the ERICA IOP 4 cyclone. *Mon. Wea. Rev.*, **122**, 341-365.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A description of the advanced research WRF Version 2. NCAR Tech Note NCAR/TN-468+STR, 88 pp.

Williamson, D. L., J. G. Olson, and B. A. Boville, 1998: A comparison of semi-Lagrangian and Eulerian tropical climate simulations. *Mon. Wea. Rev.*, **126**, 1001-1012.

# TABLES AND FIGURES

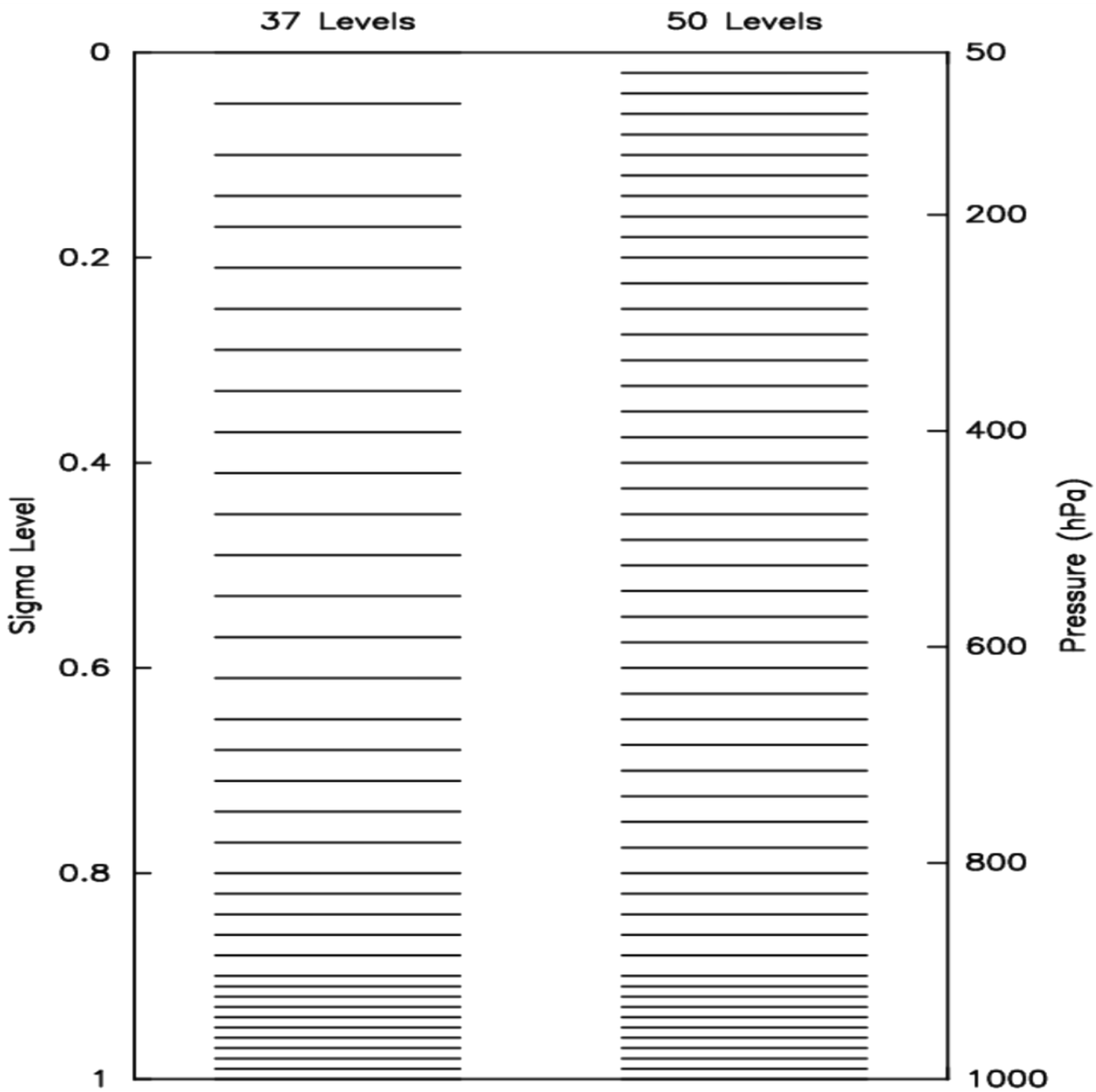


