AIRS Real-Time Sounding Profile Retrieval For IMAPP (International MODIS/AIRS Processing Package) Users

**Elisabeth Weisz, Hung-Lung Huang, Jun Li,
Suzanne Seemann, Eva Borbas, Liam Gumley**

*Cooperative Institute for Meteorological Satellite Studies,
University of Wisconsin-Madison, Madison, Wisconsin, U.S.A*

Introduction

The high-spectral-resolution Atmospheric Infrared Sounder (AIRS) was launched onboard the NASA (National Aeronautics and Space Administration) Earth Observing System (EOS)-Aqua platform on May 4, 2002. AIRS is complemented by AMSU (Advanced Microwave Sounding Unit) and HSB (Humidity Sounder Brazil) for the primary purpose to retrieve vertical temperature and humidity profiles. AIRS measures radiation in 2378 channels at a spectral resolution of $\lambda/d\lambda=1200$ in the infrared region from 3.7 to 15.4 μm. AIRS is a scanning instrument with a footprint at nadir of 13.5 km and 90 footprints per scanline. Global coverage is accomplished twice a day. Six minutes of AIRS data, known as a granule, comprises 135 scanlines.

The unprecedented accuracy and global wealth of AIRS sounding products will support climate studies and various research efforts such as weather prediction. Current work at CIMSS includes the development of a real-time Direct Broadcast (DB) processing algorithm to retrieve atmospheric parameters in the troposphere and lower stratosphere as well as surface parameters as part of the International MODIS (Moderate Resolution Imaging Spectroradiometer)/AIRS Processing Package (IMAPP). This paper reviews and discusses the pre-release version of the IMAPP-AIRS retrieval methodology. The algorithm is a statistical eigenvector regression retrieval to obtain a fast and accurate first estimate of the atmospheric state. Evaluation of the results is done by comparison with ECMWF (European Center of Medium-range Weather Forecasts) analysis data, operational AIRS products, radiosonde observations, and retrievals from GOES (Geostationary Operational Environmental Satellite) and MODIS.

Retrieval Methodology

In the regression retrieval the relationship between atmospheric state X and the measurements Y can be simply described by

$$X=CY^T \tag{1}$$

where C is the matrix of regression coefficients. The least-square regression solution proposes

$$C=XY(Y^TY)^{-1} \tag{2}$$

The coefficients, derived from a sufficiently large training set of representative atmospheric profiles and surface parameters, are then applied to any real-time measurement spectrum in the retrieval step (1). In practice, one often uses the deviations of the radiances from the sample mean for Y. The current retrieval methodology utilizes an eigenvector regression algorithm (Smith and Woolf 1976, Huang and Antonelli 2001) where regression of compressed radiances against the atmospheric state is performed. The eigenvectors (or principal components) are calculated from the covariance matrix of the training data set of simulated radiances. Performing eigenvector regression with only a few eigenvectors (those with largest eigenvalues) reduces the dimension of the regression problem, as well as the impact of random noise. Investigating retrieval results for different numbers of eigenvectors we found that about 30 eigenvectors seem sufficient for successful retrieval performance. Our current

retrieval products include atmospheric temperature, moisture and ozone profiles, Total Precipitable Water (TPW), total ozone, surface skin temperature and surface emissivity.

The global training data set comprises TIGR3 and NOAA88 profiles supplemented by special desert and polar cases. To characterize surface properties an ecosystem type has been assigned to each pixel to obtain realistic surface pressure, surface skin temperature and surface emissivity. Radiances are calculated for all profiles in the training database using the Stand-Alone Radiative Transfer Algorithm (SARTA, Strow et al. 2003). Furthermore, we classify the training set of radiances on the basis of their Brightness Temperature (BT) in Kelvin (K) at 1000 cm^{-1}. Regression coefficients are computed for each of the six classes identified by the threshold values outlined in Table 1. For the retrieval the observed spectrum is assigned to the appropriate class and the corresponding set of coefficients is used. Note that each class in the training has an overlap of 10 K with its upper and lower neighboring class. This is to ensure that when the retrieval is misclassified the use of non-optimal class retrieval coefficients (likely from an adjacent class) still produces reasonable results. Another classification we performed is based on the scanning angle. Regression coefficients have been calculated for different scanning angles between 0 and 49 degrees. We found that both classification procedures refine the output and optimize retrieval performance.

