June 2013

Validation of AIRS/AMSU - A Water Vapor and Temperature Data With In Situ Aircraft Observations From the Surface to UT/LS From 87°N–67°S

Minghui Diao
Princeton University, minghui.diao@sjsu.edu

Loayeh Jumbam
Sonoma Technology, Inc.

Justin Sheffield
Princeton University

Eric Wood
Princeton University

Mark Zondlo
Princeton University

Recommended Citation

Minghui Diao, Loayeh Jumbam, Justin Sheffield, Eric Wood, and Mark Zondlo. "Validation of AIRS/AMSU - A Water Vapor and Temperature Data With In Situ Aircraft Observations From the Surface to UT/LS From 87°N–67°S" *Journal of Geophysical Research: Atmospheres* (2013). https://doi.org/10.1002/jgrd.50483

Validation of AIRS/AMSU-A water vapor and temperature data with in situ aircraft observations from the surface to UT/LS from 87°N–67°S

Minghui Diao,[1,2] Loayeh Jumbam,[3] Justin Sheffield,[1] Eric F. Wood,[1] and Mark A. Zondlo[1,2]

Received 15 January 2013; revised 8 May 2013; accepted 9 May 2013; published 26 June 2013.

[1] Validation of the Atmospheric Infrared Sounder (AIRS)/Advanced Microwave Sounding Unit (AMSU-A) data set with in situ observations provides useful information on its application to climate and weather studies. However, different space/time averaging windows have been used in past studies, and questions remain on the variation of errors in space, such as between land/ocean and the Northern/Southern Hemispheres. In this study, in situ aircraft measurements of water vapor and temperature are compared with the AIRS/AMSU-A retrievals (Version 5 Level 2) from 87°N to 67°S and from the surface to the upper troposphere and lower stratosphere (UT/LS). By using a smaller comparison window (1 h and 22.5 km) than previous studies, we show that the absolute percentage difference of water vapor ($|d\mathrm{H_2O perc}|$) is ~20–60% and the absolute temperature difference ($|d\mathrm{Temp}|$) is ~1.0–2.5 K. The land retrievals show improvements versus Version 4 by ~5% in water vapor concentration and ~0.2 K in temperature at 200–800 mbar. The land (ocean) retrievals are colder and drier (warmer and moister) than the in situ observations in the boundary layer, warmer and drier (warmer and moister) at the UT/LS. No significant differences between hemispheres are noted. Overall, future comparisons are suggested to be done within 4 h and 100 km in order to keep the errors from window sizes within ~10%. To constrain the uncertainties in previous validation results, we show that every 22.5 km (or 1 h) increment in window sizes contributes to ~2% $|d\mathrm{H_2O perc}|$ and ~0.1 K $|d\mathrm{Temp}|$ increases.

Citation: Diao, M., L. Jumbam, J. Sheffield, E. F. Wood, and M. A. Zondlo (2013), Validation of AIRS/AMSU-A water vapor and temperature data with in situ aircraft observations from the surface to UT/LS from 87°N–67°S, *J. Geophys. Res. Atmos.*, *118*, 6816–6836, doi:10.1002/jgrd.50483.

1. Introduction

[2] Global measurements of water vapor and temperature are critical for assessing climate and weather models. The NASA Atmospheric Infrared Sounder (AIRS) and the Advanced Microwave Sounding Unit (AMSU-A) aboard the Aqua satellite provide global measurements of water vapor and temperature twice per day since May 2002. The AIRS/AMSU-A data sets also cover a wide range of vertical levels from the surface (1100 mbar) to the stratosphere (50 mbar for water vapor; 0.1 mbar for temperature). With global coverage and vertical sampling, the AIRS/AMSU-A observations provide a unique data set that has been applied in validations of multiple climate and weather forecast models [*Goldberg et al.*, 2003; *Pagano et al.*, 2004; *Chahine et al.*, 2006; *Pierce et al.*, 2006; *Fasullo and Trenberth*, 2012; *Jiang et al.*, 2012; *Tian et al.*, 2013]. The AIRS/AMSU-A water vapor and temperature measurements have also been used to analyze the global distributions of relative humidity and ice supersaturation [*Gettelman et al.*, 2006a, 2006b; *Kahn et al.*, 2007, 2009; *Lamquin et al.*, 2012]. Other analyses based on AIRS/AMSU-A observations, such as the water vapor and temperature variance scaling, have also been used to compare with [*Kahn et al.*, 2011] or improve [*Cusack et al.*, 2007] sub-grid scale model parameterizations. Besides the global-scale analyses, AIRS/AMSU-A data have been used to understand regional climatologies. For example, the AIRS/AMSU-A retrievals of land surface relative humidity and moisture provide an essential tool for understanding the hydrologic cycles over continental areas, especially for regions with a scarcity of daily meteorological measurements, such as sub-Saharan Africa [*Ferguson and Wood*, 2010].

[3] The various applications of AIRS/AMSU-A data rely on the accuracy of the observations. Thus, there is an ongoing need to assess and improve the satellite observations by comparing with ground-based or airborne measurements. Previous validations include comparisons with radiosonde observations [*Divakarla et al.*, 2006; *McMillin et al.*, 2007], dropsonde observations [*Wu*, 2009], and in situ aircraft observations

Additional supporting information may be found in the online version of this article.

[1]Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA.
[2]Center for Mid-Infrared Technologies for Health and the Environment, Princeton University, Princeton, New Jersey, USA.
[3]Sonoma Technology, Inc., Petaluma, California, USA.

Corresponding author: M. A. Zondlo, Department of Civil and Environmental Engineering, Princeton University, E206A Engineering Quad., Princeton, NJ 08544, USA. (mzondlo@princeton.edu)

2169-897X/13/10.1002/jgrd.50483

Table 1. Comparison Schemes and Data Resolution of Previous and Current Work

Reference	In Situ Data	Spatial Coverage	Resolution of Averaged In Situ Data	Data Version	Spatial Window	Temporal Window
Gettelman et al. [2004]	Aircraft (PreAVE)	5°S–40°N, Texas and Costa Rica	50 s flight data	V3	0.5° × 0.5° box	Same calendar day
Divakarla et al. [2006]	Radiosonde	Global point measurements	Vertically averaged to AIRS bin	V4	100 km radius	≤3 h
Wu [2009]	Dropsonde	16°W–34°W, 10°N–22°N	Vertically averaged to AIRS bin	V5	0.5° × 0.5° box	≤4.5 h daytime passes
Lamquin et al. [2012]	Aircraft (MOZAIC)	120°W–150°E, 30°S–90°N	1 min flight data	V5	Within AMSU footprint ~22.5 km radius	≤30 min
Current work	Aircraft (START08 and HIPPO Global #1–5) deployments	87°N–67°S, 84°W–180°W–128°E North America, central Pacific Ocean	100 s flight data	V5	Compare various sizes of spatial windows	Compare various sizes of temporal windows

[*Gettelman et al.*, 2004; *Lamquin et al.*, 2012]. For over-land validation, the AIRS/AMSU-A surface air temperature data have been compared with ground meteorological stations [*Gao et al.*, 2008] as well as other satellite remote sensing observations, such as the Advanced Microwave Scanning Radiometer for EOS [*Jones et al.*, 2010]. Radiosonde observations have also been compared with the AIRS/AMSU-A Version 4 Level 1-B water vapor and temperature data between land and ocean from the surface to the upper troposphere/lower stratosphere (UT/LS) [*Divakarla et al.*, 2006]. However, these types of comparisons are subject to two major challenges: first, the satellite observations usually have lower spatial and temporal resolutions than in situ observations, and second, because the satellite samples a relatively large volume of air with a different spatial response over a nearly instantaneous time duration, aircraft measurements will never coincide at the exact same time and location as the satellite observations. To resolve the differences in resolution, the higher-resolution in situ data are usually averaged to coarser resolution before comparing them with satellite observations. In addition, arbitrary selection criteria are usually applied to define certain spatial and temporal ranges so that only the satellite observations within these ranges will be compared with the in situ measurements. Ideally, one would like to compare the satellite observations with the in situ observations that are as close as possible in location and time. However, this would strictly limit the number of data samples. Since the data quantity also influences the statistical significance of the final assessment, the comparisons usually need to balance the size of the averaging window versus the number of data available. In previous comparisons, different spatial and temporal ranges have been used. For example, Table 1 shows the sizes of the spatial windows in previous studies from 22.5 km radius [*Lamquin et al.*, 2012] to 100 km radius [*Divakarla et al.*, 2006] and the sizes of the temporal windows from ≤30 min [*Lamquin et al.*, 2012] to within the same calendar day [*Gettelman et al.*, 2004]. The inherent variability of the atmosphere ultimately drives the sensitivity to the space and time averaging windows. Therefore, these large differences in the spatial and temporal windows set a barrier to intercompare or consolidate these previous validations. Indeed, no previous study has addressed the sensitivities of the validation results to various spatial and temporal averaging window sizes.