**Figure 1**. Vertical sigma level structure for the 37- and 50-level configurations of the WRF model.

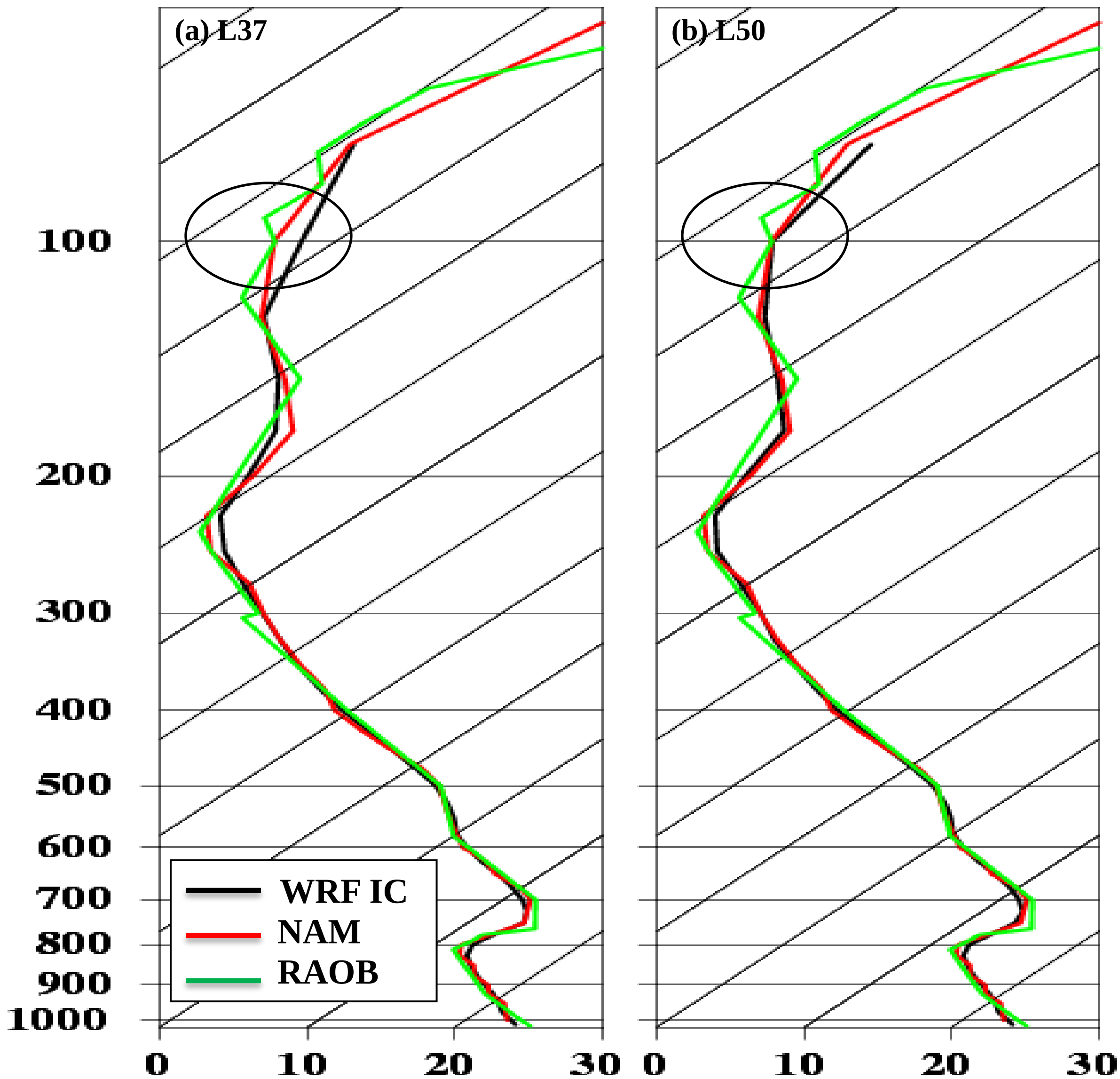


**Figure 2**. Temperature sounding profiles for (a) L37 and (b) L50 configurations at Key West, FL at 0000 UTC 17 January 2007. Black lines represent the WRF initial state, red lines the NAM analysis, and green lines the collocated rawinsonde observation. Circles highlight the level with large differences between L37 and L50.

