P5.12 ASSIMILATION OF ATMOSPHERIC INFRARED SOUNDER (AIRS) DATA IN A REGIONAL MODEL

Shih-Hung Chou*
NASA/ Marshall Space Flight Center, Huntsville, AL

Brad Zavodsky
Earth Science System Center, University of Alabama in Huntsville, Huntsville, AL

Gary J. Jedlovec, William Lapenta
NASA/ Marshall Space Flight Center, Huntsville, AL

1. INTRODUCTION

The Atmospheric Infrared Sounder (AIRS) on NASA’s Earth Observing System (EOS) Aqua satellite is currently being used to advance climate research and improve weather prediction on a global scale (Aumann et al. 2003). With access to real time AIRS data from direct broadcast receiving stations, the radiance and profile information can be used to address short-term weather problems.

The NASA Short-term Prediction Research and Transition (SPoRT) Center seeks to accelerate the infusion of NASA Earth science observations, data assimilation, and modeling research into NWS forecast operations and the decision-making process (Goodman et al. 2004). The SPoRT program has been providing output from the Weather Research and Forecasting (WRF; Michalakes et al. 2001) model to several NWS Forecast Offices in the Southern Region for the past year for use as additional guidance in their forecast and decision-making process. Non-operational versions of the SPoRT WRF have been used to prototype new forecast capabilities by assimilating real time EOS data into the WRF model. Some of these efforts are described in companion conference papers (Haines et al. 2006; LaCasse et al. 2006).

One limitation for weather forecasting is the lack of observational data over data sparse regions such as the ocean. Many regions (specifically western North America) are downstream of these data poor areas, so forecasts rely heavily on previous forecasts, which are used as new initial conditions. An approach to this problem is to use AIRS Level-2 temperature and moisture profiles to provide data in regions where traditional rawinsonde observations are not available. This paper will describe a procedure to optimally assimilate AIRS data over the ocean and the subsequent impact on the WRF forecast. The EOS science team version 4.0 temperature and moisture profiles over the eastern Pacific are used to demonstrate this capability. The Advanced Regional Prediction System (ARPS; Xue et al. 2001) Data Assimilation System (ADAS; Brewster 1996) developed at the University of Oklahoma is used to blend the AIRS temperature and moisture information with other observations and a background field provided by WRF forecasts to produce improved initial conditions. Results focus on quality control issues associated with AIRS, optimal assimilation strategies, and the impact of AIRS data on subsequent numerical forecasts.

2. AIRS DATA

Aboard the EOS polar-orbiting Aqua satellite with an early afternoon equator crossing time, AIRS coupled with the Advanced Microwave Sounding Unit (AMSU) form an integrated temperature and humidity sounding system. AIRS is a cross-track scanning infrared spectrometer/radiometer with 2378 spectral channels between 3.7 and 15.4 µm (650 and 2675 cm\(^{-1}\)). Due to its hyperspectral nature, AIRS has the ability to provide near radiosonde-quality atmospheric temperature profiles with accuracy of 1 K in 1 km vertical layers and moisture profiles with accuracy of 20% RH in 2 km vertical layers. Because AIRS footprints coincide with AMSU footprints, the AIRS Science Team uses AMSU data in the retrieval process, producing a uniform distribution of AIRS retrievals in both clear and cloudy scenes at a spatial resolution of 50 km (Aumann et al. 2003).

* Corresponding author address: Shih-hung Chou, 320 Sparkman Dr., Huntsville, AL 35805; e-mail: shih.h.chou@nasa.gov
Table 1. Description of the 6 quality indicators for version 4.0 AIRS Level-2 temperature profiles (Olsen et al. 2005).

<table>
<thead>
<tr>
<th>Quality Indicator</th>
<th>Description</th>
<th>Total % in Domain</th>
<th>Color in Fig. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>complete MW and IR (highest quality)</td>
<td>16.3</td>
<td>black</td>
</tr>
<tr>
<td>SFC Failed</td>
<td>complete MW and IR; fails QC for emissivity</td>
<td>29.2</td>
<td>blue</td>
</tr>
<tr>
<td>SFC+BOT Failed</td>
<td>complete MW and IR; fails QC below 3 km</td>
<td>13.5</td>
<td>green</td>
</tr>
<tr>
<td>SFC+BOT+MID Failed</td>
<td>complete MW and IR; fails QC below 200 hPa</td>
<td>19.5</td>
<td>orange</td>
</tr>
<tr>
<td>SFC+BOT+MID+TOP Failed</td>
<td>MW only retrieval</td>
<td>13.0</td>
<td>yellow</td>
</tr>
<tr>
<td>All Failed</td>
<td>No retrieval</td>
<td>8.4</td>
<td>red</td>
</tr>
</tbody>
</table>