[4] In addition to the uncertainties in the comparison methods, the in situ observations also have their own limitations in spatial and temporal coverage. For example, previous aircraft observations used in validation studies are mostly restricted to a horizontal layer, preventing a full vertical profile comparison with the AIRS/AMSU-A data set [*Gettelman et al.*, 2004; *Lamquin et al.*, 2012]. The vertical profile comparisons have mainly relied on radiosonde [*Divakarla et al.*, 2006; *McMillin et al.*, 2007] and dropsonde [*Wu*, 2009] observations, yet these previous dropsonde comparisons were confined to limited geographical locations. Furthermore, the radiosonde water vapor measurements in the UT/LS have been reported to suffer from slow responses at low temperatures (e.g., 27 s response time at −40°C) and ventilation problems when icing occurs [*Miloshevich et al.*, 2001; *Verver et al.*, 2006]. In particular, there were few in situ data sets available over the ocean in the Southern Hemisphere (Table 1, third column). Finally, there has been no previous validation that distinguishes between the land and ocean comparisons or between the Northern Hemisphere (NH) and Southern Hemisphere (SH) for the Version 5 Level 2 AIRS/AMSU-A data. Thus, it is unclear if the Version 5 land retrieval has improved since Version 4 [*Divakarla et al.*, 2006; *Tobin et al.*, 2006]. The sampling limitations in the vertical and horizontal dimensions limit our assessment of the performances of AIRS/AMSU-A. In order to assess the AIRS/AMSU-A performance from a global view, more advanced data sets are needed to provide a full coverage of vertical layers from the surface to the UT/LS, and also over land and ocean in both hemispheres.

[5] In this study, we investigated the sensitivity of validation errors to spatial and temporal windows and compared AIRS/AMSU-A Version 5 Level 2 data between land and ocean and between NH and SH. Our analyses are based on the in situ measurements of two aircraft flight campaigns. The flight campaigns are the National Science Foundation (NSF) HIAPER Pole-to-Pole Observations (HIPPO) Global campaign deployments #1–5 from 2009 to 2011 (flight tracks shown in Figures 1a and 1b) [*Wofsy et al.*, 2011] and the NSF Stratosphere Troposphere Analyses of Regional Transport 2008 (START08) campaign (Figure 1c) [*Pan et al.*, 2010]. The combined data sets range from 87°N to 67°S, cover both North America and the central Pacific Ocean (Table 1), and provide a large number of atmospheric vertical profiles from the surface to the UT/LS. The START08 campaign provided ~120 h flight time and ~90 vertical transects across the thermal tropopause. The HIPPO Global campaign provided ~400 h flight time and ~600 vertical transects from the surface to the tropical UT or the extratropical UT/LS. In particular, the HIPPO Global campaign was designed to achieve a

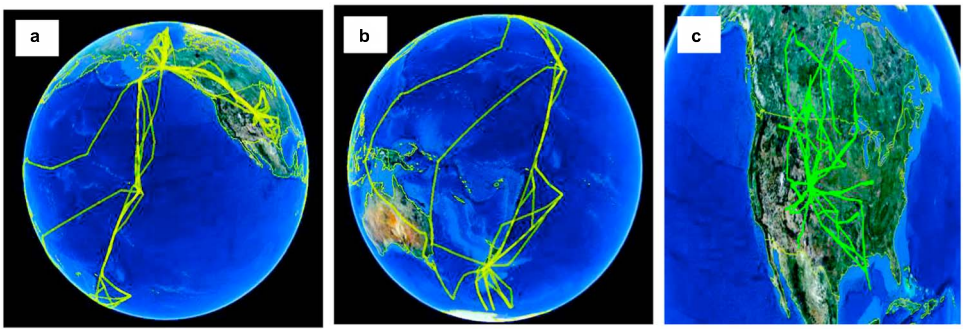

Figure 1. Google maps of the (a, b) HIPPO Global campaign #1–5 deployments (yellow tracks in Figures 1a and 1b) and the (c) START08 campaign (green tracks in Figure 1c). Only flights used in this study are shown in the figure.

vertical profile for every 2.2° of latitude, which provided unprecedented fine-grained atmospheric transects using in situ measurements. The combined data set was compared with the AIRS/AMSU-A Version 5 Level 2 water vapor and temperature at various pressure levels in the both hemispheres as well as over land and ocean. To address the sensitivities of the comparisons between AIRS/AMSU-A and in situ observations with respect to different spatial and temporal averaging windows, we propose a new method to test these sensitivities with various combinations of spatial and temporal window sizes. Not only did we conduct the comparisons under the same temporal and spatial scales of previous analyses, such as *Gettelman et al*. [2004] and *Divakarla et al*. [2006], but we also tested the sensitivities of the validation results from ± 1 h up to ± 12 h, as well as from 22.5 km up to 270 km radii.

2. Data Coverage and Handling

2.1. AIRS/AMSU Observations

[6] AIRS and AMSU-A are two of the six sensors aboard the NASA Aqua satellite. Aqua is a sun-synchronous satellite with ascending and descending orbits crossing the equator at ~1:30 A.M. local time (descending orbit) and ~01:30 P.M. local time (ascending orbit), respectively. This enables AIRS and AMSU-A to scan the Earth two times a day covering 95% of the Earth's surface [*Chahine et al*., 2006]. AIRS is a high-spectral resolution spectrometer with 2378 channels largely covering the infrared (IR) spectrum from 650 to 2665 cm^{-1} with a nominal resolving power ($\lambda/\Delta\lambda$) of 1200. It is a cross-track scanning sensor with one scan every 2.7 s, a field of view (FOV) of 1.1° and a spatial resolution of 13.5 km at nadir [*Chahine et al*., 2006; *Divakarla et al*., 2006]. Because AMSU-A scans three times as slowly (once in 8 s) as AIRS, there are nine AIRS footprints within each AMSU-A footprint. The AIRS/AMSU-A pair together measures a combined footprint at 45 km × 45 km horizontal resolution. Three product types are made available from the retrieval steps: level 1 products (radiances and brightness temperature), level 2 products (geolocated, cloud-cleared radiances and retrieved physical quantities in two or three dimensions, such as moisture and temperature), and level 3 products (1° × 1° gridded products) [*Olsen et al*., 2007b]. Table 1 shows the different versions of AIRS/AMSU-A data used in previous studies. In this study, we use Version 5 Level 2 standard products, which contain 28 standard pressure levels for temperature and 15 pressure levels for water vapor. The product type is denoted by the acronym AIRX2RET and is made available as hdf-formatted files with data recorded every 6 min (a granule of AIRS/AMSU-A data).

2.2. Aircraft Observations

[7] We use the Vertical Cavity Surface Emitting Laser (VCSEL) hygrometer onboard the NSF Gulfstream V (GV)

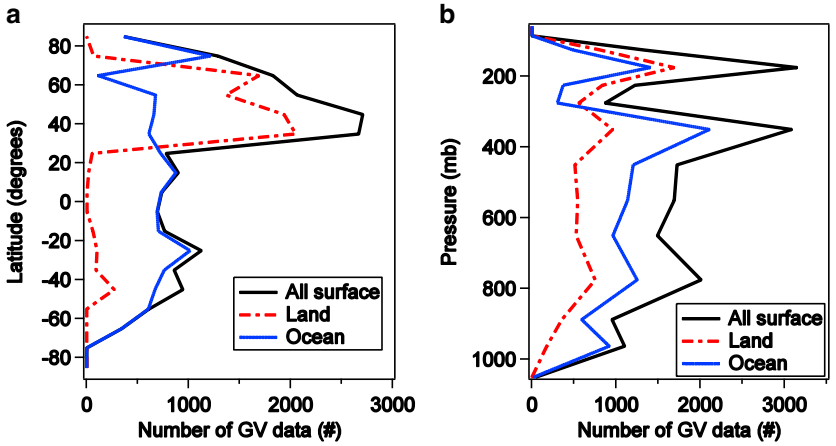

Figure 2. GV data distributions in HIPPO and START08 at different (a) latitudes and (b) pressures.