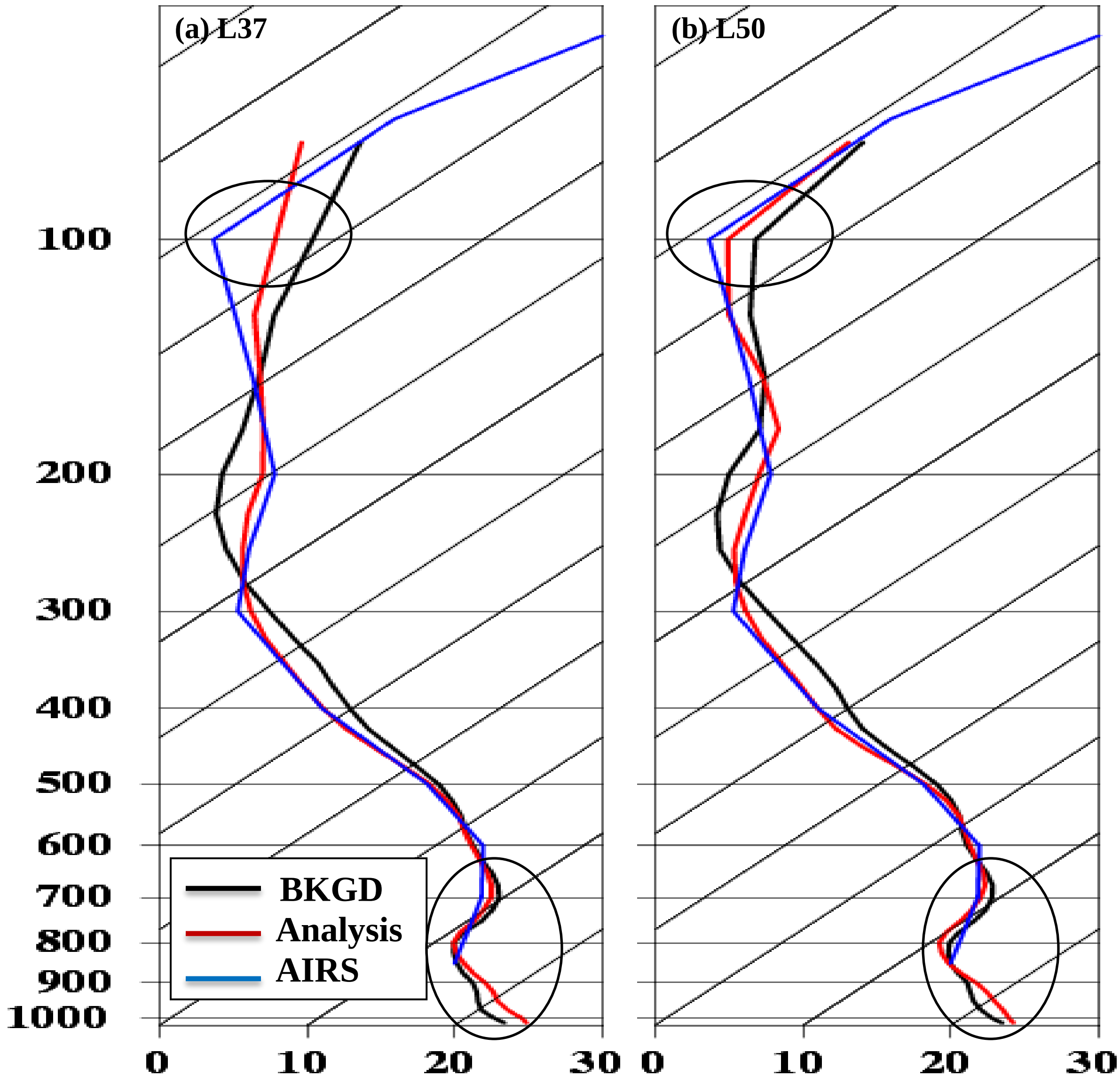


**Figure 3**. Temperature sounding profiles for (a) L37 and (b) L50 configurations at Key West, FL at 0800 UTC 17 January 2007. Black lines represent the background, red lines the WRF-Var analysis, and blue lines the AIRS profile. Circles highlight the differences between L37 and L50 near the tropopause and in the lower troposphere.