Table 1: Brightness Temperature (BT) classification scheme (values are given in Kelvin).

Class	Training: BT at 1000 cm^{-1}	Retrieval: BT at 1000 cm^{-1}
1	BT ≤ 260	BT ≤ 255
2	250 < BT ≤ 270	255 < BT ≤ 265
3	260 < BT ≤ 280	265 < BT ≤ 275
4	270 < BT ≤ 290	275 < BT ≤ 285
5	280 < BT ≤ 300	285 < BT ≤ 295
6	290 < BT	295 < BT

Results

The retrieval algorithm, as described above, has been applied to global as well as local observational data. Fig. 1 illustrates surface skin temperature and surface emissivity at wavenumber 926 cm^{-1} over the CIMSS (Cooperative Institute for Meteorological Satellite Studies) DB area for October 23, 2003.
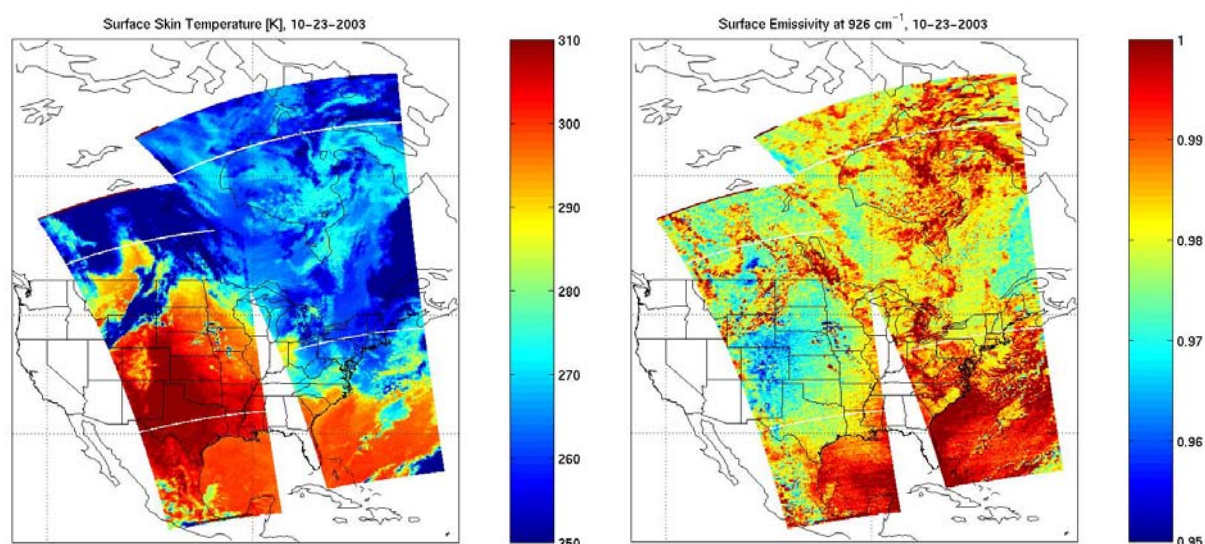

Fig. 1: Real-time DB retrieval of surface skin temperature (left) and surface emissivity at 926 cm^{-1} (right) for 10-23-2003.