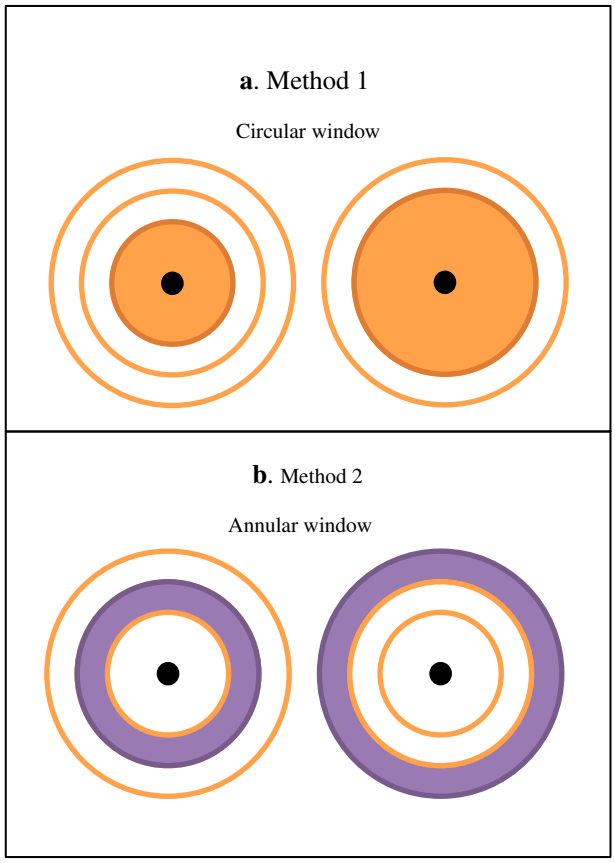

Figure 3. (a) Method 1 circular window (yellow region) and (b) Method 2 annular window (purple region) for selecting the AIRS/AMSU-A data within certain spatial and temporal ranges around each aircraft observation (black dot in the middle).

aircraft to measure water vapor from the surface to the UT/LS. The VCSEL hygrometer is an open-path, near-infrared laser working at 25 Hz [*Zondlo et al.*, 2010]. Two absorption lines are used to measure water vapor: the 1854.03 nm line for water vapor concentrations from the mid-troposphere to the UT/LS and the 1853.37 nm line for the lower troposphere. The measurement has an accuracy of 6% and a precision ≤3%. Calibrations and intercomparisons with other hygrometers have been demonstrated in the literature [*Zondlo et al.*, 2010]. The VCSEL measurements were averaged to 1 Hz in the HIPPO and START08 campaigns. Temperature measurements were recorded by a Rosemount temperature probe, which were reported at 1 Hz. Other positioning variables used for the comparisons include reference static pressure and GPS-derived longitude and latitude.

[8] We combined the HIPPO and START08 flight campaigns for the validation. The combined data set covers from 87°N to 67°S, 84°W–180°W–128°E, transecting from the surface to the UT/LS, with 299 and 214 flight hours over land and ocean, respectively. Almost all flight hours were at the daytime, as the plane usually took off at ~10 A.M. local time and landed at ~6 P.M. local time. The latitudinal and vertical distributions of the aircraft observations are illustrated in Figure 2. Our sampling is mostly limited to Northern America, and southern hemispheric land is exclusively New Zealand and Australia. Flight tracks in HIPPO were generally straight directions and took no efforts to avoid clouds except over deep convection.

3. Comparison Methods

3.1. Spatial and Temporal Comparison Windows

[9] The aircraft measurements were averaged to 100 s resolution, which is ~23 km in the horizontal and less than 1 km in the vertical. Therefore, in this study, "one GV observation" stands for one datum of the 100 s averaged timeseries. The averaging prepares the aircraft data into a coarser resolution comparable to the scale (~22.5 km) of AIRS/AMSU-A data set. We define the absolute differences between the horizontal locations of AIRS/AMSU-A and GV observations as *d*Dist (in km) and the absolute temporal differences as *d*Time (in hours).

[10] Two methods were then used to select the AIRS/AMSU-A pixel grid data which were closely "collocated" with the aircraft observations in time and space. The first method (Figure 3a) compares the aircraft observations with all the satellite observations within $\pm M$ h and $\pm N$ km (M and N are arbitrary values). The circular window method was used in all previous studies and the M and N values are shown in Table 1.

[11] The second method, which is novel, uses an annular window for comparisons (Figure 3b). The region selected by the annular window is the difference between two circular windows, which means that satellite observations within the range of $\pm M$ to $\pm(M+M')$ h and $\pm N$ to $\pm(N+N')$ km around the aircraft data were selected for comparison. For example, if we choose ±6 to ±12 h, only the satellite data that happened between 6 and 12 h before or after the in situ measurement would be compared. Using this method, we can quantify how the temporal and spatial differences between the in situ and satellite observations influence their agreement for water vapor and temperature.

[12] After selecting the AIRS/AMSU-A data, there might be more than one pixel grid inside the temporal and spatial windows around one GV observation. We define *d*Dist as the mean value of all the absolute spatial differences between the satellite data and the GV observation. Similarly, *d*Time is the mean value of all the absolute temporal differences. The calculations of *d*Dist and *d*Time are

$$d\mathrm{Dist} = \frac{\sum_{i=0}^{k} |\mathrm{AIRS}i - \mathrm{GV}|}{k}, \tag{1}$$

$$d\mathrm{Time} = \frac{\sum_{i=0}^{k} |\mathrm{AIRS}i - \mathrm{GV}|}{k}. \tag{2}$$

[13] Here "AIRS_i" denotes the AIRS/AMSU-A pixel grid data within the comparison window; k denotes the number of AIRS/AMSU-A data around one GV observation satisfying both the spatial and temporal window criteria. We note that

Table 2. Quality Control of AIRS/AMSU-A Temperature and Water Vapor Data

	Quality Control
Temperature data	All unit data (grids) within each FOV retrieved at Pressure > PGood are deleted, as recommended by previous literature [*Susskind*, 2007]
Water vapor data	1. All FOVs with Qual_H_2O = 2 are deleted, since it means that the entire water vapor column or FOV is of bad quality 2. All grid data with H_2OMMRStdErr > 0.5 × H_2OMMRStd are deleted, as recommended previously [*Olsen et al.*, 2007a] 3. All grid data retrieved at Pressure > PBest are deleted

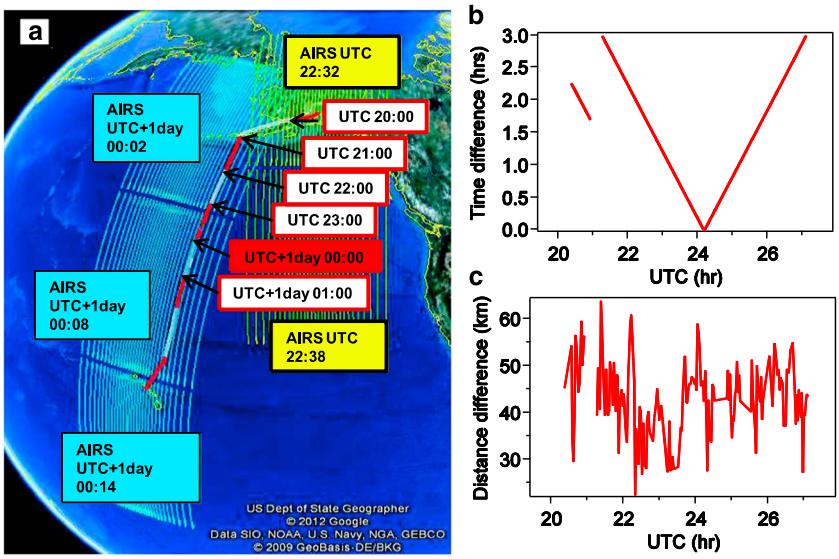

Figure 4. An example of spatial and temporal differences between AIRS/AMSU-A and GV aircraft observations in HIPPO#1 RF04. The time stamp is based on UTC (in h) on the date that the aircraft took off. The UTC time in the next day is labeled as UTC + 1 day. (a) The GV flight track in red and light blue, where the start of each colored segment represents the start of a new hour in UTC. The UTC time for the start of each segment of flight track is labeled in the white text box, such as UTC 20:00. The segment with the best co-location in space and time between the aircraft and AIRS/AMSU-A is labeled in the red text box. Two swaths are illustrated in Figure 4a with yellow and blue colors, respectively. The UTC time of each granule is labeled in yellow or blue square boxes according to the swath's color. (b) The 100 s resolution time series of the averaged *d*Time value for all AIRS/AMSU-A data within 67.5 km and 3 h of each GV observation; (c) similar to Figure 4b but for *d*Dist.

