Retrieval validation during the European Aqua Thermodynamic Experiment

Daniel K. Zhou,[a]* William L. Smith,[b,c] Vincenzo Cuomo,[d] Jonathan P. Taylor,[e†]
Christopher D. Barnet,[f] Paolo Di Girolamo,[g] Gelsomina Pappalardo,[d] Allen M. Larar,[a] Xu Liu,[a]
Stuart M. Newman,[e†] Clare Lee[e†] and Stephen A. Mango[h]

[a] *NASA Langley Research Center, Hampton, VA, USA*
[b] *Hampton University, Hampton, VA, USA*
[c] *University of Wisconsin-Madison, Madison, WI, USA*
[d] *IMAA-CNR, Tito Scalo (Potenza), Italy*
[e] *Met Office, Exeter, UK*
[f] *NOAA/NESDIS, Camp Springs, MD, USA*
[g] *DIFA, Università degli Studi della Basilicata, Potenza, Italy*
[h] *NPOESS Integrated Program Office, Silver Spring, MD, USA*

ABSTRACT: Atmospheric and surface thermodynamic parameters retrieved with advanced hyperspectral remote sensors aboard Earth observing satellites are critical to weather prediction and scientific research. The retrieval algorithms and retrieved parameters from satellite sounders must be validated to demonstrate the capability and accuracy of both observation and data processing systems. The European Aqua Thermodynamic Experiment (EAQUATE) was conducted not only for validation of the Atmospheric InfraRed Sounder on the Aqua satellite, but also for assessment of validation systems of both ground-based and aircraft-based instruments that will be used for other satellite systems, such as the Infrared Atmospheric Sounding Interferometer on the European MetOp satellite, the Cross-track Infrared Sounder from the National Polar-orbiting Operational Environmental Satellite System (NPOESS) Preparatory Project and the continuing series of NPOESS satellites. Detailed intercomparisons were conducted and presented using different retrieval methodologies: measurements from airborne ultraspectral Fourier transform spectrometers, aircraft *in situ* instruments, dedicated dropsondes and radiosondes, ground-based Raman lidar, as well as the European Centre for Medium-range Weather Forecasting modelled thermal structures. The results of this study not only illustrate the quality of the measurements and retrieval products, but also demonstrate the capability of the validation systems put in place to validate current and future hyperspectral sounding instruments and their scientific products. Copyright © 2007 Royal Meteorological Society

1. Introduction

The Earth Observing System (EOS) Aqua satellite with the Atmospheric InfraRed Sounder (AIRS) on board was launched on 4 May 2002 (Pagano *et al.*, 2003; Aumann *et al.*, 2003a; Chahine *et al.*, 2006). The AIRS sounding goals are a temperature of 1K r.m.s. per 1 km layer average and 15% r.m.s. of moisture per 2 km layer average; the specifications are found elsewhere (Susskind *et al.*, 2006). A great deal of calibration and

validation activity has been done to understand the instrument and to quantify its data products (Aumann *et al.*, 2003a; Divakarla *et al.*, 2006; Tobin *et al.*, 2006a). The AIRS Science Team has developed a retrieval scheme capable of producing atmospheric and surface thermodynamics parameters. Validation of retrieved temperature and moisture profiles has been made through comparisons with the European Centre for Medium-range Weather Forecasting (ECMWF) model analysis and with matching radiosonde measurements (Divakarla *et al.*, 2006; Susskind *et al.*, 2006; Tobin *et al.*, 2006a). More information about the ECMWF can be obtained from the web site http://www.ecmwf.int/. In the meantime, the National Polar-orbiting Operational Environmental Satellite System (NPOESS) Airborne Sounder Testbed – Interferometer (NAST-I) and the Scanning-High resolution Interferometer Sounder (S-HIS) both flown on high altitude aircraft such as the ER-2 and Proteus (Revercomb *et al.*, 1998), have gone through

* Correspondence to: Daniel K. Zhou, Mail stop 401A, NASA Langley Research Center, Hampton, VA 23681, USA.
E-mail: Daniel.k.zhou@nasa.gov
† The contributions of J. P. Taylor, S. M. Newman and C. Lee were prepared as part of their official duties as employees of the UK Government. It is published with the permission of the Controller of Her Majesty's Stationery Office and the Queen's Printer for Scotland. The contributions of D.K.Zhou, A.M. Larar, X. Liu, C.D. Barnet, and S.A. Mango were prepared as part of their official duties as United States Federal Government employees.

numerous field campaigns (Tobin *et al.*, 2006b). These instruments provide experimental observations needed for finalizing specifications, testing proposed designs, and developing data-processing algorithms for the Cross-track Infrared Sounder (CrIS), which will fly on NPOESS satellites. Detailed descriptions of NAST-I instrumentation, data processing methodologies, and data products can be found elsewhere (Cousins and Smith, 1997; Smith *et al.*, 1999; Larar *et al.*, 2002; Zhou *et al.*, 2002; Smith *et al.*, 2005a). Selected AIRS radiance datasets have been put through testing using the NAST-Team retrieval algorithm since AIRS data first became available in July 2002 (Smith *et al.*, 2004). Through the process of these validation and evaluation activities, the AIRS retrieval algorithm has improved greatly over the past few years and now produces more accurate retrievals.

Aircraft instrumentation flown under the orbital pass of Aqua has been providing both radiance and retrieval evaluation (Gunshor *et al.*, 2003; Smith *et al.*, 2004; Tobin *et al.*, 2006b). An international experiment, European Aqua Thermodynamic Experiment (EAQUATE), performed in both Italy and the United Kingdom during September 2004, demonstrated certain ground-based and airborne systems useful for validating hyperspectral sounding observations from satellites throughout this decade and the next. The focus of this initial experiment was primarily placed on the validation of the AIRS instrument data from the EOS Aqua satellite (Cuomo *et al.*, 2005; Smith *et al.*, 2005b; Taylor *et al.*, 2007). A great deal of effort has been given to the data collected during EAQUATE. This paper reports on validation results of the thermodynamic parameters or the atmospheric temperature and moisture profiles. Intercomparison efforts have been made using four different perspectives:

(1) retrieval algorithm intercomparison through retrieval products,
(2) retrieval intercomparison between two instruments (i.e. AIRS and NAST-I) using the same retrieval algorithm,
(3) profile retrieval validation through comparisons with radiosonde, dropsonde, and Raman lidar measurements, and

(4) atmospheric structure intercomparison with that displayed by the ECMWF model analysis.

It is worth pointing out that the ECMWF model analysis is partially influenced by multi-sensor observations and/or assimilated instrumental data including AIRS and radiosondes. A conclusion is made by drawing a comparison between the various results of this validation activity and illustrating the quality of the measurements and retrieval products as well as the capability of the validation systems used.

2. Retrieval algorithms

In this study, a major objective is to validate retrieval algorithms applied to AIRS data. It is equally important to validate retrievals as well as the retrieval algorithm producing the geophysical products. Two retrieval algorithms are used here for intercomparison; they are described in detail elsewhere (Zhou *et al.*, 2002; Susskind *et al.*, 2003; Susskind *et al.*, 2006). However, the differences and similarities of these retrieval algorithms are addressed here in brief. Retrievals from NAST-I and AIRS using the same retrieval methodology (called the NAST-Team retrieval methodology, hereafter denoted as N-T) are compared with AIRS retrievals obtained by the latest version (version 4.0) of the AIRS Science Team retrieval methodology (hereafter denoted as A-T). Table I summarizes the major differences and similarities of these two retrieval algorithms.

The N-T retrieval methodology uses eigenvector regression to obtain initial profiles (Zhou *et al.*, 2001; Zhou *et al.*, 2002). The radiance eigenvectors are generated from radiances calculated using a forward radiative transfer model from a regional and seasonal climatology of radiosonde data. For NAST-I retrievals, 4424 spectral channels are used, whereas for AIRS 1526 spectral channels are used in N-T retrievals. The N-T retrieval uses only infrared radiance data from NAST-I or AIRS within a retrieval procedure that directly accounts for the influence of clouds on the observed radiances, but the retrieval validity is generally restricted to that data obtained from above cloud-top level. The regression result is used as

Table I. Major similarities and differences between A-T and N-T retrieval methodologies.

	A-T AIRS/AMSU	N-T AIRS	N-T NAST-I
Forward model	SARTA	SARTA	OSS
Regression retrieval methodology	EOF	EOF	EOF
Regression training profile	Global ECMWF	Regional radiosondes	Regional radiosondes
Regression training radiance	AIRS measured (clear)	SARTA simulated	OSS simulated
Regression channel number	1680	1526	4424
Physical retrieval methodology	Regularization	Regularization	Regularization
Physical retrieval procedure	Sequential	Simultaneous	Simultaneous
Physical retrieval channel number	156	575	697
Dealing with clouds	Cloud clearing with AMSU	Retrieval to cloud top	Retrieval to cloud top
Surface emissivity retrieval	Regression followed by physical	EOF regression	EOF regression

a first guess to a matrix inverse solution of the radiative transfer equation. In the physical matrix inverse retrieval step, the N-T uses an iterative simultaneous matrix inverse solution for all variables based on a selection of 697 NAST-I channels and 575 AIRS channels for the NAST-I and AIRS physical retrievals, respectively. Surface and cloud spectral emissivity is determined by EOF (empirical orthogonal function) regression and used in the matrix inverse solution. The N-T retrieval algorithm is applied to both AIRS and NAST-I data to minimize the impact of retrieval algorithm differences on the retrieval products. However, it should be noted that the forward radiative transfer models used differ in that the optimal spectral sampling (OSS) fast molecular radiative transfer model is used for the NAST-I retrieval (Liu *et al.*, 2003; Moncet *et al.*, 2003) while the Stand-alone AIRS Radiative Transfer Algorithm (SARTA) is used for the AIRS retrieval (Strow *et al.*, 2003).