Note, the retrieval is shown for all pixels regardless of whether they are cloudy or clear. Since our retrieval algorithm is designed to process clear-sky radiances only, we need a reliable cloud mask to identify clear and cloudy field-of-views. To address this issue we are making use of the 1km-MODIS cloudmask product (Ackerman et al. 1998). Daytime granule 192 on September 2, 2003 has been selected to illustrate clear scene identification. Fig. 2 and Fig. 3 show retrievals of temperature at 850 mbar and TPW, respectively, for all pixels (left panel) as well for clear pixels only (right panel). The retrievals are being compared with the corresponding ECMWF analysis fields. Note that the analysis data has been spatially interpolated to AIRS measurement locations, and the difference in time is approximately one hour. For temperature we find some differences in the fine-scale structure between retrievals and analysis but the spatial average is in good agreement. Differences are more drastic in the case of humidity (not shown) and TPW (Fig. 3), although some features are similarly depicted.

The Root-Mean-Square (RMS) error of the difference between analysis and retrieval for all clear pixels (approximately 4500 in number) is given in Fig. 4 for temperature and humidity profiles at each pressure level as well as for skin temperature and TPW. Largest values occur at and near the surface, partly caused by surface parameters (in particular surface skin temperature and surface emissivity) not being sufficiently and/or incorrectly represented in the training data set. Another reason is the assumption of linearity between measurements and atmospheric state in the regression retrieval problem. An iterative non-linear physical retrieval method will yield significantly smaller errors in the boundary layer. It should be mentioned that the 1-km (for temperature) and 2-km (for humidity) "layer mean RMS" profiles are commonly used by the AIRS science team. Layer mean RMS is significant smaller than the corresponding level RMS for both temperature and humidity retrievals.

Intercomparison with an analysis filed may not be sufficient to characterize the retrieval error. The residual, which is the difference between the observations and the calculated radiances (the latter calculated by the forward model using the retrieved profile as the input), is more indicative of how the retrieved and the model data represent the actual state of the atmosphere. Fig. 5 shows the RMS error, the standard deviation and the mean (or bias) of the residuals of the clear pixels as a function of wavenumber. Results are shown for the retrieved granule (left) and for the analysis (right) in the spectral range from 500 to 1700 cm^{-1}. All three quantities are significant smaller for the retrieval than for the analysis, in particular in the long-wave atmospheric window region where the difference can be as large as 3 K.
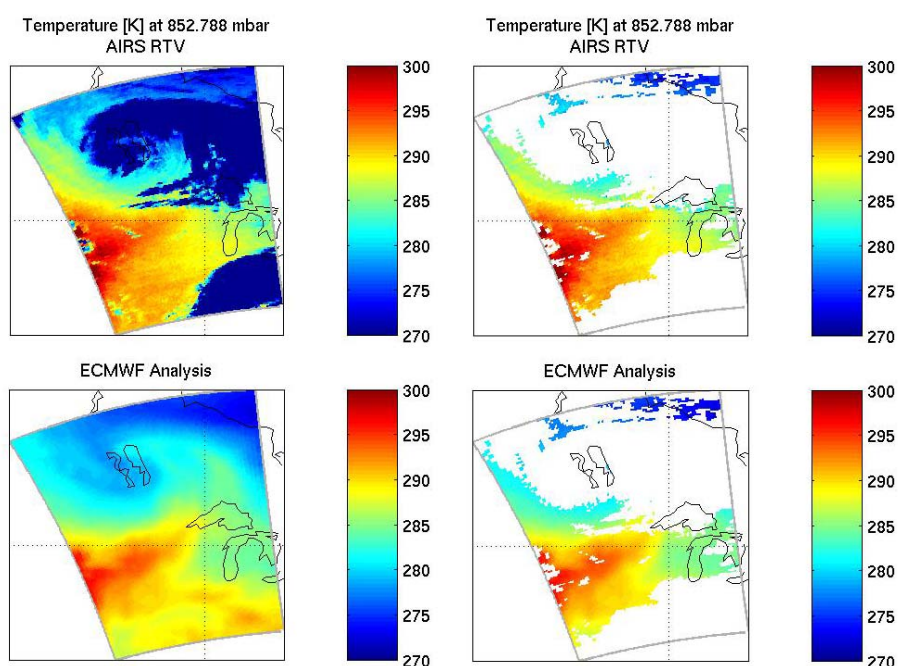