*d*Dist and *d*Time are divided by the same denominator *k* since they select the same AIRS/AMSU-A data for calculation.

3.2. Vertical Interpolation

[14] Water vapor and temperature reported by AIRS/AMSU-A were further interpolated onto the same pressure level as the GV flight level. The interpolation is a necessary step before the comparison, because the vertical mismatches between the two data sets can lead to different readings of water vapor and temperature. To help minimize the impact of vertical dislocations between the two data sets, we interpolated the satellite data by assuming a log-scale water vapor distribution and a linear-scale temperature distribution in the pressure profile.

[15] Water vapor was interpolated by $\log_{10}$(water vapor mixing ratio) $= a + b \times$ Pressure. The coefficients a and b were determined from AIRS/AMSU-A data at two layers: the first layer is where the GV flies through, and the second layer is where its midlevel is the second closest to the GV flight level besides the first layer. We only used two layers for the interpolation because the more layers we use in interpolation, the higher chance that one of them will not satisfy the quality control criteria, and also because the interpolation assumption of a log-scale distribution of water vapor may not apply to a large vertical range.

[16] As an example, if the GV flies at 370 mbar, it is within the AIRS/AMSU retrieval layer of 400–300 mbar. The layers above and below the flight level are 300–200 mbar and 500–400 mbar, respectively. To decide which of these two layers should be used for the interpolation, we compare their midlevels and select the one closer to the aircraft flight level. In this case, the midlevel of the 500–400 mbar layer (i.e., 450 mbar) is closer to the flight level (370 mbar) than the midlevel of the 300–200 mbar layer (i.e., 250 mbar). The reason that we compare the midlevels to the flight level is because the AIRS/AMSU-A water vapor measurement reported at each layer represents the average water vapor concentration inside the layer, even though AIRS/AMSU-A Level 2 data are labeled with the bottom level of the layer. For example, the 500–400 mbar layer is labeled as "500 mb," while the 400–300 mbar layer is labeled as "400 mb." The interpolation example for water vapor is

$$\text{AIRS } H_2O \text{ interpolated value (g/kg)} = 10^{\left(\left(\log_{10}(\text{AIRS } H_2O_{500\text{mb}}) - \log_{10}(\text{AIRS } H_2O_{400\text{mb}})\right) \times \frac{\text{Pressure}_{GV} - 350\text{mbar}}{450\ \text{mbar} - 350\text{mbar}} + \log_{10}(\text{AIRS } H_2O_{400\text{mb}})\right)}; \quad (3)$$

AIRS $H_2O_{500\ mb}$ and AIRS $H_2O_{400\ mb}$ stand for the retrievals of the mean water vapor concentration at 500–400 mbar and 400–300 mbar, respectively.

[17] Temperature was interpolated by Temperature $= c + d \times$ Pressure. Different from the water vapor retrieval, the temperature retrieval represents the temperature at the exact

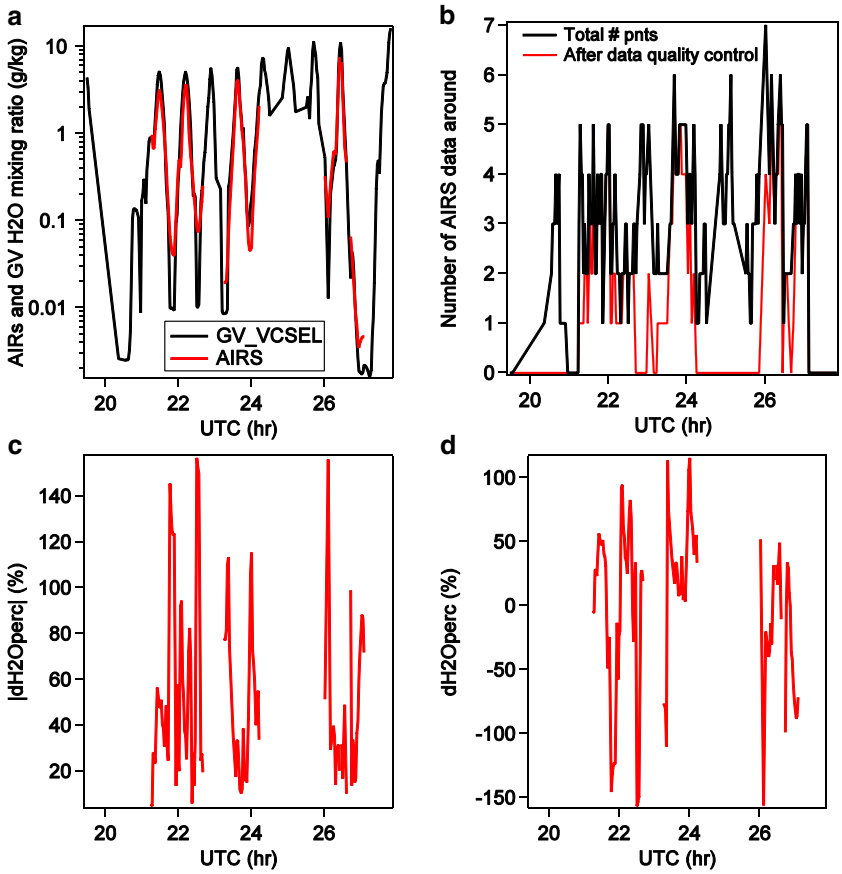

Figure 5. An example of water vapor comparison for HIPPO#1 RF04. (a) The time series of log-scale water vapor mass mixing ratios (g/kg) for VCSEL hygrometer (black) and AIRS/AMSU-A (red). (b) The number of AIRS data satisfying the spatial and temporal selection windows (black line) and the remaining number after applying the water vapor quality control criteria (red line); (c) the time series of $|d\mathrm{H_2O}\mathrm{perc}|$, i.e., the absolute percentage difference of water vapor mass mixing ratio; (d) similar to Figure 5c, but for $d\mathrm{H_2O}\mathrm{perc}$.

pressure level. Therefore, the temperature interpolation coefficients c and d were calculated from the two AIRS/AMSU-A levels where the GV flew in between. For example, if the GV flies at 370 mbar, the AIRS/AMSU-A 300 mbar and 400 mbar levels would be used for the interpolation. The calculation of the linear interpolation onto the GV flight level is

$$\text{AIRS Temperature interpolated value} = \left(\text{AIRS } T_{400\text{mb}} - \text{AIRS } T_{300\text{mb}}\right) \cdot \frac{\text{Pressure}_{\text{GV}} - 300\text{mbar}}{400\text{mbar} - 300\text{mbar}} + \text{AIRS } T_{300\text{mb}}; \quad (4)$$

AIRS $T_{400\ \mathrm{mb}}$ and AIRS $T_{300\ \mathrm{mb}}$ stand for the AIRS/AMSU-A temperature retrievals at 400 mbar and 300 mbar, respectively.

3.3. AIRS/AMSU-A Data Quality Control

[18] After selecting the AIRS/AMSU-A data within a certain horizontal spatial and temporal windows around one GV point, we quality controlled the selected satellite observations of temperature and water vapor with different criteria (Table 2). The presence of clouds in the FOV of the satellite sensor can complicate the retrieval processes [*Susskind et al.*, 2003, 2006]. Therefore, for each variable, we checked the quality of the data at the two pressure levels being used in the vertical interpolations. If one of the two levels did not pass the quality control criteria, the retrieval would not be used to compare with this GV observation. This filtering process selects satellite retrievals that are of good quality and also reduces the computational cost by eliminating the nonqualified data before moving to the next step of comparison. We note that there are higher ratios of data at the lower altitudes being filtered out, since the retrievals at lower altitudes are more likely to be complicated by the presence of clouds.