The A-T retrieval methodology uses eigenvector regression to obtain initial profiles as well (Goldberg *et al.*, 2003; Susskind *et al.*, 2003; Susskind *et al.*, 2006). The A-T regression is based on a global database of profiles extracted from ECMWF analyses colocated with actual AIRS cloud-cleared radiances using 1680 AIRS spectral channels. In the case of the A-T retrievals, microwave data from the Advanced Microwave Sounding Unit (AMSU) instrument (Aumann *et al.*, 2003a) aboard the Aqua satellite are used to clear the radiance data of clouds and provide additional sounding radiance information for retrieval in a cloudy atmosphere. The regression coefficients, relating the atmospheric state variables to the radiance eigenvector amplitudes, are generated. In the physical retrieval, the A-T uses a sequential approach and a total of 156 AIRS spectral channels in which 65 spectral channels are used for temperature, 42 for water vapour, 26 for ozone, and 23 for surface temperature. The retrieval of surface emissivity is specified using synthetic radiance regression relationships similar to the N-T retrieval followed by a physical approach (Susskind *et al.*, 2003). The SARTA is used for the A-T AIRS retrieval.

3. Validation on thermodynamics products

There were two phases of the EAQUATE campaign: one was performed in Italy and the other in the UK. The Proteus aircraft performed four flights passing over the ground station (Potenza) where numerous ground-based instruments were operational (Cuomo *et al.*, 2005). However, the only Potenza ground-based measurements used in this work were VaisalA-Type radiosondes and two Raman lidar systems (Pappalardo *et al.*, 2003; Cuomo *et al.*, 2004; Di Girolamo *et al.*, 2004; Di Girolamo *et al.*, 2006). One Raman lidar station (labelled Lidar 1) is located at Potenza (40°39′N latitude, 15°48′E longitude, 730 m above sea level), while another Raman lidar station (labelled Lidar 2) and a radiosonde station are

located at Tito Scalo, Potenza (40°36′N latitude, 15°44′E longitude, 760 m above sea level). Lidar 1 measures both water vapour and temperature, while Lidar 2 measures only water vapour. For this latter system, relative humidity (RH) is computed with a coincident radiosonde temperature. Two Proteus flights dedicated to AIRS validation and flown over Potenza were scheduled precisely at the time when the Aqua satellite was also passing over Potenza during the nights of 7 and 9 September 2004. During the second phase over the UK, the Proteus aircraft, together with the UK Facility for Airborne Atmospheric Measurements (FAAM) BAE 146 aircraft (Taylor *et al.*, 2007), made two flights dedicated to AIRS validation during the days of both 14 and 18 September 2004. Numerous *in situ* sensors and remote-sensing instruments flew on the BAE 146 aircraft, and dropsondes were released from the BAE 146 during the experiments. The wealth of data from both ground-based and aircraft-sensing instruments collected during this experiment was meant to accurately capture the atmospheric state observed by the satellite instruments during the same period. These data have been used for detailed investigation, evaluation, and validation in this study. However, with such an enormous amount of data collected, only a fraction of data is presented herein to support the concluding results of this work.

3.1. 7–8 September over Potenza, Italy

The Aqua satellite passed over Potenza at 0108 UTC on 8 September 2004. The Proteus aircraft passed over a few times that same day with one overpass, made at 0110 UTC, being close to the time of the Aqua satellite overpass. Radiosondes were launched from Potenza at 2300 UTC on 7 September and 0200 UTC on 8 September. Both Raman lidar systems were operational. Raman lidar data were processed using a 10-minute integration time and variable vertical averaging in order to reach a higher altitude for retrieval validation as Aqua and Proteus were passing over. During that time, the sky was generally clear with a few scattered clouds. Figure 1(a) plots the A-T derived surface skin temperature image of the AIRS granule with the Proteus flight track. The two Raman lidar stations and one radiosonde station at Potenza are marked by the triangle. Because the observation times between these radiosondes and AIRS are not quite the same, ECMWF model analysis (with a 1° horizontal grid and 6-hour temporal resolution) is linearly interpolated to the AIRS/AMSU field of view (FOV) and observation time. ECMWF model analysis is plotted as an additional reference. Detailed side-by-side intercomparison is shown presenting a slight difference between the retrieval algorithms. The temperature, water vapour mixing ratio, and RH profiles from the Potenza area are shown in Figure 1(b) to (d). AIRS retrievals from both A-T and N-T retrieval algorithms use the same cloud-cleared radiance produced by the AIRS team. The A-T cloud clearing procedure depends on AMSU measurements. Therefore, the AIRS spatial resolution has to be

degraded to the AMSU spatial resolution, about 3×3 AIRS single FOV and having a diameter of 45 km at nadir. All NAST-I retrievals from NAST-I single FOV falling into AIRS 3×3 FOV (or AMSU FOV) are averaged, the time being when Proteus flew over Potenza at about 0200 UTC on 8 September. The ECMWF model analyses were interpolated to AIRS FOV location and time. Two Potenza radiosonde observations ($\sim$2300 UTC and $\sim$0200 UTC) were averaged and plotted in Figure 1 for intercomparison. Raman lidar data for the Aqua satellite overpass at 0108 UTC and for the Proteus aircraft overpass at 0215 UTC were extensively processed; the averaged profiles of these two overpasses are plotted in Figure 1. The profile differences (deviation from the A-T AIRS retrieval) are shown in the figures. Radiosonde and Raman lidar data are linearly interpolated into the AIRS and/or NAST retrieval grid. Differences between N-T NAST and N-T AIRS retrievals are expected from differences in NAST and AIRS instrument FOV, spectral resolution, and noise level. NAST-I retrievals for this case are matched more closely to the measurements of radiosonde, Raman lidar, and the analysis of the ECMWF model than the AIRS retrievals. As for the comparison of the water vapour mixing ratio and RH profiles, two Raman lidar profiles indicate some difference, especially in the region near 3 km where a few scattered clouds were observed. A large difference between Raman lidar and other profiles in that region is due to the difference in vertical and horizontal resolution; the Raman lidar has a higher signal-to-noise ratio and the vertical resolution is about 15 to 30 m. The AIRS water vapour mixing ratio and RH profiles retrieved with A-T and N-T algorithms are in good agreement, i.e. within instrument sounding goals (Smith *et al*, 2005a; Susskind *et al*., 2006). In the region near 9 km, the radiosonde indicates a lower humidity than that indicated by all other sensors. The NAST-I retrieval compares well with the radiosonde in the altitudes below 8 km. The NAST-I and N-T AIRS retrievals agree well except in the region below 3 km, which could be a result of the difference in instrument spatial resolution. After carefully considering the particulars addressed above, as a single profile comparison for this case, the outcome of the intercomparison is consistent with the retrievals being within the retrieval accuracy claimed for the AIRS and NAST-I instruments.

The dataset of a whole AIRS granule, shown in Figure 1(a), is used for intercomparison between two different retrieval algorithms and the thermal structure of ECMWF model analysis. Thermal structure covering a relatively large geophysical location of atmosphere is revealed and used to assess how well the measurements and retrieval systems can capture the features of atmospheric variation. The cross-sections of temperature and RH deviations from the means along the Aqua ground track (i.e. AIRS nadir viewing) are illustrated in Figure 2(a) to (f). The cross-section of the temperature deviation from the mean reveals the details of temperature spatial variation. The general patterns (i.e. the
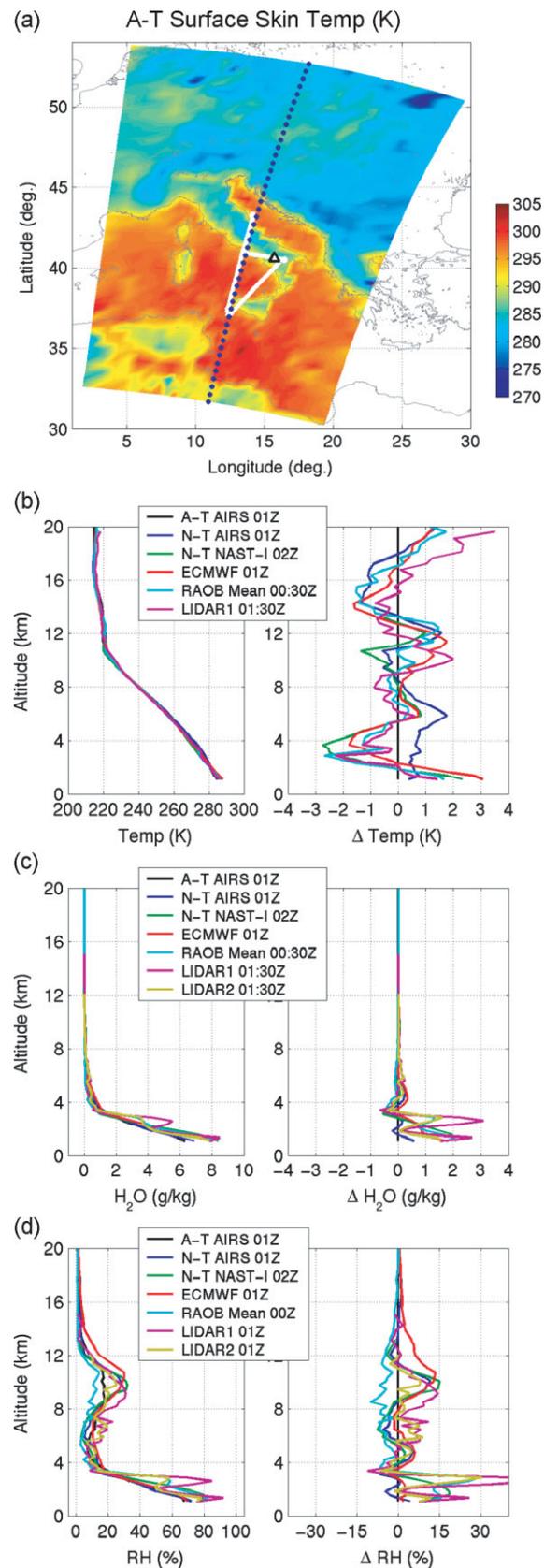