Fig. 2: Retrieved temperature [K] at ~850 mbar (top) for granule 192 (09-02-2003) compared with ECMWF analysis (bottom) with cloud filtering (right) and without (left).
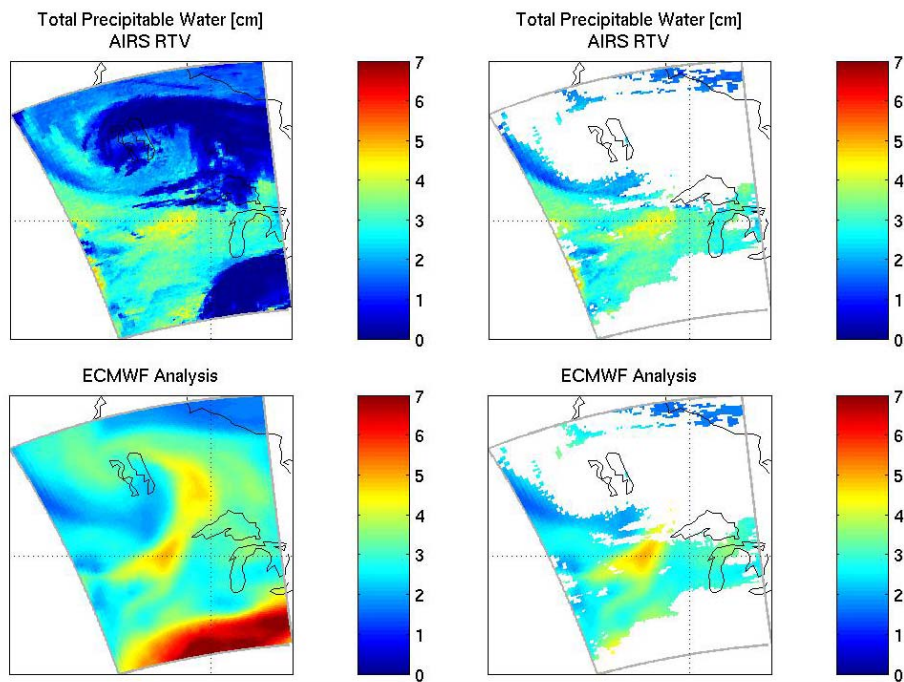
Fig. 3: Retrieved TPW [cm] (top) for granule 192 (09-02-2003) compared with ECMWF analysis (bottom) with cloud filtering (right) and without (left).
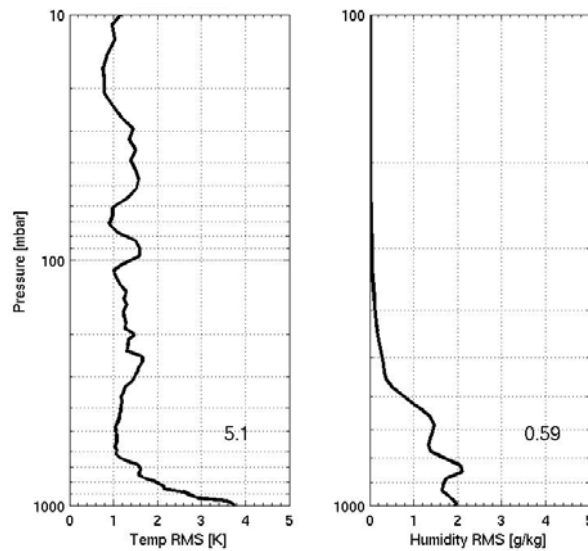