3.4. Definitions of Water Vapor and Temperature Differences Between AIRS/AMSU-A and GV Measurements

[19] For one GV observation, there might be more than one AIRS/AMSU-A data around it that satisfy all the selection window criteria and the quality control criteria. In the first step, the comparison was conducted between this GV observation and each of the AIRS/AMSU-A data surrounding it. Then all of these individual comparisons were averaged to represent the final comparison result for this GV point. The reason for taking the average for all AIRS/AMSU-A data around one GV observation is to make sure that the final comparison result is not heavily weighted over a few GV observations with a large number of AIRS/AMSU-A data around them.

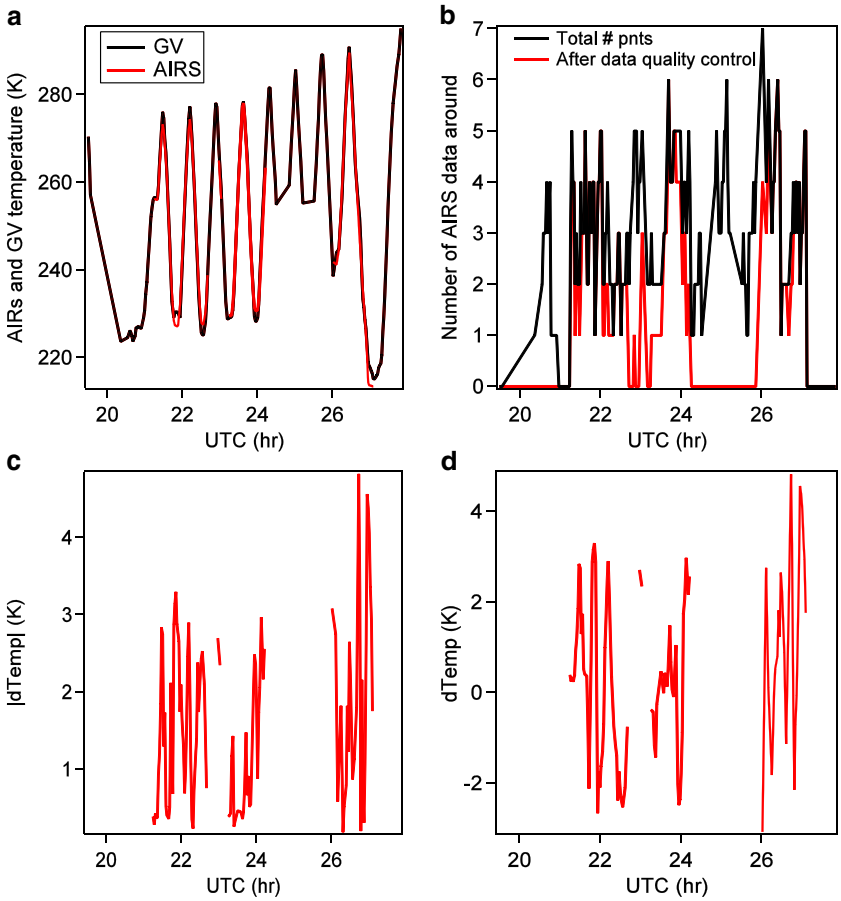

Figure 6. Similar to Figure 5 but for the temperature comparison example of HIPPO#1 RF04.

[20] The difference in water vapor between AIRS/AMSU-A and the VCSEL hygrometer was represented in two ways: the absolute difference of water vapor in percentage, defined as $|d\mathrm{H_2Operc}|$ (%) (equation (5a)), and the difference of water vapor in percentage, defined as $d\mathrm{H_2Operc}$ (%) (equation (5b)). The percentage difference was used instead of the difference (in g/kg) to allow assessment of the full vertical profile of water vapor retrieval. Water vapor concentrations vary by 5 orders of magnitude from the surface to the UT/LS, and the difference (in g/kg) at the UT/LS is orders of magnitude

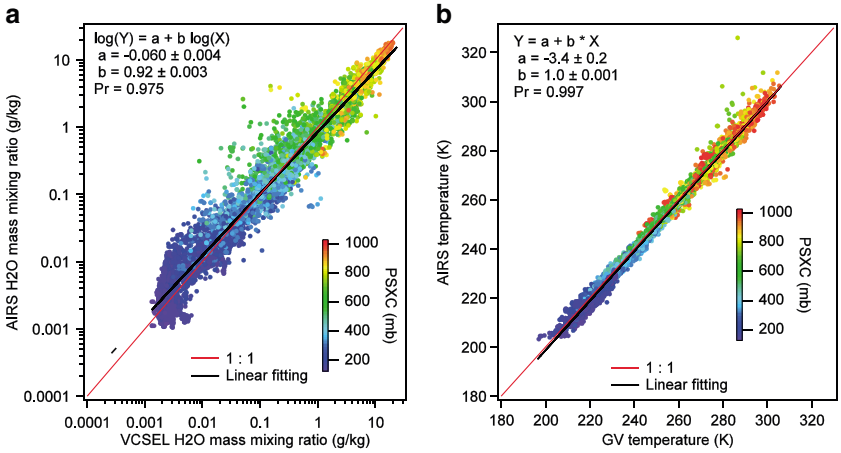

Figure 7. Comparisons of (a) water vapor and (b) temperature for all flights. The scatterplots show the correlation between AIRS/AMSU-A (ordinate) and GV data (abscissa). The color coding represents the pressure of the GV aircraft from the UT/LS (purple) to the surface (red). For water vapor comparison, the linear fit (red line) is $\log_{10}(\mathrm{AIRS}) = a + b \times \log_{10}(\mathrm{VCSEL})$. For temperature, the linear fit (red line) is $\mathrm{AIRS} = a + b \times (\mathrm{GV})$. "AIRS" denotes the whole AIRS/AMSU-A data set. "Pr" denotes the Pearson-R value. When the Pr value is closer to 1, it stands for a stronger correlation between the two data sets.

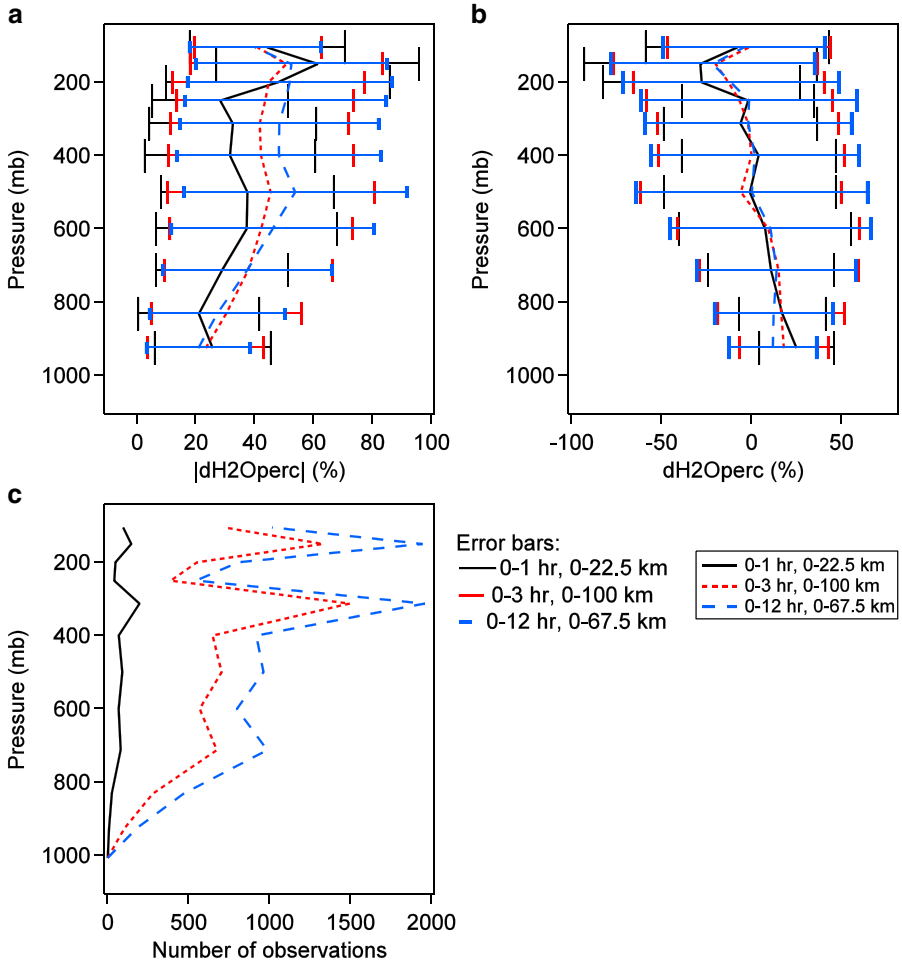

Figure 8. Water vapor comparisons in each pressure layer for all flight data. Different colors represent different selection windows in space and time: ±1 h and ±22.5 km (solid black line), ±3 h and ±100 km (dotted red line, same sizes as *Divakarla et al*. [2006]), ± 12 h and ± 67.5 km (dashed blue line, similar sizes to within 1 calendar day and ±0.5° in *Gettelman et al*. [2004]). (a) The $|d\mathrm{H_2Operc}|$ values for all the GV observations in each pressure layer from the surface to the UT/LS. Error bars denote one standard deviation for all the $|d\mathrm{H_2Operc}|$ inside each pressure level. (b) Similar to Figure 8a but for $d\mathrm{H_2Operc}$. (c) The number of GV data being used for the water vapor comparisons in each pressure layer.