Figure 1. (a) A-T retrieved surface skin temperature (K) from one AIRS granule dataset. Potenza is indicated by a triangle, the Proteus flight track is plotted in white, and the blue dots represent AIRS nadir views. Intercomparison of (b) temperature, (c) water vapour mixing ratio, and (d) RH profiles from Potenza during the night of 7 September 2004.

spatial features) shown in the ECMWF model analysis compare favourably with AIRS retrievals. However, as shown in Figure 2(g) and (h), the profile discrepancy between the ECMWF and AIRS retrievals can be much greater than that which is required for AIRS retrieval accuracy or sounding goals (Susskind *et al.* 2006), especially for moisture up to RH of 50% or more, at some locations (Figure 2(h)). Some of the shown differences might be due to the different vertical and horizontal resolutions. Detailed thermodynamic structure of the atmosphere is captured by AIRS measurements in both A-T and N-T retrievals. As shown in Figure 2, there is a correspondence between the spatial features revealed by both temperature and RH cross-sections despite the differences between the two retrieval algorithms. It is noted, however, that slightly larger spatial gradients result with

the N-T algorithm. As shown in Figure 2(i) and (j), the cross-section of retrieval difference between N-T and A-T reveals a small, but detailed structure. A standard deviation (STD) and a mean derived from the cross-section of retrieval difference are plotted in Figure 2. For temperature, the standard deviation error (STDE) lies in the range 0.5–1.0 K and the mean difference (i.e. bias) is within ± 1.0 K shown in Figure 2(g). For RH, the standard deviation profile is less than 10% at altitudes above 3 km, but it exceeds 10% in the terrestrial boundary layer (TBL). The mean difference profile is within $\pm 8\%$. The statistical results are within the requirements of this type of sounder (Smith *et al.*, 2005a; Susskind *et al.*, 2006). These small differences can be explained by the different vertical resolutions achieved by the different retrieval algorithms, which are also manifested in
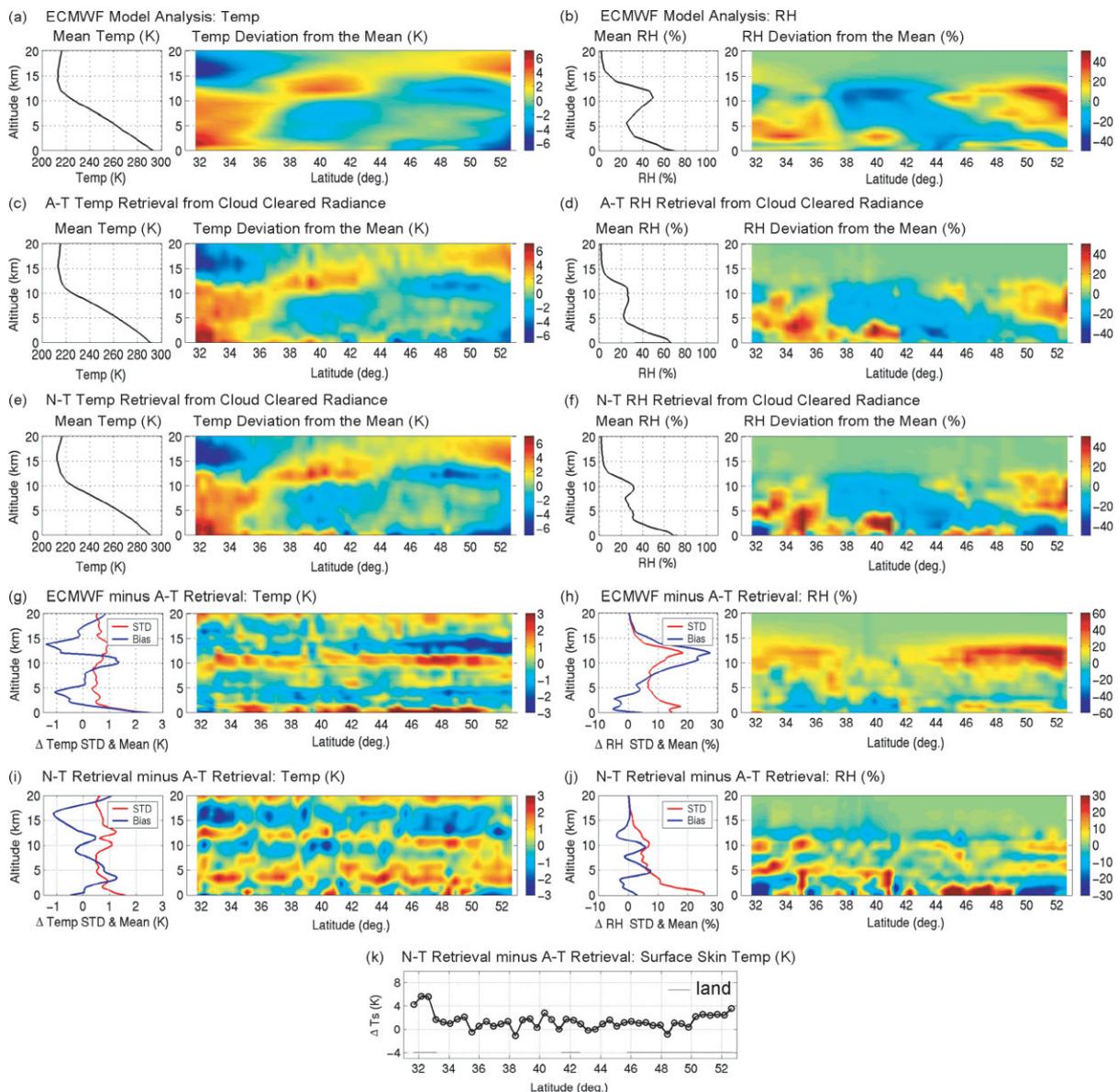

Figure 2. Temperature and RH cross-sections of AIRS granule at nadir, during the night of 7 September 2004, as indicated by blue dots in Figure 1(a). (a, b) ECMWF model analyses, (c, d) A-T retrievals, and (e, f) N-T retrievals. (g) and (h) show the cross-sections of temperature and RH differences between A-T retrievals and ECMWF. (i) and (j) show the cross-sections of temperature and RH differences between A-T and N-T retrievals. (k) shows the surface skin temperature difference between A-T and N-T, with the bars indicating measurements over land.

the profile validation over Potenza (Figure 1). A relatively large difference in RH near the surface, found mostly in observations over land, is possibly a result of the difference in the manner in which surface emissivity is handled by the two different retrieval schemes. Other factors, such as the A-T cloud-clearing procedure and the regression training dataset used for retrieval, could also be contributors. A relatively large difference in surface skin temperature, shown in Figure 2(k), could be initiated by an error in the land surface emissivity of the retrieval. In a related study, Zhou *et al.* (2006) found that the error in the surface emissivity (or skin temperature) resulted in a large moisture error near the surface, compensating for land surface uncertainty. For example, a large skin temperature difference existing in various points over land (e.g. latitudes less than 33 °N and greater than 50 °N in Figure 2(k)) is associated with the large RH difference in the TBL as shown in Figure 2(j). As seen in the figure, a large skin temperature difference between two retrievals occurs in the vicinity of the Sahara Desert (in this case, around

33 °N or lower). In that area, the surface emissivity retrieved from the two retrieval schemes is also quite different (not shown). In general, A-T retrieved skin temperature is cooler than N-T, while A-T emissivity is higher than N-T. Accurately retrieving land surface properties is still a challenge, and further studies are needed to improve retrieval methodology for land surface conditions.