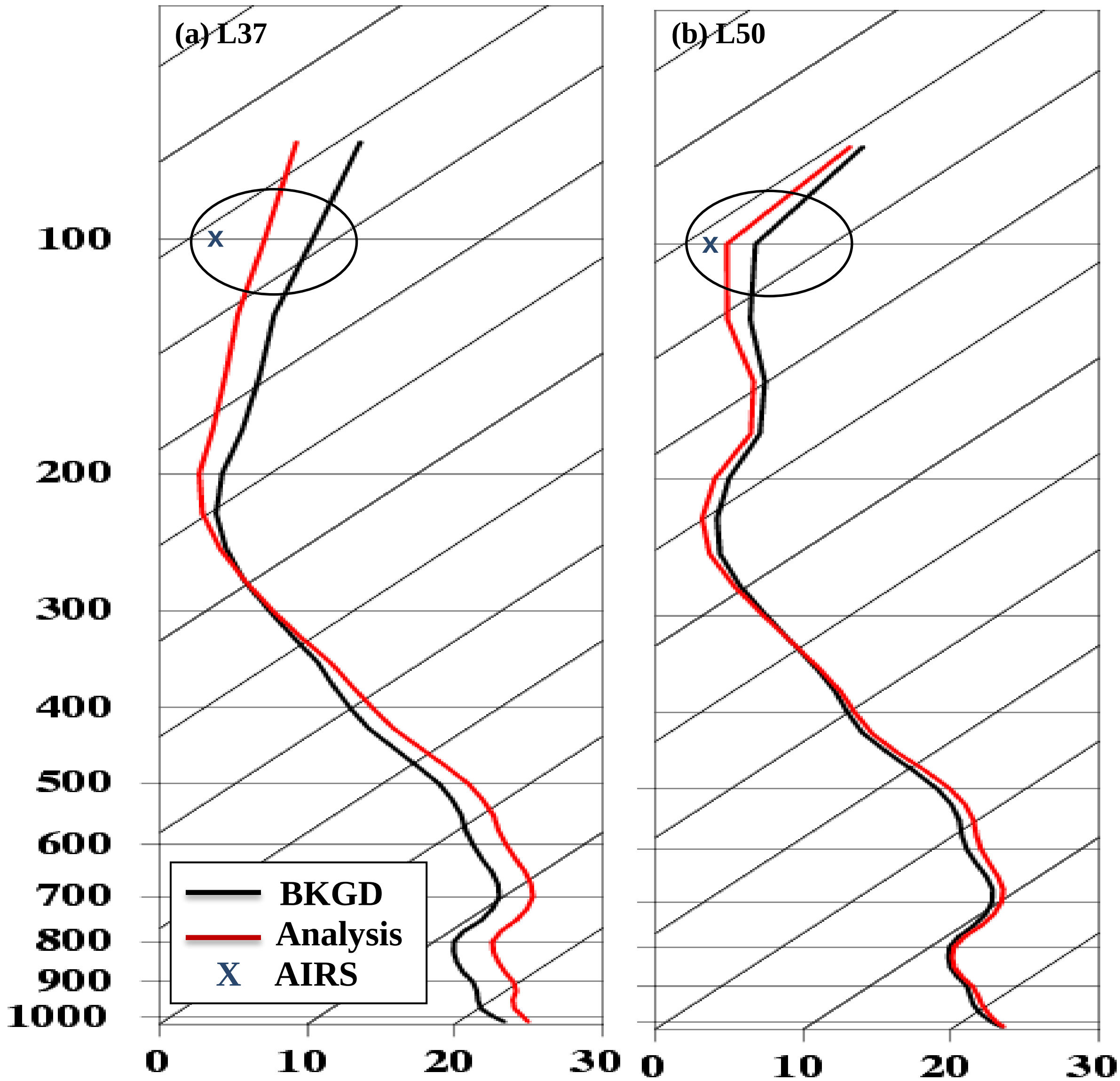


**Figure 4**. Same as Fig. 3, except that only 100 hPa AIRS data are assimilated. The blue 'X' indicates the 100 hPa temperature of the nearest AIRS profile.