smaller than the difference at the surface. For each AIRS/AMSU-A and GV comparison pair, we calculated one $|d\mathrm{H_2Operc}|$ and one $d\mathrm{H_2Operc}$ value, and the final $|d\mathrm{H_2Operc}|$ (or $d\mathrm{H_2Operc}$) represents the mean of all the $|d\mathrm{H_2Operc}|$ (or $d\mathrm{H_2Operc}$) values for this GV observation. The equation is shown as below, for which we use "AIRS" to denote the whole AIRS/AMSU-A data set:

$$|d\mathrm{H_2Operc}| = \frac{\sum_{i=0}^{a}\left(\frac{|\mathrm{AIRS}_i - \mathrm{VCSEL}|}{\mathrm{VCSEL}} \times 100\%\right)}{a}; \quad (5a)$$

$$d\mathrm{H_2Operc} = \frac{\sum_{i=0}^{a}\left(\frac{\mathrm{AIRS}_i - \mathrm{VCSEL}}{\mathrm{VCSEL}} \times 100\%\right)}{a}. \quad (5b)$$

[21] Similarly, the difference in temperature between the AIRS/AMSU-A and the GV measurement was represented in two ways: absolute temperature difference, $|d\mathrm{Temp}|$ (in Kelvin) (equation (6a)), and temperature difference, $d\mathrm{Temp}$ (in Kelvin) (equation (6b)):

$$|d\mathrm{Temp}| = \frac{\sum_{i=0}^{b}|\mathrm{AIRS}_i - \mathrm{GV}|}{b}; \quad (6a)$$

$$d\mathrm{Temp} = \frac{\sum_{i=0}^{b}\mathrm{AIRS}_i - \mathrm{GV}}{b}. \quad (6b)$$

[22] We caution here that the number of AIRS/AMSU-A water vapor data (a in equations (5a) and (5b)) and the number of temperature data (b in equations (6a) and (6b)) used to compare with the GV observation might not be the same, since these two variables require different quality control criteria as shown in Table 2.

3.5. Averaged $d\mathrm{H_2Operc}$ and $d\mathrm{Temp}$ Values in Each Pressure Level

[23] After $|d\mathrm{H_2Operc}|$, $d\mathrm{H_2Operc}$, $|d\mathrm{Temp}|$ and $d\mathrm{Temp}$ were calculated for each GV observation using equations

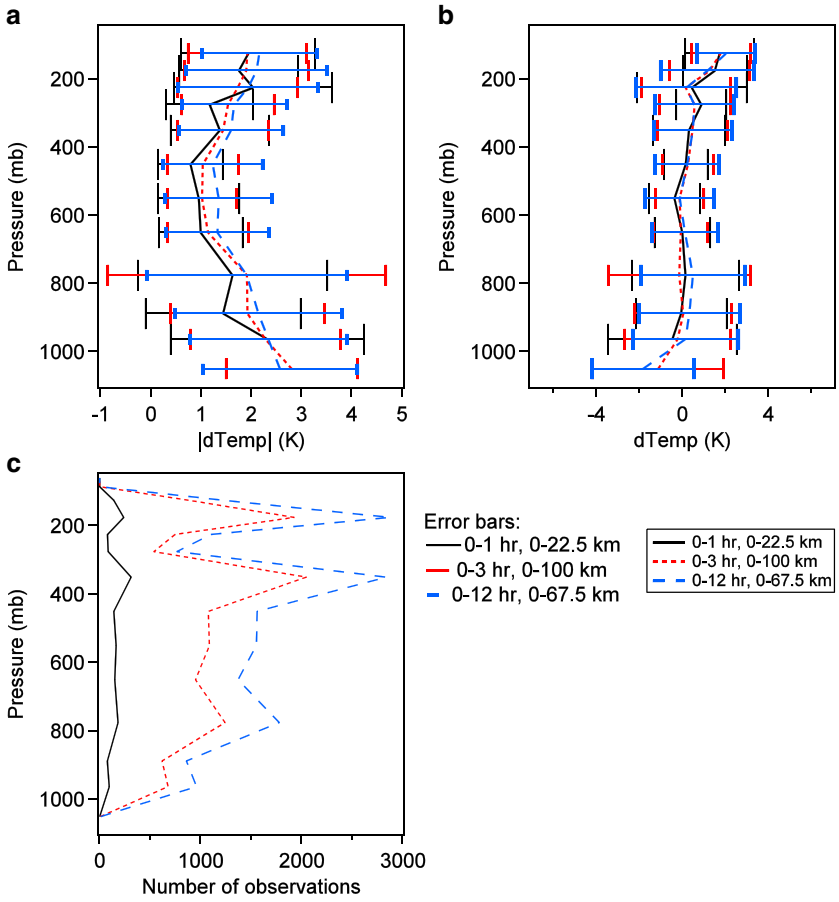

Figure 9. Similar to Figure 8 but for temperature comparison. (a) |*d*Temp|, (b) *d*Temp, and (c) the number of GV observations for temperature comparisons at each pressure level.

(5a), (5b), (6a), and (6b), we calculated their average values in each pressure level as

$$|d\mathrm{H_2Operc}| \text{ in each pressure level} = \frac{\sum_{j=0}^{c} |d\mathrm{H_2Operc}|_j}{c}, \quad (7a)$$

$$d\mathrm{H_2Operc} \text{ in each pressure level} = \frac{\sum_{j=0}^{c} d\mathrm{H_2Operc}_j}{c}, \quad (7b)$$

$$|d\mathrm{Temp}| \text{ in each pressure level} = \frac{\sum_{j=0}^{d} |d\mathrm{Temp}|_j}{d}, \quad (8a)$$

$$d\mathrm{Temp} \text{ in each pressure level} = \frac{\sum_{j=0}^{d} d\mathrm{Temp}_j}{d}. \quad (8b)$$

[24] Here c and d denote the number of GV observations in each pressure level that have been compared with AIRS/AMSU-A water vapor and temperature data, respectively. Again, we note that c and d might not be the same, since these two variables require different quality control criteria.

4. Results

4.1. Comparison Examples From HIPPO#1 RF04

[25] A typical flight example of HIPPO#1 RF04 is used to illustrate the geographical positions of the GV aircraft and the AIRS/AMSU-A FOVs around the aircraft track (Figure 4a). During this flight, the aircraft took off from Anchorage, Alaska, at UTC 19:30 (Anchorage local time 11:30 A.M.) and landed at Honolulu, Hawaii, at UTC +1 day 03:50 (Hawaii local time 5:50 P.M.). The flight duration was about 8 h and 20 min. The horizontal location of the GV flight track is shown as the thick line in red and light blue in Figure 4a. The color change between light blue and red represents the beginning of a new hour in UTC. Thus, each segment of the flight track in light blue or red represents 1 h of flight, except that the first and last segments are not full hours. The UTC time of the beginning of each segment is labeled with white text boxes with black arrows pointing to the start of each UTC hour. The red text box for UTC + 1 day 00:00 shows the segment of the flight where a co-location in both space and time happened between the GV and satellite observations. The closest two swaths around the GV flight track are plotted in yellow and blue (Figure 4a). The UTC time for each granule is labeled by a yellow or blue box, and each granule is 6 min apart. Within the AIRS granule "UTC + 1 day 00:08," the GV and satellite data were collocated in both space and time. During this 8 h flight, only a few data samples were available when the aircraft and the satellite observations happened at the same location and time, illustrating the difficulty of comparing the observations between aircraft and satellite, as they constantly change relative positions in space and time.