Fig. 4: RMS (at each pressure level) of ECMWF analysis minus retrieval for temperature (left) and humidity (right) shown for clear pixels of granule 192 (09-02-2003). The numbers inside the panels refer to the RMS error of skin temperature [K] and of total precipitable water [cm] respectively. Note the lowest pressure level shown is 10 mbar for temperature and 100 mbar for humidity.
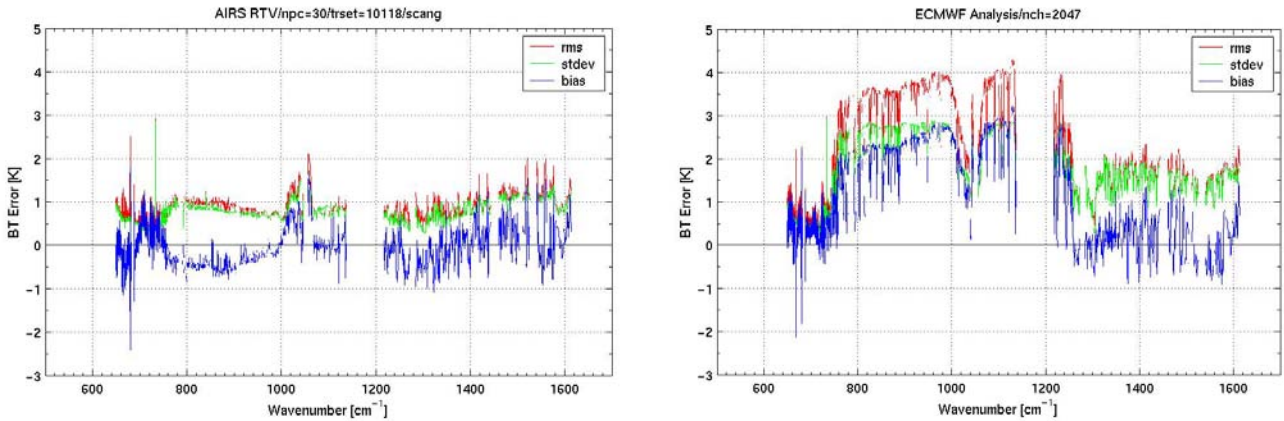
Fig. 5: Root-mean-square error (red), standard deviation (green) and bias (blue) of brightness temperature residuals as a function of wavenumber for the retrieval (left) and the ECMWF analysis (right).

Comparison with MODIS (Seemann et al. 2003) and GOES retrieval (Schmit et al. 2002) is displayed in Fig. 6 for temperature and moisture at 620 mbar for the same granule as above. This figure demonstrates that the IMAPP AIRS retrieval is in good agreement with operational and well-validated retrieval products. It should be mentioned that the different coverage is due to different cloudmasks and different spatial resolution of the measurements.
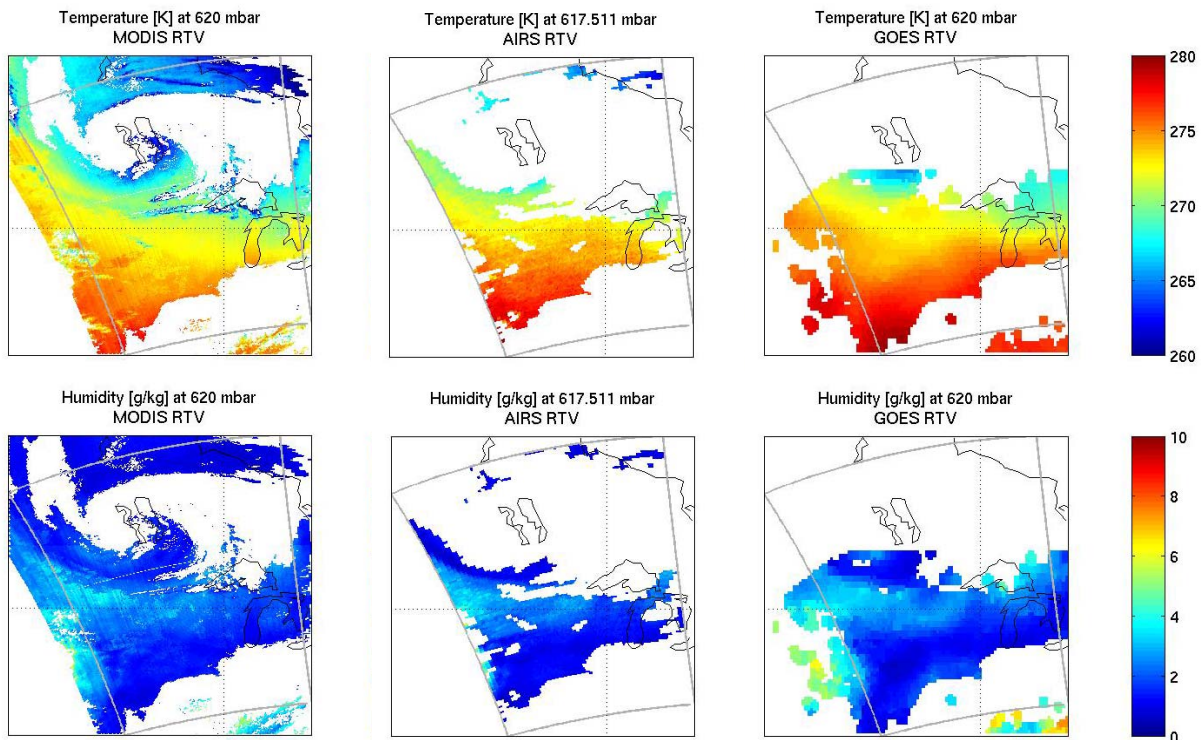