Intercomparison between AIRS and NAST-I has been conducted using the N-T retrieval algorithm with AIRS original single-FOV data. NAST-I retrievals were spatially degraded to AIRS single-FOV footprints for intercomparison. ECMWF model analyses were also interpolated to AIRS single-FOV footprints. Figure 3(a) and (b) show the effective skin temperature retrieved by AIRS and NAST-I with a full spatial resolution of ∼15 km and ∼2 km, respectively. Mean temperature and RH profiles of the section (indicated by the open circles) along with a VaisalA-Type radiosonde and two Raman lidar observations from Potenza are plotted in Figure 3(c)
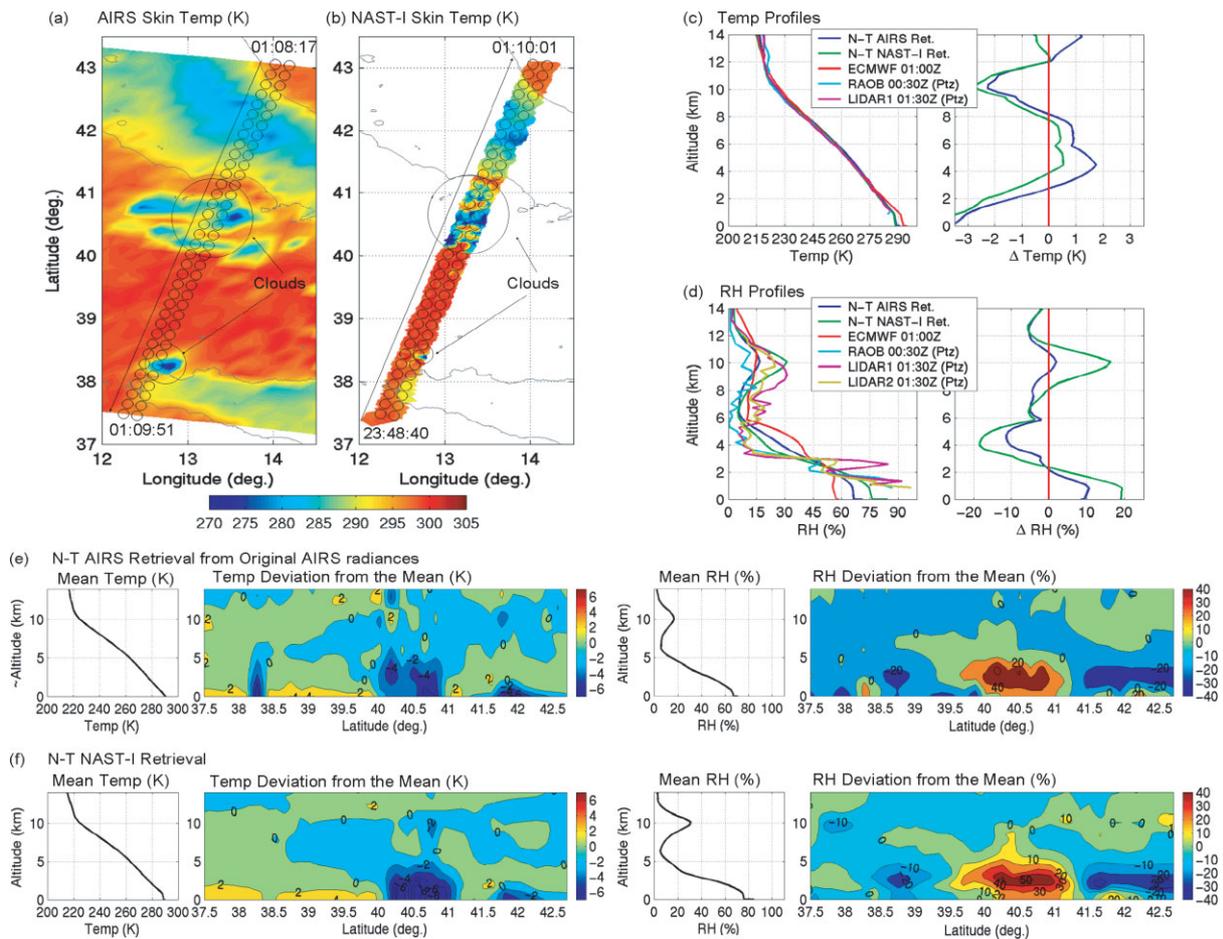

Figure 3. Effective surface skin temperature retrieved from (a) AIRS and (b) NAST-I, during the night of 7 September 2004. The circles are AIRS single FOVs within NAST-I ground swath, and lines with arrows indicate the Aqua and Proteus flight directions with associated times. The small circles represent AIRS single FOV within the NAST-I ground-track swath width, and the cloudy regions are indicated by larger open circles. Section mean profiles of (c) temperature and (d) RH using AIRS and NAST-I data processed through N-T retrieval scheme; radiosonde and Raman lidar profiles from Potenza, and section mean ECMWF model analysis are plotted as references. (e) AIRS and (f) NAST-I cross-sections of temperature and RH deviation from the means.

and (d); the deviations of NAST-I and AIRS cross-section means from the ECMWF cross-section mean are also plotted. It is worth pointing out that the averaged retrieval profile with a reduced vertical resolution of a single FOV retrieval has a lower vertical resolution while radiosonde and Raman lidar observations have a much higher vertical resolution, and a potentially dry bias in the humidity measurement from a Vaisala-type radiosonde at altitudes above ∼8 km has been noticed here as well as in other validations.

The cross-sections of temperature and RH deviation from the means as shown in Figure 3(e) and (f) reveals that very similar atmospheric spatial structures were retrieved from both AIRS and NAST-I observations, indicating a correspondence between the spatial features revealed in both temperature and RH cross-sections. Despite the cloudy regions (e.g. ∼38.2 °N and 40–41 °N), the atmospheric features are captured by both AIRS and NAST-I sounders as well as the ECMWF modelled structure. Small differences between AIRS and NAST-I are noticeable and are probably due to the observation time differences between AIRS and NAST-I. The fine vertical structures (i.e. resolution) of the retrieved profiles are partially due to instrumentation differences such as the spectral resolution and instrument noise, and these could cause a difference in the retrieved profiles. Overall, the retrievals from the two different sounders compare favourably with each other and show correlated patterns, which implies that both instruments are well calibrated and that the two radiative transfer models used are accurate.

3.2.   9–10 September over Potenza, Italy

The second Aqua validation flight over Italy was conducted during the night of 9 September 2004. Coincident profile comparisons over the Potenza ground site, comparable to Figure 1, are shown in Figure 4. A dedicated Potenza radiosonde observation at approximately 0035 UTC is used. Raman lidar data were processed for the Aqua satellite overpass at 0035 UTC and for the Proteus aircraft overpass at 0053 UTC; the averaged profiles of these two datasets are plotted. The profile differences (deviation from the A-T AIRS retrieval) are also shown in Figure 4(b) to (d). It is noted that radiosonde and Raman lidar data are linearly interpolated into the AIRS and/or NAST retrieval grid. In the comparison of the RH profiles, a relatively large difference between Raman lidar and other profiles in the lower tropospheric region is contributed to the difference in vertical and horizontal resolution. AIRS RH profiles retrieved with A-T and N-T algorithms agree well (within the sounding goal) except in the region below 3 km, where more moisture is shown in the NAST-I and N-T AIRS retrievals. The radiosonde shows a lower humidity than that indicated by all other sensors in the region of 8–12 km. Both NAST-I and N-T AIRS retrievals agree within their sounding goals.
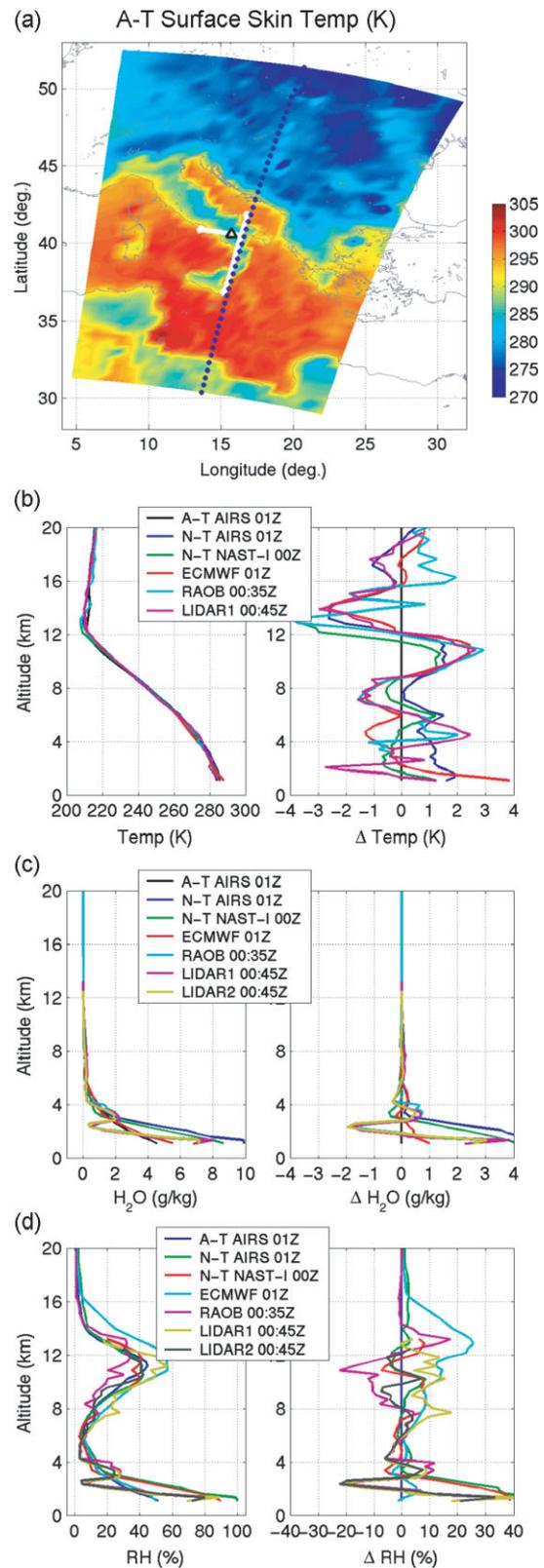

Figure 4. As Figure 1, but for 9 September 2004.