[26] The time series of the differences in time and space for this example flight (HIPPO#1 RF04) are illustrated in

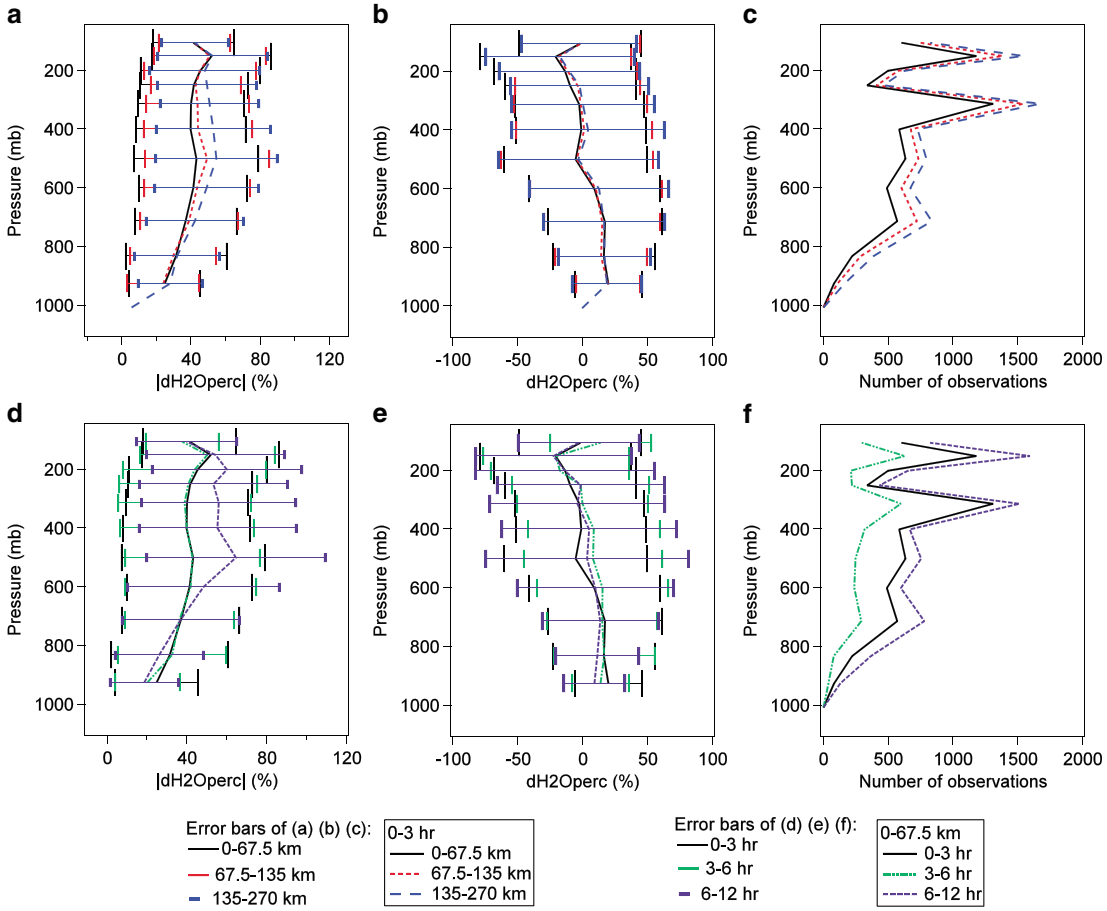

Figure 10. Influences of spatial and temporal window sizes on the water vapor comparison results. (a–c) Comparisons with fixed temporal window size of 0–3 h and increasing horizontal spatial window sizes. Different colors in Figures 10a–10c denote 0–67.5 km (solid black), 67.5–135 km (dotted red), and 135–270 km (dashed blue). Figures 10a–10c are vertical profiles of $|d\mathrm{H_2Operc}|$, $d\mathrm{H_2Operc}$, and the number of GV observations involved in the comparison at each pressure level, respectively. (d–f) Similar to Figures 10a–10c but for increasing temporal window sizes at fixed spatial window of 0–67.5 km. Different colors in Figures 10d–10f denote different temporal window sizes: 0–3 h (solid black), 3–6 h (dash-dotted green), and 6–12 h (short dashed purple). All error bars denote one standard deviation for $|d\mathrm{H_2Operc}|$ or $d\mathrm{H_2Operc}$ values at each pressure level.

Figures 4b and 4c, respectively. The *d*Dist and *d*Time were calculated based on equations (1) and (2), respectively. Here we use the ±3 h and ±67.5 km window, which is the subset of two previous windows: 3 h and 100 km [*Divakarla et al.*, 2006], and 12 h and 67.5 km [*Gettelman et al.*, 2004]. This window pair is comparable to the mean horizontal wind speed of 6.7 m/s in the midtroposphere [*Peixoto and Oort*, 1992]. By choosing this combination of spatial and temporal windows, we consider the possible transport of the air within 3 h of the flight track at a mean horizontal speed of 6.7 m/s, although not the direction of the transport which is beyond the scope of current work. In Figure 4c, *d*Time reached zero at UTC + 1 day 00:11, but *d*Dist at this time stamp was not zero (Figure 4b) because there were more than one AIRS/AMSU-A data used to calculate the average *d*Dist within 67.5 km (equation (1)). The *d*Time from UTC 20:22 to 20:54 was discontinuous because two different swaths of AIRS/AMSU-A were used for the comparisons. During the earlier segment, the GV flight track was closer to the yellow granules of AIRS/AMSU-A data, while in the latter segment, the GV flew into a different location and was closer to the blue granules. As the two sets of granules were ~1.5 h apart, the averaged *d*Time had an abrupt change. The gradual changes of *d*Time after UTC 21:15 represent the continuous changes in aircraft time.

[27] We further calculated the time series for water vapor (Figure 5) and temperature comparisons at 100 s resolution using the 67.5 km and 3 h window (Figure 6). The time series of $|d\mathrm{H_2Operc}|$ (%) (Figure 5c) was calculated by equation (5a), that is, the average value of all the absolute percentage differences between all the AIRS/AMSU-A data around one GV observation. Similarly, Figure 5d shows $d\mathrm{H_2Operc}$ (%) as calculated by equation (5b). The $|d\mathrm{H_2Operc}|$ values range from ~10% to 160%, and $d\mathrm{H_2Operc}$ values range from −160% to 120%. Generally, the larger $|d\mathrm{H_2Operc}|$ happens at the UT/LS, not only because a slight variation at low water vapor concentrations can result in a large percentage difference but also due to the low sensitivity of AIRS in this region [*Fetzer et al.*, 2008; *Liang et al.*, 2010]. The water vapor concentration of AIRS/AMSU-A in Figure 5a was calculated by