Fig. 6: Retrievals of temperature (top) and moisture (bottom) at 620 mbar for the MODIS (left), AIRS (middle) and GOES (right) instrument.

An ongoing task is to compare the IMAPP AIRS retrieval results with the L2 Standard Retrieval Product (called operational product hereafter) which has been made public available by the AIRS Science Team in September 2003. The operational products are obtained by using cloud-cleared radiances calculated for an AMSU footprint (Susskind et al. 2003). An AMSU footprint encompasses a 3x3 array of AIRS footprints. To be able to compare our retrieval with the operational product we apply our regression coefficients to the operational L2 cloud-cleared radiance product to obtain the

same spatial resolution. We demonstrate this type of evaluation for temperature (daytime granule 7, 09-02-2003) and for humidity (daytime granule 175, 09-02-2003) in Fig.7. Both variables are shown over ocean at ~850 mbar for the operational product (left), IMAPP AIRS retrieval (middle) and the ECMWF analysis (right). Note, only pixels with successful and validated cloud-clearing and retrieval performance are displayed. Operational and IMAPP AIRS retrievals have similar temperature fields, and the verification against the analysis suggests that good regression retrieval estimates are achievable. The main features in the humidity field as depicted in the analysis (Fig. 7, bottom right) are well identified by both retrievals, but the absolute values differ to some extent. For the case shown we find that the IMAPP AIRS regression retrieval (middle panel) is quite successful in describing the moist regions quantitatively. Nevertheless, the differences found in this study make further comparison and investigation necessary.