The intercomparison of AIRS retrieval amongst the A-T and N-T algorithms and the ECMWF modeled structure is summarized in Figure 5(a) to (f). The differences between the A-T retrievals and ECMWF model analysis are shown in Figure 5(g) and (h) for temperature and

RH profiles. The differences between the A-T and N-T retrievals are shown in Figure 5(i) to (k) for temperature, RH profiles, and surface skin temperature, respectively. Results similar to the previous case are obtained, indicating a correspondence between the spatial features revealed by both temperature and RH cross-sections. Again from this case, the spatial features shown in the ECMWF model analysis agree well with AIRS retrievals; however, the profile discrepancy between ECMWF and AIRS retrievals can be greater than the required AIRS retrieval accuracy, despite the difference in vertical resolution and spatial resolution.

Intercomparison between AIRS and NAST-I has also been performed for this case with the N-T retrieval algorithm using AIRS original single-FOV data. NAST-I retrievals were spatially degraded to AIRS single-FOV footprints. The atmospheric thermal variation was captured by both AIRS and NAST-I. Similar to the previous

case (Figure 3), a comparison between the atmospheric features captured by both AIRS and NAST-I sounders is reported in Figure 6. Small differences between AIRS and NAST-I retrievals are noticed:

(1) the observation time is different between AIRS and NAST-I;
(2) the fine vertical structures (i.e. resolution) of the retrieved profiles are partially due to the instrumentation differences, such as spectral resolution and instrument noise, which could cause a difference in the retrieved profiles; and
(3) in the NAST-I cross-section, a discontinuity occurred at a latitude of 40.5 °N where the NAST-I data has a time gap as noticed in Figure 6(b).

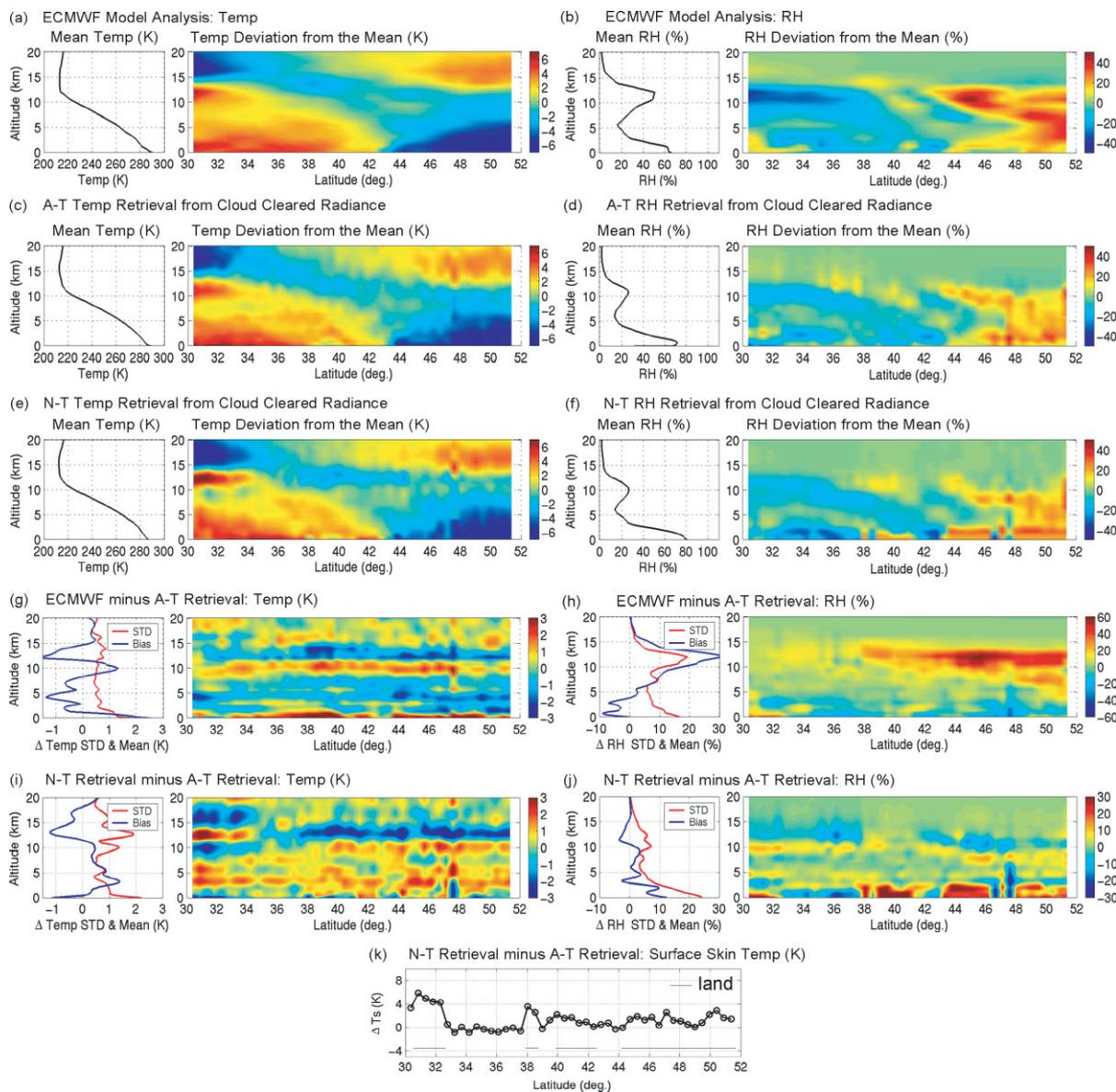Overall, the retrievals from the two different sounders compare favourably.

Figure 5. As Figure 2, but for the night of 9 September 2004 (blue dots in Figure 4(a)).

3.3. 14 and 18 September over Celtic Sea, UK

Similar analyses have been conducted for two AIRS granule data sets while Proteus and BAE 146 were flown beneath Aqua on 14 and 18 September 2004. Proteus and BAE 146 were flying at approximately the same location and time for these two dates. A series of five and nine dropsondes were released from BAE 146 in a close location and time on 14 and 18 September, respectively (Taylor *et al.*, 2007). Figure 7 shows the different atmospheric conditions from day to day captured by the measurements. Figure 7(a) and (e) show the A-T effective surface skin temperature retrieved from the AIRS cloud-cleared radiance. It is noticed that clouds affecting actual skin temperature retrievals were present in the sky, which is reflected in the data. However, the sky was almost clear at the locations where Proteus and BAE 146 aircraft were flying. The intercomparisons for temperature, water mixing ratio, and RH profiles were made with an AMSU FOV (or AIRS 3 × 3 Single FOV) located at 50.84 °N, 6.68 °W on 14 September, and at 50.99 °N, 6.85 °W on 18 September. Data from available nearby radiosonde stations (marked with open triangles in Figure 7(a) and (e)) were used in the profile comparison. The dropsondes from BAE 146 were dropped during the same period that NAST-I was taking the measurements. The approximate time of the

profile (or profile mean) is indicated in the figure legends. Overall, the agreement between these profiles (i.e. A-T AIRS, N-T AIRS, NAST-I, ECMWF, radiosonde, and dropsondes) is equivalent to that shown for Potenza, Italy. The difference is within what is expected for instrumental limitations and retrieval uncertainties. However, there is a vertical feature in the temperature profile of 14 September for the tropopause region (10 to 16 km). A somewhat larger variation in NAST-I and N-T AIRS retrievals is pronounced; this variation is also shown in the radiosonde profile but is not as pronounced as in the retrievals. A relatively large temperature discrepancy is noted in this region and it is also indicated, but less pronounced, by other datasets (e.g. shown in Figures 2 and 5). Other detailed evaluations between the different algorithms, different instrumental measurements, and ECMWF model analyses have been performed with these two AIRS granules. Results similar to those shown for the two Italian granules were obtained.