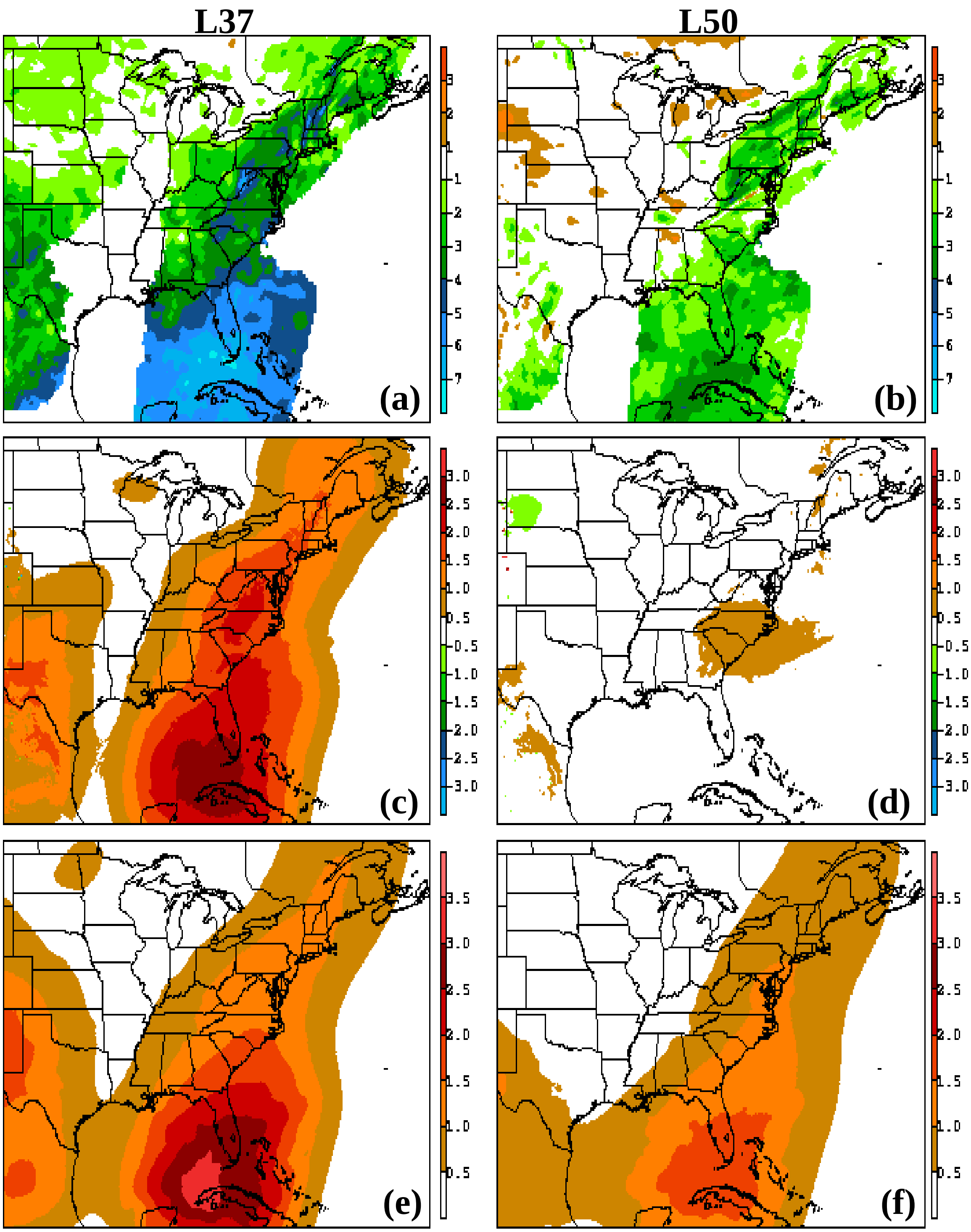


**Figure 5**. Spatial distribution valid at 0800 UTC 17 January 2007 of (a, b) 100 hPa temperature innovation ($^{o}$C), (c, d) 850 hPa temperature increment ($^{o}$C) assimilated with the 100 hPa AIRS data only, (e, f) surface pressure increment (hPa) assimilated with the 100 hPa AIRS data only. Panels (a), (c), and (e) are for L37 and (b), (d), and (f) are for L50.