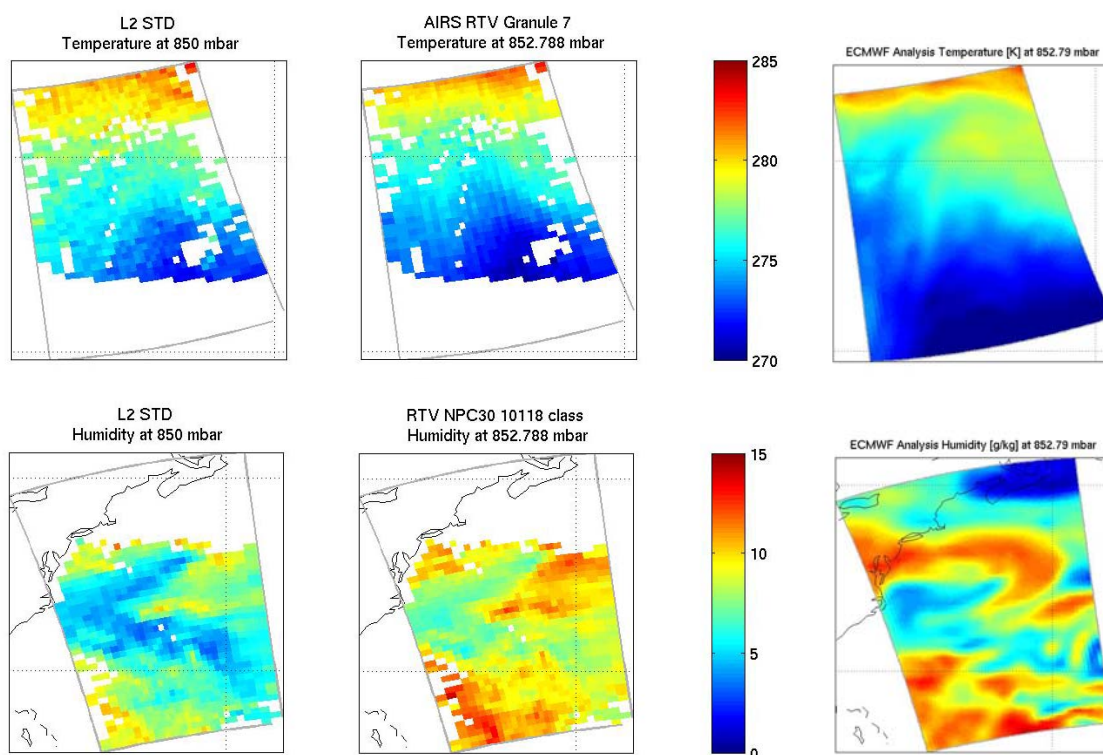

Fig. 7: L2 operational retrieval (left), AIRS IMAPP retrieval (middle) and ECMWF analysis (right) for temperature (top) and humidity (bottom) at 850 mbar.

Radiosonde observations provide an important opportunity to investigate retrieval performance, although temporal and spatial mismatch between radiosonde and satellite measurements have to be kept in mind when comparing the observed radiosonde profile with the retrieved profiles. In Fig. 8 one example of a single-profile retrieval compared with the radiosonde profile is presented. The upper panels display (from left to right) the temperature profiles, the corresponding differences of RAOB minus the retrieved profile profile, humidity profiles and their differences. Retrieved profiles include GOES, operational AIRS and IMAPP AIRS retrieval. The associated brightness temperature residuals are shown in the lower panels. Both retrieval results and residuals are comparable. Current efforts include retrieval validation against a large number of co-located radiosondes to quantify retrieval performance in a statistical sense.
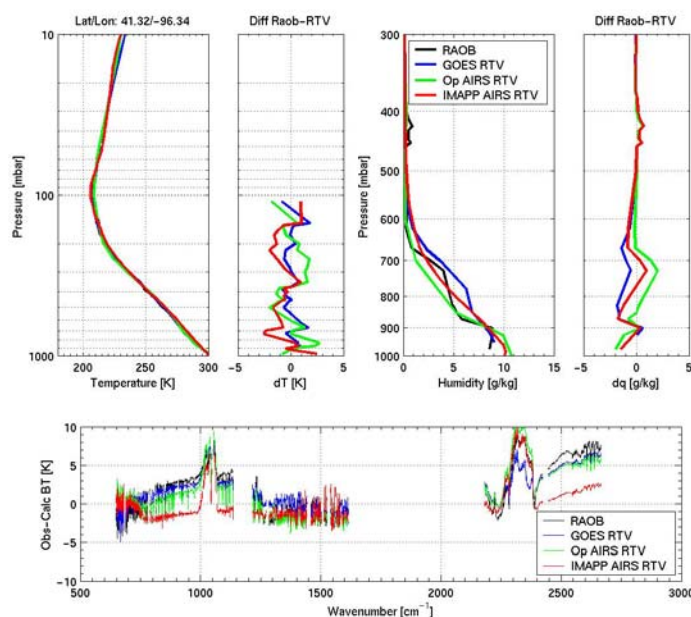
Fig. 8: Top: Profiles of temperature, temperature difference (radiosonde profile minus retrieval), humidity, and humidity difference for radiosonde (black), retrieved profiles of GOES (blue), operational AIRS (green) and IMAPP AIRS (red). Bottom: associated BT residuals in K.