Because these two cases were made with the BAE 146 aircraft flying as low as 30 m above the sea, the surface properties, like the sea surface temperature (SST) and emissivity, are observed with the Airborne Research Interferometer Evaluation System (ARIES) on board (e.g. Newman *et al.*, 2005; Taylor *et al.*, 2007). NAST-I and ARIES observations within the AIRS FOV are used for
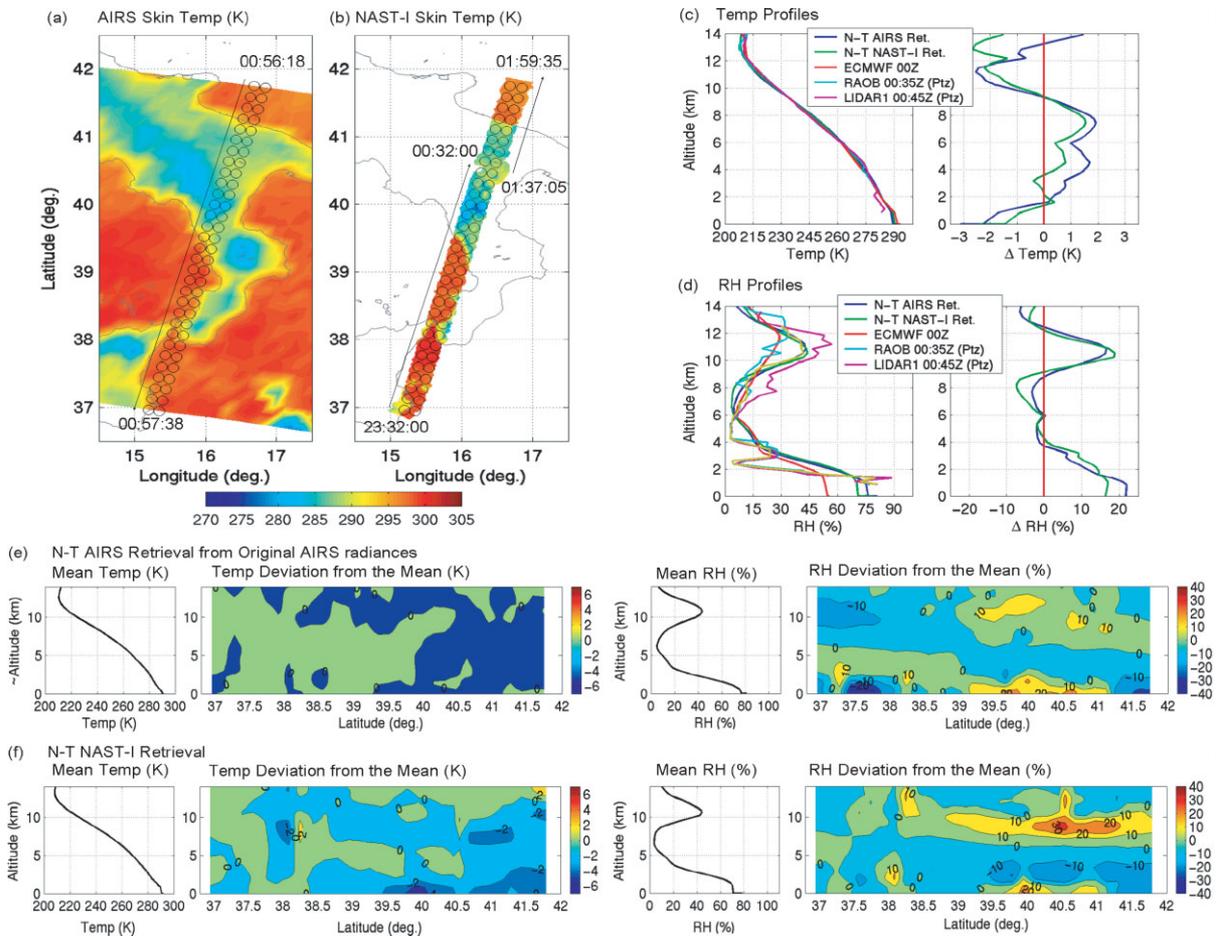

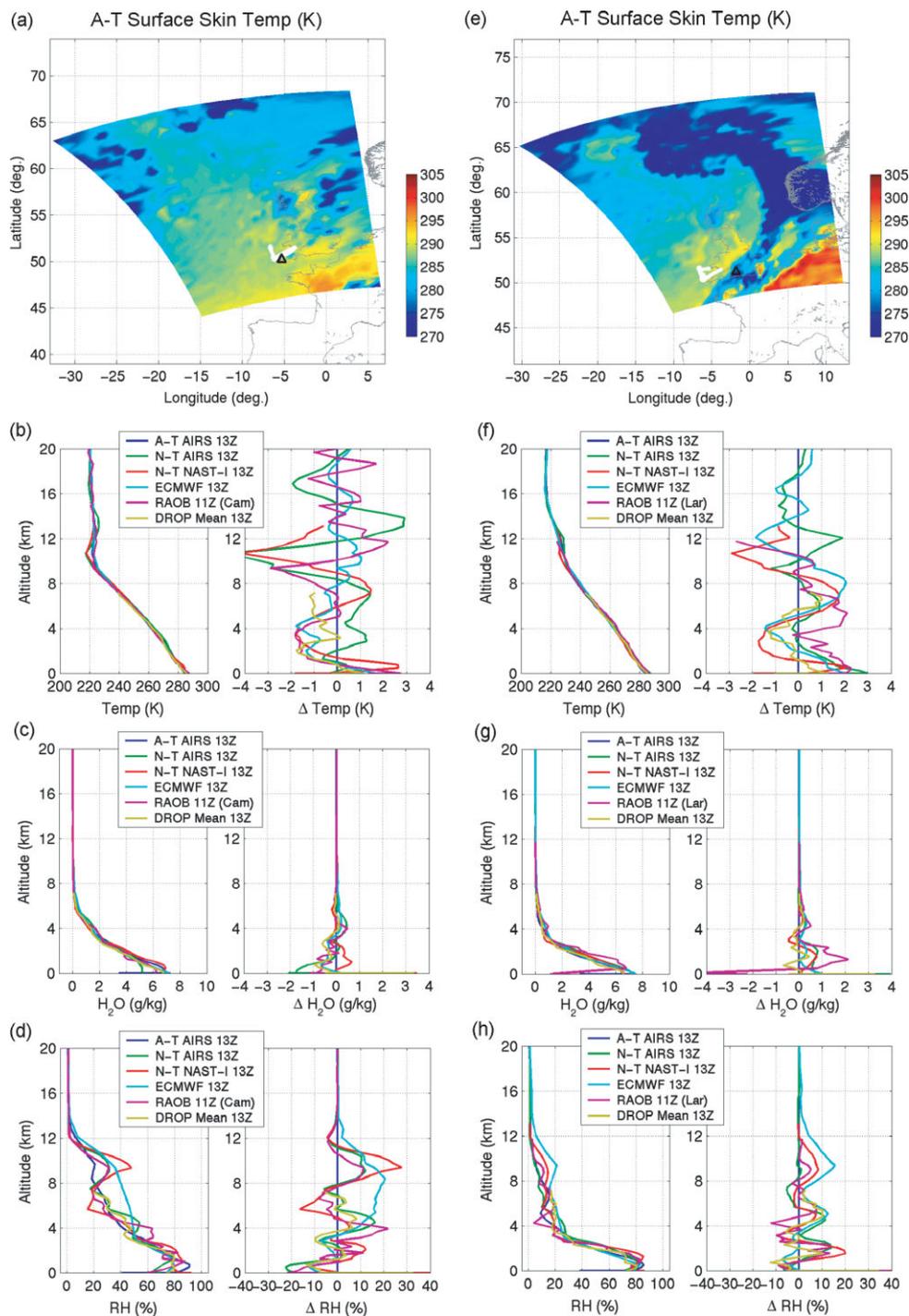Figure 6. As Figure 3, but for 9 September 2004.

Figure 7. (a) Effective skin temperatures of AIRS granule for 14 September 2004; the Camborne radiosonde station is indicated by a triangle, and the Proteus flight tracks are plotted in white. Profile intercomparisons are shown for (b) temperature, (c) water mixing ratio and (d) RH. (e) to (h) are as (a) to (d), but for 18 September 2004; the triangle in (e) shows the position of the Larkhill radiosonde station.

intercomparisons. The geophysical locations of observations are illustrated in Figure 8; the AIRS and ECMWF data are interpolated to the mean location of ARIES. The AIRS SST observations are contributed by a fairly large area in comparison with NAST-I and ARIES observations. AIRS SST was retrieved with A-T and N-T retrieval algorithms, and the SST means of both NAST-I and ARIES retrievals are listed in Table II along with ECMWF analyses. ARIES SST findings are a little higher

than those of AIRS and NAST-I retrievals by a mean of 0.37 K. Retrieved SSTs are lower than ECMWF SST, which may be related to the bulk temperatures reported by the floating buoys. The SSTs retrieved from AIRS, NAST-I, and ARIES radiance are expected to be lower than the sea surface bulk temperature due to evaporative cooling (Schluessel *et al*., 1987; Smith *et al*., 1996). A bias (normally within a few tenths of a degree) existing between infrared sensed and *in situ* measured SST

is the physical difference between sea surface skin temperature and *in situ* SST measured at some depth (Smith *et al.*, 1996). Results of AIRS SST from the two cases revealed here are consistent with what was reported by Aumann *et al.* (2003b). The associated surface emissivity spectra are plotted against laboratory-measured seawater emissivity (Salisbury and D'Aria, 1992) in Figure 8. As seen, the ARIES emissivity agrees very well with laboratory-measured seawater emissivity for both cases. The emissivity spectra derived from NAST-I and ARIES are nearly unchanged from day to day. Despite the difference of the AIRS viewing angle (i.e. satellite zenith angle) and wind speed from 14 September to 18 September, the seawater emissivity spectra derived from different instruments, platforms, viewing geometry, and retrieval algorithms are in fair agreement for these cases. Small differences, such as ARIES emissivity and

laboratory-measured seawater emissivity being slightly higher (i.e. ~0.006–0.010) than AIRS and NAST-I, are shown. The retrieved seawater emissivity is relatively good as a diverse surface database was used to handle different surface types in the retrieval process. A compensation for the emissivity difference may explain why the SST of AIRS and NAST-I are slightly lower (~0.35 K) than ARIES. Nevertheless, these differences are within the expectation of these instruments (i.e. AIRS and NAST-I) and their data processing procedures because atmospheric effects or contributions are included in the retrieval schemes.