The AIRS science team version 4.0 data (Fetzer et al. 2005) used herein provides more accurate retrievals than its predecessors. The 4.0 temperature profiles have been validated by the science team over both land and ocean with RMS profile errors ranging from 0.6-1.0 K (ocean ±50° of the equator) to 0.9-1.3K (global ocean and land) for complete retrievals (corresponding to roughly 40-60% of all profiles). Although the moisture profiles are improved over previous versions, 4.0 moisture data has not yet been fully validated over land due to surface emissivity issues in the retrieval process. AIRS science team retrieval accuracy goal for relative humidity is 20% over 2-km layers.

The version 4.0 profiles contain 6 quality indicators, which provide qualitative information on the accuracy of the retrievals. Table 1 provides a description of each quality indicator. The quality indicators provide information on the retrieval process and the status of internal quality checks of the temperature data for various layers of the atmosphere.

3. ANALYSIS SCHEME

The ADAS is designed to address regional-scale data assimilation in real-time numerical weather forecasting. It provides a means to merge different sources of local meteorological data into a single, coherent three-dimensional description of the atmosphere. The data can be from remotely sensed sources or Meteorological Assimilation Data Ingest System (MADIS) (i.e., rawinsonde, ACARS, METAR, buoys, wind profiler, etc.) and are combined using the Bratseth (1986) successive correlation method.

The Bratseth method uses observations to iteratively update a first-guess field provided by a model forecast. The gridpoint estimate for an analysis variable, $\phi$ (i.e., temperature, moisture, pressure, or wind), is

$$
\phi_x(k+1) = \phi_x(k) + \sum_{i=1}^{N_{obs}} \alpha_{x_{i}} [\phi_{i}^{obs} - \phi_{x}(k)],
$$

where $\phi_x(k+1)$ is the analysis for the $k^{th}$ iteration, $\phi_x(k)$ is the analysis value at the grid point for the previous iteration (background value if $k=1$), $[\phi_{i}^{obs} - \phi_{x}(k)]$ is the innovation (observation minus $k^{th}$ iteration grid point estimate), and $\alpha_{x_{i}}$ is a weighting factor. The weighting factor is a function of the distance of the observations from each grid point and the ratio of the observation and background error variances.

Correlation with respect to distance between observations and grid points is assumed to be Gaussian in nature:

$$
\exp\left(-\frac{r_{ij}^2}{R^2}\right) \exp\left(-\frac{\Delta z_{ij}^2}{R_z^2}\right),
$$

where $r_{ij}$ and $\Delta z_{ij}$ are respectively the horizontal and vertical distances between observations and gridpoints and $R$ and $R_z$ are respectively the horizontal and vertical scaling factors, which are a function of data decorrelation and define how much smoothing will occur in the analysis. The error variances—incorporated into ADAS via error tables—quantitatively describe the degree of confidence of each data set as a function of pressure. The errors in these tables combine instrument error and representativeness of the data (i.e., how information from a point observation is spread across a grid cell). A smaller error indicates larger weighting on a data type in the final analysis. For instance, if the error for an AIRS profile is less than for the background, the final analysis will more closely reflect the AIRS value.

Generally, data sets are assimilated from largest impact to smallest impact with data sets possessing similar characteristics assimilated in the same pass. Usually, data are assimilated in
4. EXPERIMENTAL DESIGN

4.1 Case Study

The 14-17 January 2004 period was selected as a case study because it represented a typical situation where AIRS profiles could provide additional information to the forecast process. The area of interest was the western portion of the continental U.S. and the eastern Pacific where a low-pressure system over the central Pacific was propagating northeastward. A weak ridge over the eastern Pacific produced a relatively cloud-free environment allowing for a large number of high-quality AIRS soundings to be produced. Over the

more than one pass to guarantee convergence of the Bratseth method. The tuning of the error tables and scaling factors for ADAS used in this study will be described in Section 4.3.
48h forecast period, the trough moves to the northeast eventually bringing precipitation to the Pacific Northwest. Figure 1 shows analyses from the National Center for Environmental Prediction’s (NCEP) Global Forecast System (GFS) (Kanamitsu et al. 2002) at the beginning of the forecast cycle and at the 48h forecast. Figure 2 shows an infrared image of the conditions at 00 UTC on 15 January 2004 showing the conditions at the beginning of the cycle and the propagation of the trough.