Summary and Conclusions

We have presented algorithm and results of the pre-release version of the IMAPP AIRS regression retrieval. Results obtained so far promise acceptable retrieval performance in real-time. On-going efforts to improve these results include preparation of the training database, investigation of eigenvector compression and noise-filtering techniques. Furthermore, we will continue to validate IMAPP AIRS sounding retrievals against radiosonde observations, analysis fields, operational retrieval products from AIRS and other instruments such as GOES and MODIS.

The main drawback of the regression retrieval method is the assumption of a linear relationship between the measured radiances and the atmospheric state. To address this problem a more sophisticated non-linear physical inversion that includes information about first guess (background), error statistics and weighting functions will be implemented. We expect significant improvements from the physical method in particular for humidity soundings. Nevertheless, the linear regression retrieval will continue to be indispensable in achieving the 1K/1km and 10%/1km requirements for tropospheric temperature and humidity, respectively.

Other future research efforts include the optimal use of AIRS cloud-cleared radiances and improved handling of cloud contamination within the single and multiple field-of-views retrieval. Significant effort will be made related to single field-of-view humidity retrieval to provide regional meso-scale real-time applications for international direct broadcast users.

Acknowledgements

References

Ackerman, S.A., Strabala, K.I., Menzel, W.P., Frey, R.A., Moeller, C.C., Gumley, L.E. 1998. Discriminating clear sky from clouds with MODIS. *J. Geophys. Research-Atmospheres,***103**, 32141-32157.

Hung-Lung, H., Antonelli, P. 2001. Application of principal component analysis to high-resolution infrared measurement compression and retrieval. *J. Appl. Met.*, **40**, 365-388.

Seemann, S.W., Li, J., Menzel, W.P., Gumley, L.E. 2003. Operational retrieval of atmospheric temperature, moisture, and ozone from MODIS infrared radiances. *J. Appl. Met.*, **42**, 1072-1091.

Schmit, T.J., Feltz, W.F., Menzel, W.P., Jung, J., Noel, A.P., Heil, J.N., Nelson, J.P., Wade, G.S. 2002. Validation and use of GOES sounder moisture information. *Wea. Forecasting*, **17**, 139-154.

Smith, W.L., Woolf, H.M. 1976. The use of eigenvectors of statistical covariance matrices for interpreting satellite sounding radiometer observations. *J. Atmos. Sci.*, **33**, 1127-1140.

Strow, L.L., Hannon, S.E., De Souza-Machado, S., Motteler, H.E., Tobin, D. 2003. An overview of the AIRS radiative transfer model. *IEEE Trans. Geosci. Remote Sensing*, **41**, 303-313.

Susskind, J., Barnet, C.D., Blaisdell, J.M. 2003. Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds. *IEEE Trans. Geosci. Remote Sensing*, **41**, 390-409.

Proceedings
of the
Thirteenth International
TOVS Study Conference
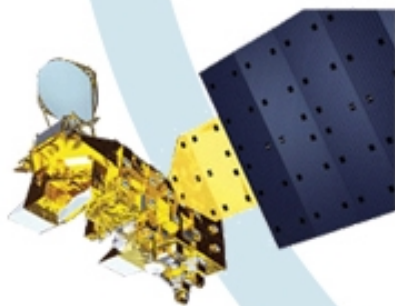
**Sainte-Adèle, Québec, Canada
29 October – 4 November 2003**