4. Conclusions

The international experiment, EAQUATE, was successful in testing both ground-based and airborne systems for
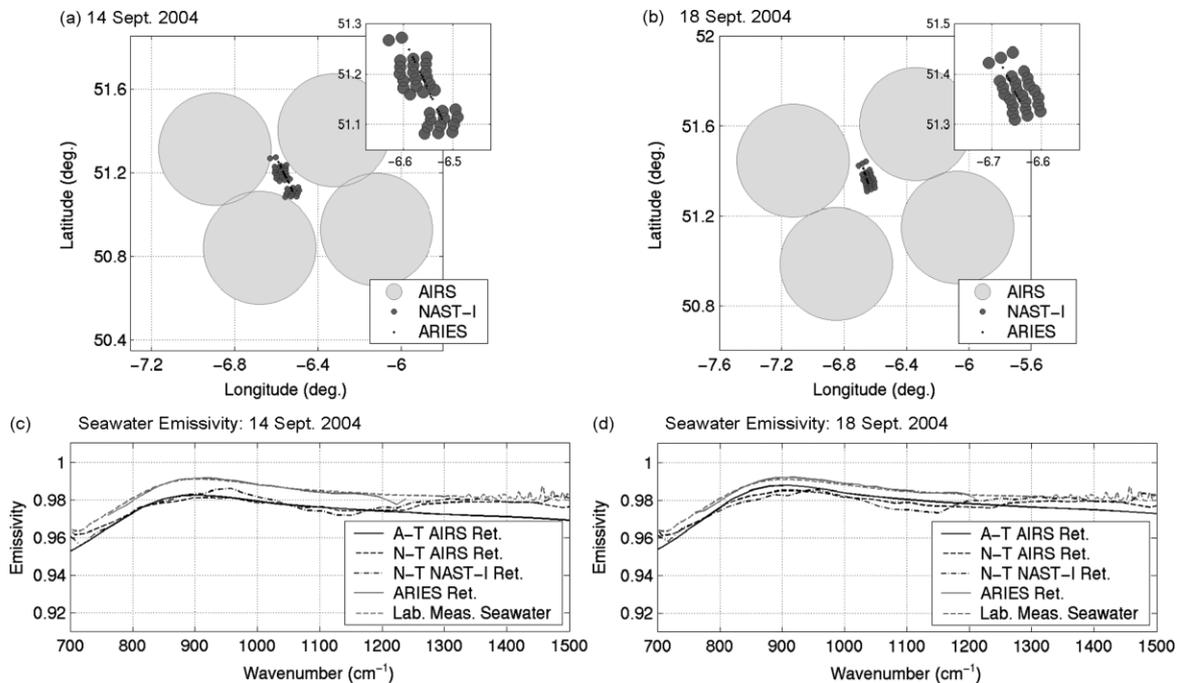

Figure 8. Geographical locations and approximate FOVs from AIRS, NAST-I, and ARIES used for intercomparison of surface properties for (a) 14 and (b) 18 September 2004. Large grey circles are AIRS 3 × 3 single FOVs equivalent to AMSU FOV; four AIRS/AMSU FOVs are used to interpolate to the value centred at ARIES observations. The footprints of NAST-I and ARIES are enlarged at top right. Seawater emissivity intercomparisons between AIRS retrievals from A-T and N-T, NAST-I retrieval, ARIES retrieval, and laboratory measurements are shown for (c) 14 and (d) 18 September 2004 (see text).

Table II. SST intercomparison.

	14 September 2004				18 September 2004			
	Lat. (°N)	Long. (°W)	SST (K)	ΔSST (K)*	Lat. (°N)	Long. (°W)	SST (K)	ΔSST (K)*
A-T AIRS/Aqua	51.18	6.56	287.91	−0.37	51.37	6.66	287.81	−0.35
N-T AIRS/Aqua	51.18	6.56	287.94	−0.34	51.37	6.66	287.82	−0.34
NAST-I/Proteus	51.17	6.56	287.96	−0.32	51.37	6.65	287.66	−0.50
ARIES/BAE 146	51.17	6.55	288.28	+0.00	51.37	6.66	288.16	+0.00
ECMWF	51.18	6.56	289.04	+0.76	51.37	6.66	288.61	+0.45

* SST minus ARIES-retrieved SST

validating hyperspectral satellite measurements. Experimental data were collected during the campaign. These data are used for AIRS data validation as demonstrated in this paper. This study demonstrates the need for both high-density *in situ* observations for intercomparison and high-altitude aircraft sounders like the NAST-I or S-HIS in order to provide instant radiance and retrieval validation. High-altitude aircraft remote sounders provide broad area coverage with high spatial resolution and continuous spectral coverage as are needed to validate satellite observations. It is worth pointing out that, since the retrieval problem with remote sensors such as AIRS and NAST-I is ill-posed, the spatial and temporal resolution needs to be considered for sounding profile validation if the sounding retrieval goals are to be achieved. The fine vertical structure shown by the radiosonde and lidar observations might not perfectly represent what was really measured or retrieved as an 'averaged' profile within a relatively larger FOV.

Retrieval validation studies have been conducted in a manner enabling the effects of retrieval algorithm accuracy and satellite measurement accuracy to be separated. A variety of independent validation systems including radiosondes, dropsondes, ground-based Raman lidar, and the ECMWF model analysis is used. Several conclusions can be made from this study.

(1) Very similar surface and atmospheric spatial structures, shown in Figures 3 and 6, were retrieved from AIRS and NAST-I observations.
(2) General retrieval product agreement is obtained for two different retrieval algorithms, one from the AIRS team and the other from the NAST team. The standard deviation of differences is under 1 K for temperature, less than 10% for RH above the TBL, and 10 to 25% for RH in the TBL (sections 3.1 and 3.2).
(3) Many validation cases with Proteus and Aqua overpasses above Potenza having dedicated radiosonde and Raman lidar observations reveal that there is a potentially dry bias in VaisalA-Type radiosonde observations in an altitude range of 8 to 11 km; this bias is revealed by the intercomparisons between radiosonde, AIRS retrievals, NAST-I retrievals, and two Raman lidar observations (sections 3.1 and 3.2).
(4) The spatial features shown in the ECMWF model analysis compare favourably with AIRS retrievals. However, the profile discrepancy between ECMWF and AIRS retrieval could be much greater than that required to validate AIRS retrieval accuracy, especially for RH up to 0.5 or more in some areas shown in Figures 2 and 5.
(5) The surface properties over water are retrieved within expected accuracy (0.4 K and 0.01 for SST and emissivity) and compared favourably with observations by all instruments from the variety of platforms. However, AIRS surface skin temperature retrieved from the two retrieval algorithms has an offset that is greatly pronounced over land (up to

6 K; Figures 2(k) and 5(k)). This surface skin temperature difference is due to the manner in which surface emissivity is treated by the two different algorithms. The retrieval uncertainties in the surface properties (i.e. emissivity and temperature) affect the accuracy of the thermodynamic profiles in the TBL.

The overall outcome indicates that the latest version (i.e. version 4.0) of the archived AIRS/AMSU retrievals is within the AIRS sounding goal, at least for clear-sky oceanic conditions. This experiment has demonstrated the use of ground-based and aircraft-based validation for current and future hyperspectral satellite sounders. This work has also established the need for validating retrieval algorithms to ensure the accuracy of the retrieval and that the same retrievals can be produced from alternative retrieval algorithms.

References

Aumann HH, Chahine MT, Gautier C, Goldberg MD, Kalnay E, McMillin LM, Revercomb H, Rosenkranz PW, Smith WL, Staelin DH, Strow L, Susskind J. 2003a. AIRS/AMSU/HSB on the Aqua mission: Design, science objective, data products, and processing systems. *IEEE Trans. Geosci. Remote Sensing* **41**: 253–264.
Aumann HH, Chahine MT, Barron D. 2003b. Sea surface temperature measurements with AIRS: RTG.SST comparison. *Proc. SPIE* **5151**: 252–261.
Chahine C, Pagano TS, Aumann HH, Atlas R, Barnet C, Blaisdell J, Chen L, Divakarla M, Fetzer EJ, Goldberg M, Gautier C, Granger S, Hannon S, Irion FW, Kakar R, Kalnay E, Lambrigtsen BH, Lee S-Y, Le Marshall J, McMillan WW, McMillin L, Olsen ET, Revercomb H, Rosenkranz P, Smith WL, Staelin D, Strow LL, Susskind J, Tobin D, Wolf W, Zhou L. 2006. AIRS: improving weather forecasting and providing new insights into climate. *Bull. Am. Meteorol. Soc.* **87**: 911–926.
Cuomo V, Amodeo A, Cornacchia C, Mona L, Pandolfi M, Pappalardo G. 2004. ENVISAT validation campaign at IMAA-CNR.

Proceedings of the Second Validation Workshop on the Atmospheric Chemistry Validation of ENVISAT (ACVE-2). **SP-532**: European Space Agency: Noordwijk.