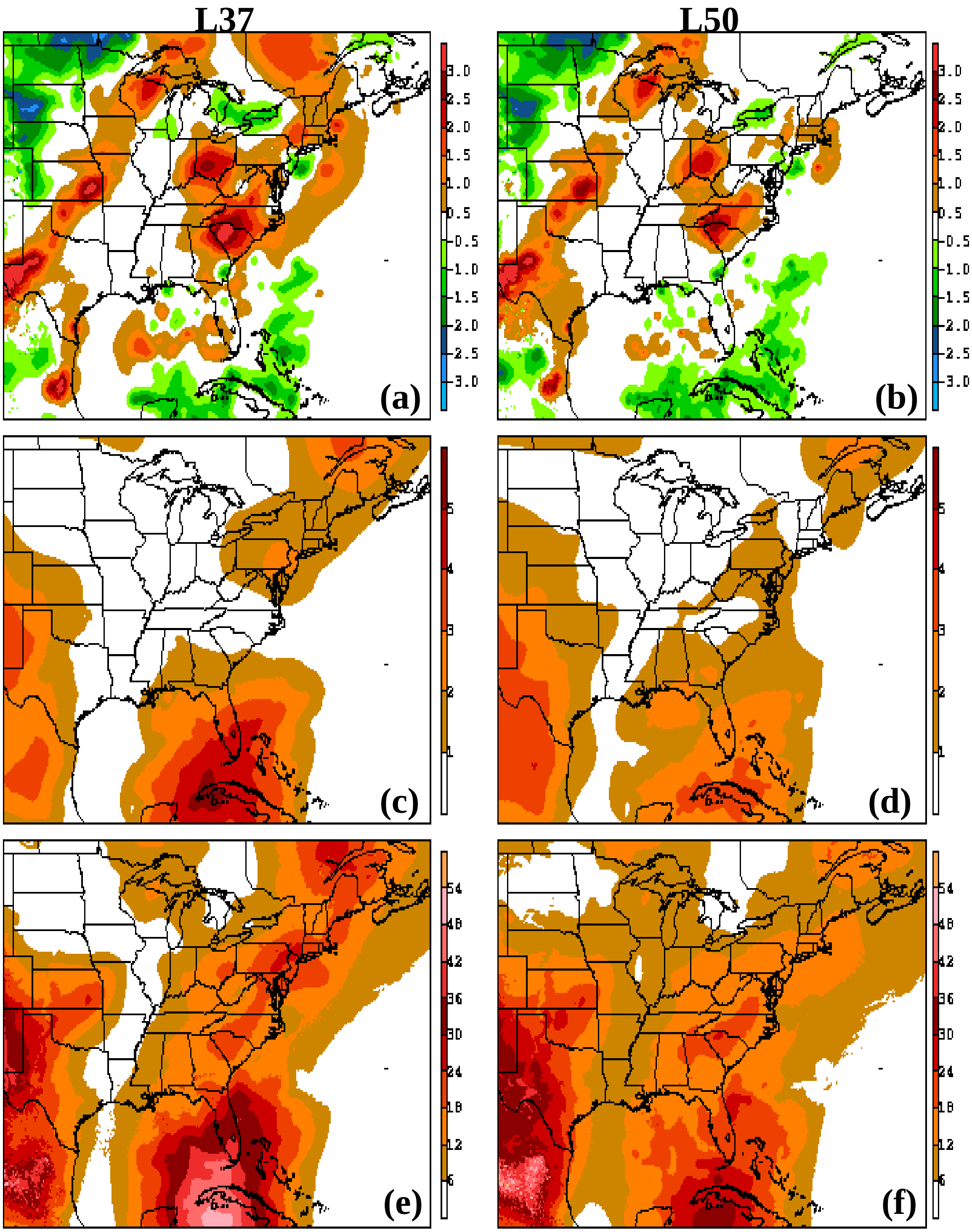


**Figure 6**. Spatial distribution valid at 0800 UTC 17 January 2007 of (a, b) 850 hPa temperature increment ($^{o}$C), (c, d) surface pressure increment (hPa), and (e, f) 850 hPa geopotential height increment (meters) when all the AIRS profiles are assimilated. Panels (a), (c), and (e) are for L37 and (b), (d) and (f) are for L50.