The Aqua satellite overpass dictates the timing of the assimilation of AIRS data into the forecast model. Two consecutive AIRS passes over the eastern Pacific occur within 2 hours of each other as seen in Fig. 3. The first is at 22 UTC on 14 January 2004 and the second is at 00 UTC on 15 January 2004. Within the model domain, there are 3842 AIRS profiles. The approximate location of the AIRS overpasses with respect to the cloud field is shown by the red outlines in Fig. 2; the exact locations of each AIRS profile along with its appropriate quality indicator are shown in Fig. 3. The quality indicators show a broad spatial distribution of AIRS retrievals of various quality in the two passes. Table 1 presents a tally of the number of profiles in each quality category in the WRF model domain. While only about 16% of the retrievals pass all quality checks, nearly half (46% as indicated by the black and blue retrieval locations in Fig. 3) passed atmospheric temperature quality checks. The majority of the retrievals that fail some form of atmospheric quality check occur in overcast conditions (compare Figs. 2 and 3).

### 4.2 Forecast Model Configuration

The forecast model used herein is the WRF model, a next-generation mesoscale numerical weather prediction system designed to serve both operational forecasting and atmospheric research needs. It is a limited-area, non-hydrostatic primitive equation model with multiple options of physical parameterization schemes. The model domain used in this study consists of a 360 × 200 grid with 30 km spacing and covers the central Pacific and CONUS. It has 31 staggered terrain-following sigma levels with the top-level pressure at 100 hPa and finest resolution near the boundary layer. Table 2 summarizes the physical options used for this study.

#### 4.3 ADAS Configuration

Special processing of the WRF and ADAS was required to link the assimilation system to the forecast model. The ADAS has 43 sigma levels separated by an average of 500 meters with emphasis on more levels near the top and bottom. Because ADAS does not have the capability to directly define a vertical domain to match the WRF, vertical interpolation is necessary to adjust the background (first-guess) field.

A series of sensitivity tests have been run with AIRS temperature and moisture profiles in ADAS to understand the effect on the resulting analysis of two specific tuning parameters: horizontal and vertical scaling factors and error variance tables. The result led to the selection of ADAS parameters that were deemed appropriate for the AIRS profile characteristics. It is reasonable to select scaling factors to match the estimated decorrelation of each data type. For the configuration used in this study, the AIRS data is assimilated in two passes with a final horizontal scale factor of 120 km—approximately 3 times the spacing of the nominal AIRS profiles—resulting in a smooth yet representative analysis. For the vertical scaling factor, AIRS data was assigned a value of 750 m to take into account its layer nature. Testing indicates that a vertical scale factor of 1000 m produces too smooth of a temperature profile in the vertical but that 500 m introduces noise. At the present time, these factors are applied equally to temperature and moisture profile data although the data quality and scale of variability of moisture data may require separate values in future studies. Table 3 shows the data assimilated in each pass along with the appropriate scaling factors used in the study.

Error tables for the MADIS data for all the experiments in this paper were left with the specifications recommended by the University of Oklahoma. The only exception is with the background error table. Even though 2- and 4-hour WRF forecasts are used as first guesses, error tables for 3-hour Rapid Update Cycle (RUC) forecasts are used because of their availability. This is not an issue as model forecasts errors are comparable (Brewster, personal communication).
Table 3. Overview of ADAS assimilation parameters.

<table>
<thead>
<tr>
<th>Data Assimilated</th>
<th>Vert. Scale Factor (m)</th>
<th>Horiz. Scale Factor (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pass 1 AIRS, RAOB, WPF</td>
<td>750</td>
<td>180</td>
</tr>
<tr>
<td>Pass 2 AIRS</td>
<td>750</td>
<td>120</td>
</tr>
<tr>
<td>Pass 3 RAOB, ACARS, WPF</td>
<td>400</td>
<td>100</td>
</tr>
<tr>
<td>Pass 4 ACARS, BUOY, METAR, SAO</td>
<td>N/A</td>
<td>80</td>
</tr>
<tr>
<td>Pass 5 BUOY, METAR, SAO</td>
<td>N/A</td>
<td>60</td>
</tr>
</tbody>
</table>

Fig. 4. Error tables used for ADAS analysis of AIRS data for a) temperature and b) relative humidity. The black line is for the WRF forecast background, the red line is for the AIRS full data, and the blue line is for the AIRS surface- and bottom-rejected data. Note that where the bottom of the AIRS profiles are invalid the error is 3-4 times larger than the background indicating that the background will have more weighting in the analysis.