Cuomo V., A. Amodeo PB. Antonelli A. Boselli A. Bozzo C. Cornacchia G. D'Amico M. Di Bisceglie F. Esposito P. Di Girolamo G. Grieco AM. Larar L. Leone F. Madonna T. Maestri R. Marchese G. Masiello G. Meoli L. Mona M. Pandolfi G. Pappalardo G. Pavese G. Pisani R. Restieri R. Rizzi F. Romano E. Rossi F. Rossi D. Sabatino C. Serio WL. Smith N. Spinelli D. Summa G. Todini D. Villacci X. Wang Zhou DK. 2005. The Italian phase of the EAQUATE measurement campaign. *Proc. SPIE* **5979**: 396–409.

Cousins D, Smith WL. 1997. National Polar-Orbiting Operational Environmental Satellite System (NPOESS) Airborne Sounder Testbed-Interferometer (NAST-I). *Proc. SPIE* **3127**: 323–331.

Divakarla MG, Barnet CD, Goldberg MD, McMillin LM, Maddy E, Wolf WW, Zhou L, Liu X. 2006. Validation of Atmospheric Infrared Sounder temperature and water vapor retrievals with matched radiosonde measurements and forecasts. *J. Geophys. Res.* **111**: D09S15, DOI:10.1029/2005JD006116.

Di Girolamo P, Marchese R, Whiteman DN, Demoz BB. 2004. Rotational Raman Lidar measurements of atmospheric temperature in the UV. *Geophys. Res. Lett.* **31**: L01106, DOI:10.1029/2003GL018342.

Di Girolamo P, Behrendt A, Wulfmeyer V. 2006. Pure rotational Raman lidar measurements of atmospheric temperature and extinction from space: Performance simulations. *App. Opt.* **45**: 2474–2494.

Gunshor MM, Tobin D, Schmit TJ, Menzel WP. 2003. 'First satellite intercalibration comparing high spectral resolution AIRS with operational geostationary imagers'. In preprints to 12th Conference on Satellite Meteorology and Oceangraphy, Longbeach, CA, 9–13 February 2003. American Meteorological Society: Boston.

Goldberg MD, Qu Y, McMillin LM, Wolf W, Zhou L, Divakarla M. 2003. AIRS near real-time products and algorithms in support of operational numerical weather prediction. *IEEE Trans. Geosci. Remote Sensing* **41**: 379–389.

Larar AM, Smith WL, Zhou DK. 2002. Spectral radiance validation studies using NAST-I and other independent measurement systems. *Proc. SPIE* **4485**: 81–90.

Liu X, Moncet J-L, Zhou DK, Smith WL. 2003. A fast and acurate frward mdel for NAST-I Instrument. In *Fourier Transform Spectroscopy and Optical Remote Sensing of Atmosphere.* Technical Digest Series, Optical Society of America: Washington DC.

Moncet JL, Liu X, Rieu-Isaacs H, Snell H, Zaccheo S, Lynch R, Eluszkiewicz J, He Y, Uymin G, Lietzke C, Hegarty J, Boukabara S, Lipton A, Pickle X. 2003. Algorithm theoretical basis document (ATBD) for the Cross-track Infrared Sounder (CrIS) environmental data records (EDR), v1.2.3. AER Document Number P882-TR-E-1.2.3-ATBD-03-01. http://eic.ipo.noaa.gov/IPOarchive/SCI/atbd/cris_atbd_03_09_01.pdf

Newman SM, Smith JA, Glew MD, Rogers SM, Taylor JP. 2005. Temperature and salinity dependence of sea surface emissivity in the thermal infrared. *Q. J. R. Meteorol. Soc.* **131**: 2539–2557.

Pagano TS, Aumann HH, Hagan DE, Overoye K. 2003. Prelaunch and in-flight radiometric calibration of the Atmospheric Infrared Sounder (AIRS). *IEEE Trans. Geosci. Remote Sensing* **41**: 265–273.

Pappalardo G, Amodeo A, Amoruso S, Mona L, Pandolfi M, Cuomo V. 2003. One year of tropospheric Lidar measurements of aerosol extinction and backscatter. *Ann. Geophys.* **46**: 401–413.

Revercomb HE, Walden VP, Tobin DC, Anderson J, Best FA, Ciganovich NC, Dedecker RG, Dirkx T, Ellington SC, Garcia RK, Herbsleb R, Knuteson RO, Laporte D, McRae D, Werner M. 1998. Recent results from two new aircraft-based Fourier-transform interferometers: The Scanning High-resolution Interferometer Sounder and the NPOESS Atmospheric Sounder Testbed Interferometer. Pp. 249–254 in Proceedings of 8th Internat. Workshop on Atmospheric Science from Space using Fourier Transform Spectrometry (ASSFTS), Toulouse, France, 16–18 November 1998.

Salisbury JW, D'Aria DM. 1992. Emissivity of terrestrial material in the 8–14 μm atmospheric window. *Remote Sens. Environ.* **42**: 83–106.

Schluessel P, Shin HY, Emery WJ, Grassl H. 1987. Comparison of satellite-derived sea surface temperatures with *in situ* skin measurements. *J. Geophys. Res.* **92**: 2895–2874.

Smith WL, Knuteson RO, Revercomb HE, Feltz W, Howell HB, Menzel WP, Nalli NR, Brown O, Brown J, Minnett P, McKeown W. 1996. Observations of the infrared radiative properties of the ocean – Implications for the measurement of sea surface temperature via satellite remote sensing. *Bull. Am. Meteorol. Soc.* **77**: 41–51.

Smith WL, Larar AM, Zhou DK, Sisko CA, Li J, Huang B, Howell HB, Revercomb HE, Cousins D, Gazarik MJ, Mooney D. 1999. NAST-I: Results from revolutionary aircraft sounding spectrometer. *Proc. SPIE* **3756**: 2–8.

Smith WL, Zhou DK, Revercomb HE, Huang HL, Antonelli P, Mango SA. 2004. Hyperspectral atmospheric sounding. *Proc. SPIE* **5235**: 389–396.

Smith WL, Zhou DK, Larar AM, Mango SA, Howell HB, Knuteson RO, Revercomb HE, Smith WL Jr. 2005a. The NPOESS Airborne Sounding Testbed Interferometer – Remotely sensed surface and atmospheric conditions during CLAMS. *J. Atmos. Sci.* **62**: 1117–1133.

Smith WL, Taylor JP, Cuomo V, Larar AM, Antonelli P, Mango SA, Serio C, Zhou DK, Pappalardo G. 2005b. EAQUATE – An international experiment for hyper-spectral atmospheric sounding validation. In *Fourier Transform Spectroscopy/Hyperspectral Imaging and Sounding of the Environment;* Topical Meeting on CD-ROM. JMA4. Optical Society of America: Washington DC.

Susskind J, Barnet CD, Blaisdell JM. 2003. Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in presence of clouds. *IEEE Trans. Geosci. Remote Sensing* **41**: 390–409.

Susskind J, Barnet CD, Blaisdell JM, Iredell L, Keita F, Kouvaris L, Molnar G, Chahine MT. 2006. Accuracy of geophysical parameters derived from Atmospheric Infrared Sounder/Advanced Microwave Sounding Unit as a function of fractional cloud cover. *J. Geophys. Res.* **111**: DOI:10.1029/2005JD006272.

Strow LL, Hannon SE, De Souza-Machado S, Motteler HE, Tobin D. 2003. An overview of the AIRS radiative transfer model. *EEE Trans. Geosci. Remote Sensing* **41**: 303–313.

Taylor JP, Smith WL, Cuomo V, Larar AM, Zhou DK, Serio C, Maestri T, Rizzi R, Newman S, Antonelli P, Mango SA, Di Girolamo P, Esposito F, Grieco G, Summa D, Restieri R, Masiello G, Romano F, Pappalardo G, Pavese G, Mona L, Amodeo A, Pisani G. 2007. EAQUATE – An international experiment for hyper-spectral atmospheric sounding validation. *Bull. Am. Meteorol. Soc.* In press.

Tobin DC, Revercomb HE, Knuteson RO, Lesht BM, Strow LL, Hannon SE, Feltz WF, Moy LA, Fetzer EJ, Cress TS. 2006a. Atmospheric Radiation Measurement site atmospheric state best estimates for Atmospheric Infrared Sounder temperature and water retrieval validation. *J. Geophys. Res.* **111**: D09S14, DOI:10.1029/2005JD006103.

Tobin DC, Revercomb HE, Knuteson RO, Best FA, Smith WL, Ciganovich NN, Dedecker RG, Dutcher S, Ellington SD, Garcia RK, Howell HB, LaPorte DD, Mango SA, Pagano TS, Taylor JK, van Delst P, Vinson KH, Werner MW. 2006b. Radiometric and spectral validation of Atmospheric Infrared Sounder observations with the aircraft-based Scanning High-Resolution Interferometer Sounder. *J. Geophys. Res.* **111**: D09S02, DOI:10.1029/2005JD006094.

Zhou DK, Smith WL, Larar AM. 2001. Surface temperature and emissivity from airborne measurements of IR radiance spectra. *Eos Trans. AGU* **82**(47).

Zhou DK, Smith WL, Li J, Howell HB, Cantwell GW, Larar AM, Knuteson RO, Tobin DC, Revercomb HE, Mango SA. 2002. Thermodynamic product retrieval methodology for NAST-I and validation. *App. Opt.* **41**: 6957–6967.

Zhou DK, Larar AM, Smith WL, Liu X. 2006. Surface emissivity effects on thermodynamic retrieval of IR spectral radiance. *Proc. SPIE* **6405**: 64051H.