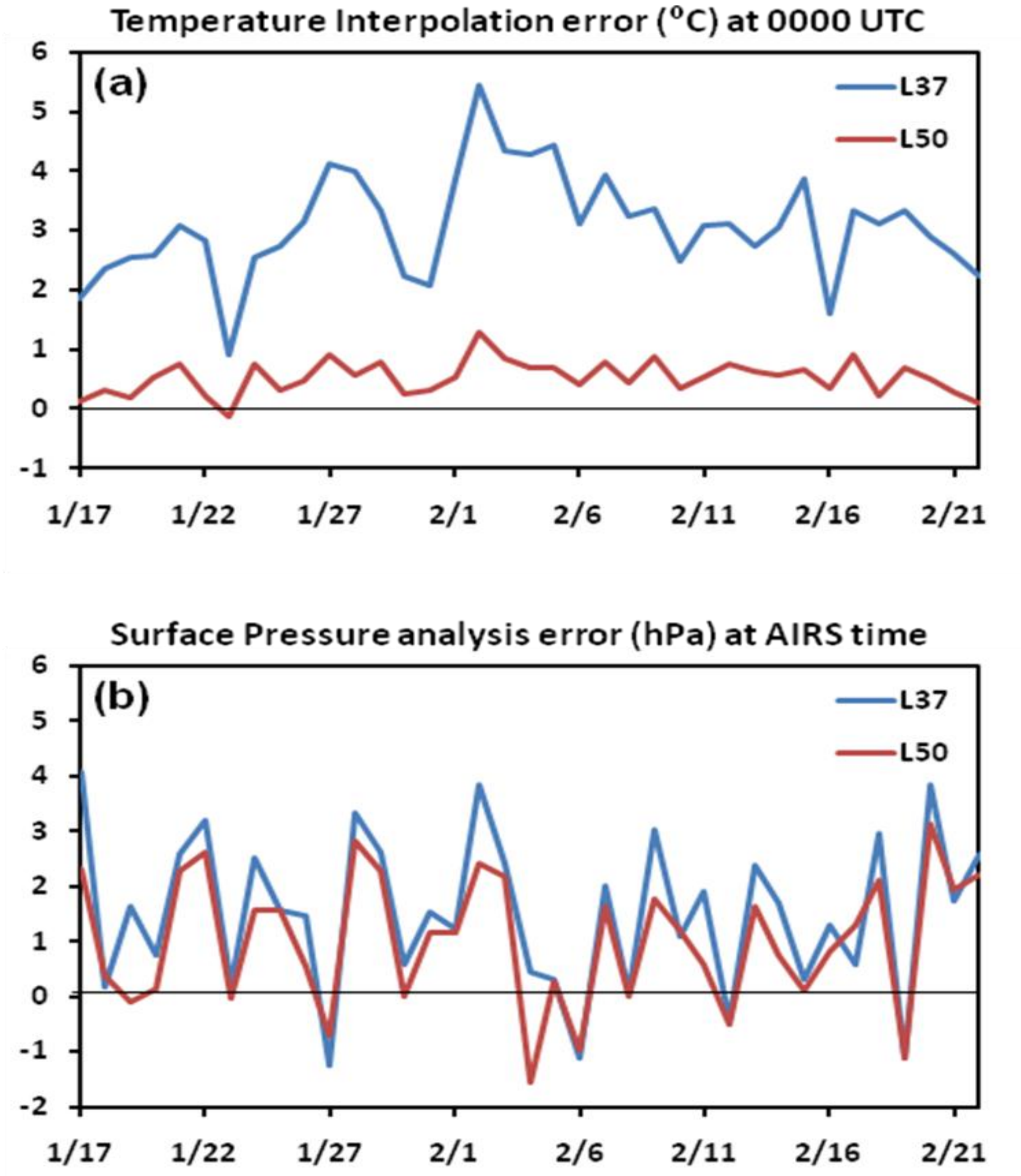


**Figure 7.** Time Series of (a) temperature interpolation errors (°C) against NAM analysis at 0000 UTC and (b) surface pressure error (hPa) against METAR at AIRS analysis time. The test case period is from 17 January to 22 February 2007 and the verification station is Key West, FL.