The black lines in Fig. 4 denote the background error table values. The error table used for AIRS is essentially the instrument error described in Olsen et al. (2005). The red lines in Fig. 4 are the error tables used for the first three experiments described in Section 4.4; the blue lines in Fig. 4 are the error tables used for the last experiment described Section 5.2.

4.4 Numerical Experiments

The first-guess field and boundary conditions for the analysis are provided by the GFS, which is available every 6 hours. Currently, the operational version of the GFS assimilates radiance data from instruments aboard GOES and NOAA polar-orbiting satellites. The infrared instruments (GOES sounders and HIRS) provide limited spatial coverage in cloud free regions, while the microwave instruments (AMSU) provide a more global coverage. This implies that the first-guess field used to initialize the WRF provides large-scale thermodynamic information that may be sporadically influenced by satellite data.

The SPoRT ADAS/WRF assimilation cycle begins at 18 UTC on 14 January 2004 with a 4-hour WRF forecast initialized with the GFS analysis. This forecast is, then, used as the first-guess field for the ADAS analysis at 22 UTC. All valid MADIS and AIRS data are assimilated at this time, and the analysis and boundary conditions are created. The analysis and boundary conditions are used to produce a 2-hour WRF forecast, which becomes the first-guess for the ADAS analysis at 00 UTC on 15 January 2004. This second ADAS analysis and boundary conditions are used for a 48-hour WRF forecast with no further data assimilation. Figure 5 shows the ADAS/WRF assimilation/forecast process.

The impact of the AIRS profiles is examined through four experiments using AIRS data of varying quality. The control (CNTL) experiment used only MADIS data for the analysis; this represented the best results without additional data over the eastern Pacific. To assess the impact of the best AIRS retrievals but not necessarily the best data coverage, full retrievals and MADIS data were combined in a second experiment (FULL). A third experiment, adding profiles that only failed surface and/or low-level checks, provided additional data coverage (SFBT). The spatial distribution and coverage of the AIRS profiles with quality flags was presented in Fig. 3. This coverage was presumably at the risk of including lower quality data. A final experiment (NEWT) incorporates upgraded AIRS error table
Fig. 5. Schematic of the ADAS/WRF assimilation/forecast procedure.

Fig. 6. Skew-T plots for the FULL (red), SFBT (green), and NEWT (blue) experiments after 22 UTC ADAS analysis. Also shown are the background (black) and the nearest full-retrieval AIRS profile (orange).

Fig. 7. Location of AIRS full-retrieval soundings (orange) and RAOB validation stations (red). Numbers denote 800 hPa temperature of AIRS and RAOB stations. X marks the sounding location in Fig. 6.

into the analysis to reflect the different quality flags associated with AIRS data. The difference between the NEWT experiment and the SFBT experiment is that the error for the AIRS data is increased in the lowest levels for the surface-failed and surface- and bottom-failed data to reflect the degradation in the quality of the AIRS profiles that have been rejected below 3 km (blue lines in Fig. 3).

5. MODEL RESULTS

5.1 Effect of AIRS Profile on Initial Analysis

The effect of AIRS profiles on the initial analysis provided by ADAS is illustrated in Fig. 6, which shows the 22 UTC soundings of the FULL, SFBT, and NEWT analyses with the background and closest AIRS full-retrieval sounding. The model grid at 42.86 N and 135.89 W (marked x in Fig. 7) is selected because it is surrounded by lesser-quality AIRS data with the nearest full-retrieval AIRS data approximately 1° latitude away. This grid point provides an opportunity to examine the effect of AIRS data of various quality and the AIRS error tables on the ADAS analysis. As shown in Fig. 6, at the grid point, the temperature profiles are very similar among the experiments, while the moisture profiles show more variation. Using only the full-retrieval AIRS data, the analysis deviates only slightly from the background due to the distance between the analysis point and the observation. When surface-failed, and surface- and bottom-failed AIRS data are also included in the analysis, the resulting profile (green) moves further away from the background and closer to the full-retrieval AIRS profile. Notice that the SFBT profile displays more vertical variation than the AIRS profile, indicating the influence of the lesser-quality but nearer AIRS data on the analysis.

5.2 Verification Statistics

The upper-air verification statistics for the 48 h forecast (at 12 hour intervals) are computed by comparing the rawinsonde value to the model forecast values interpolated to the location of that specific observation. In this study, verifications are based on 12 RAOB stations in the western states (see Fig. 7), where AIRS data will have the largest impact on the 0-48h forecast.

Results for the 24h forecast are presented in Fig. 8. The control forecast (black line) performs
fairly well at all levels below 200 hPa with RMS values ranging from 1 to 2°C for temperature and generally less than 1.0 g kg\(^{-1}\) for moisture. The largest RMS temperature errors occur at 500 and 200 hPa and in the lowest layers for moisture. The inclusion of a small number of high quality (full-retrieval) AIRS data (red line) yields a little improvement in RMS temperature errors and a slight increase in RMS error for low-level moisture. In the mid-troposphere, temperature RMS error has a reduction of approximately 0.2°C at 500 hPa.

When lesser-quality AIRS data (the surface-failed and surface-and-bottom-failed) are also included in the analysis, there is a substantial reduction in RMS temperature errors in the layers between 700 and 300 hPa with a slight increase in error near the surface. Inclusion of these additional profiles provided spatial information not contained in the high quality retrievals because of their limited coverage (see black points in Fig. 3). However, a noticeable degradation in the moisture forecast (i.e., increased RMS) occurs below 700 hPa with the inclusion of the lesser-quality profiles. Presumably some of the moisture information from these additional AIRS profiles is not correct. While the inclusion of these profiles degrades the forecast in the lower layers, they improve the forecast quality between 250 and 600 hPa by approximately 0.5°C for temperature and 0.1 g kg\(^{-1}\) for moisture. The assimilation challenge is to develop a way to assimilate these profiles into WRF to maximize their impact. For retrievals that have failed the surface and bottom quality checks, reducing the weighting below 700 hPa may provide a way to do this. To test this approach, a revised AIRS error table was developed which reduced the weight (influence) of the AIRS data in the lower levels (the blue lines in Fig. 4). The data was re-assimilated with the new error weights and a WRF 48 h forecast was produced. At the initial assimilation time, the resulting analysis profile using the new error table shifted closer to the analysis profile obtained from the FULL experiment (compared to the SFBT) as expected (see Fig. 6). The validation results for this new run (NEWT) are presented in Fig. 8 (blue lines). Using revised AIRS tables tailored specifically to the lesser-quality data, positive results were obtained. Overall, the RMS temperature and moisture improvement made with the previous data inclusion were maintained, but the suppression of the poor quality low-level information wasn’t fully achieved. New error tables, which further reduce the impact of the low level AIRS data for these profiles, needs to be used to completely suppress this information.

Based on the above experiments, there are several ways to improve the implementation of AIRS data in an analysis model. When an observation fails a quality check for temperature, both temperature and moisture are flagged, irrespective of their individual quality. By applying separate quality indices for temperature and moisture, the number of usable AIRS profiles can be increased. The other improvement can be achieved by assigning quality indices to each individual level instead of the layer structure of the current error table (e.g., bottom-failed for below 700 hPa). Version 5.0 (available in Summer 2006) will improve over version 4.0 by providing level-specific quality indicators independently for
temperature and moisture (Susskind, personal communication). These improvements may mitigate some of these problems and improve forecast quality.

6. CONCLUSIONS

A procedure has been developed to incorporate AIRS Level-2 version 4.0 temperature and moisture sounding data into the ADAS analysis using an approach to maximize the utility of the available data. Preliminary experiments show that this technique, when applied to AIRS data in the eastern Pacific, can improve weather forecasts for western North America. The fact that the test case was chosen because of data availability and not because of bad forecast indicates that the AIRS data has a potential to improve the routine forecast. Several further studies are currently under consideration. Since the impact of the Pacific AIRS soundings do not reach the West Coast until 48 hours later, the model domain can be reduced to cover only the western half of the CONUS and a finer grid spacing can be utilized. This new domain can also be applied to test cases on different dates. Another possible test case can be set up to examine the impact of AIRS data over the Gulf of Mexico on the forecast for the southeastern United States.

Acknowledgements

The authors would like to thank Kate LaCasse and Dr. Steven Lazarus for their assistance in setting up the ADAS runs and their helpful comments during the course of the research. This research was funded by the NASA Science Mission Directorate’s Earth-Sun System Division in support of the Short-term Prediction and Research Transition (SPoRT) program at Marshall Space Flight Center.

REFERENCES