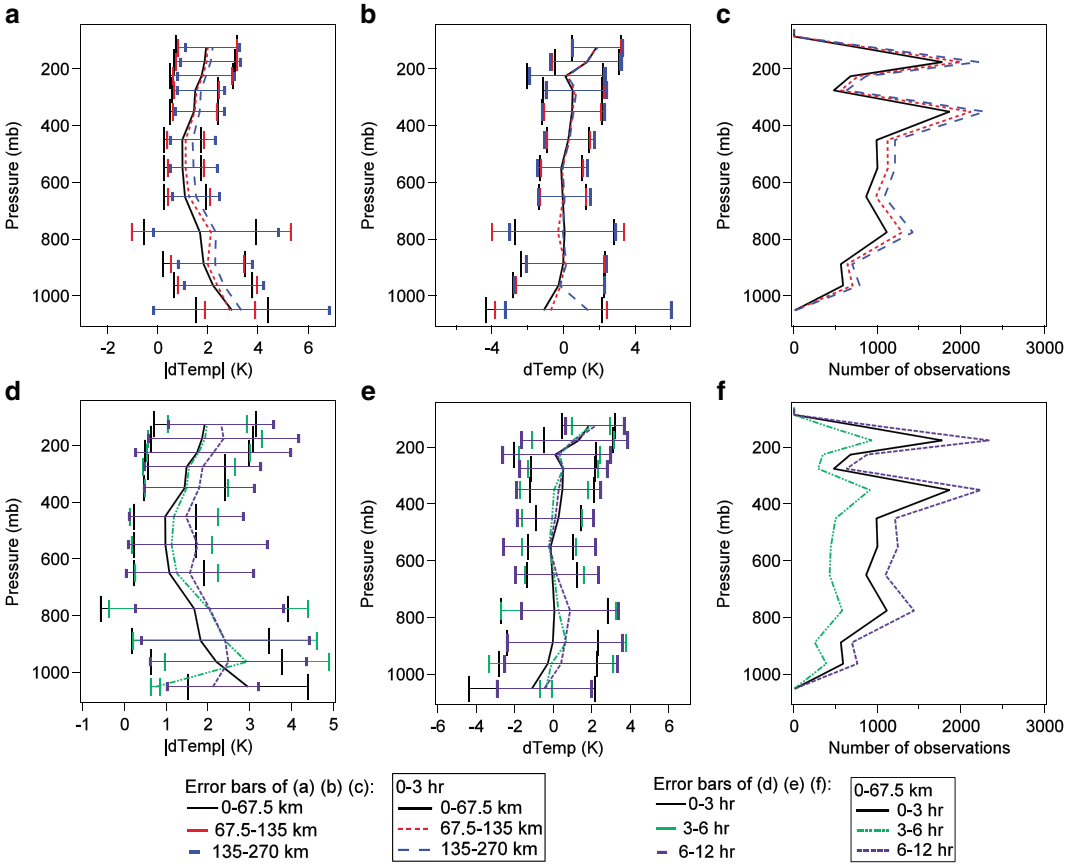

Figure 11. Similar to Figure 10, but for the influences of *d*Dist and *d*Time on the temperature comparison results.

taking the average of all qualifying AIRS/AMSU-A points before calculating any differences, while the $|d\mathrm{H_2Operc}|$ in Figure 5c (or $d\mathrm{H_2Operc}$ in Figure 5d) was calculated by taking the absolute percentage differences (or percentage difference) first and then averaging them. Similarly, the example of temperature comparisons is illustrated in Figure 6. |*d*Temp| and *d*Temp in Figures 6c and 6d were calculated by equations (6a) and (6b), respectively. In this example, |*d*Temp| ranges from ~0 to 5 K, and *d*Temp ranges from ~ −3 to 5 K.

4.2. Syntheses of Water Vapor and Temperature Comparisons for All Flights

[28] The comparisons between AIRS/AMSU-A and GV data were conducted for the whole data set using the 3 h and 67.5 km window. Figures 7a and 7b show the correlation of water vapor and temperature, respectively, with color coding of pressure from the surface (red) to the UT/LS (purple). Each AIRS/AMSU-A datum in Figure 7 represents the mean value of all qualifying satellite data around one GV observation. Water vapor mixing ratio (g/kg) correlation was fitted by $\log_{10}(\mathrm{H_2O_{AIRS}}) = a + b \times \log_{10}(\mathrm{H_2O_{VCSEL}})$, where the slope is 0.92 ± 0.003 and the intercept is -0.060 ± 0.004 (Figure 7a). Temperature correlation was fitted by $T_{\mathrm{AIRS}} = a + b \times T_{\mathrm{GV}}$, where the slope is 1.0 ± 0.001 and the intercept is -3.4 ± 0.2 (Figure 7b). The slope of the temperature linear fit is closer to 1 than the water vapor slope, implying a better agreement between the satellite and aircraft temperatures data than the water vapor data. The Pearson-R (Pr) correlation coefficients for water vapor and temperature comparisons are 0.975 and 0.997, respectively. *Gettelman et al.* [2004] showed that for the validation of Version 3 Level 2 data in the UT/LS region (pressure ≤500 hPa), the Pr values of water vapor and temperature comparisons were 0.91 and 0.98, respectively [*Gettelman et al.*, 2004, Figure 2]. In comparison, at 3 h and 67.5 km window, our validation results of pressure ≤500 hPa show that the Pr of water vapor and temperature are 0.953 and 0.976, respectively. And for pressure >500 hPa, the Pr of water vapor and temperature are 0.926 and 0.987, respectively. Therefore, compared with AIRS Version 3 Level 2 data in *Gettelman et al.* [2004], AIRS Version 5 Level 2 data show better agreement with the in situ observations for water vapor and similar agreement for temperature.

4.3. Water Vapor and Temperature Comparisons Using Circular Comparison Windows

[29] The comparison results of water vapor and temperature in each AIRS/AMSU-A pressure level are shown in Figures 8 and 9, respectively. $|d\mathrm{H_2Operc}|$ and $d\mathrm{H_2Operc}$, and |*d*Temp| and *d*Temp represent the mean values in each pressure level (equations (7a) and (7b), and (8a) and (8b), respectively). The AIRS/AMSU-A pressure levels involved in the comparison are from Level #1 (1100 mbar) to Level #13 (100 mbar). The maximum and minimum pressure values of the whole aircraft data set are 1021 mbar and 133 mbar,

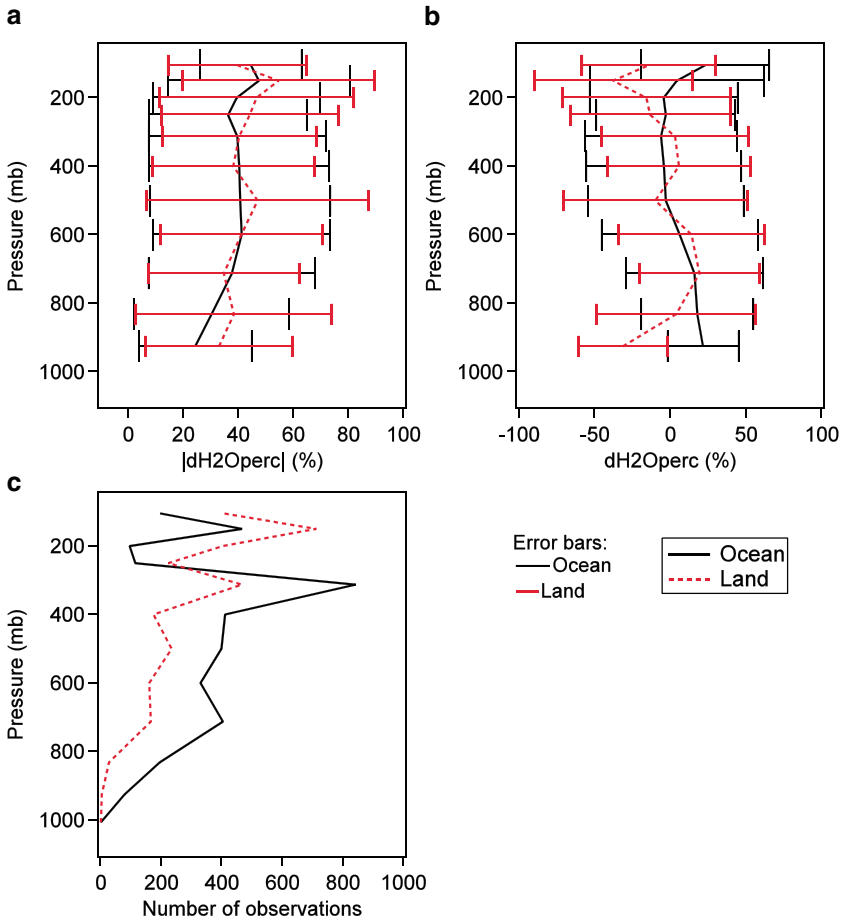

Figure 12. Comparisons of water vapor measurements over ocean (solid black line) and land (dotted red line). The comparisons are within the 3 h and 67.5 km windows. Vertical profiles of (a) $|d\mathrm{H_2O}\mathrm{perc}|$ and (b) $d\mathrm{H_2O}\mathrm{perc}$. (c) The number of GV observations in each pressure level. Error bars denote one standard deviation of $|d\mathrm{H_2O}\mathrm{perc}|$ or $d\mathrm{H_2O}\mathrm{perc}$ in each pressure layer.

respectively. We note that although in previous studies, water vapor data at pressure >300 mbar were usually used for the comparison with climate models [*Jiang et al*., 2012; *Tian et al*., 2013], there is still a strong need to show the water vapor retrieval uncertainties in the UT/LS region since other studies have also used water vapor at 300–150 mbar for analyses of relative humidity and ice supersaturation distributions [*Gettelman et al*., 2006a; *Lamquin et al*., 2012]. In addition, we do not exclude the intercomparisons at low pressures (pressure < 300 mbar) because the sensitivity of AIRS is more dependent on the water vapor concentration (~10–20 ppmv) [*Gettelman et al*., 2004; *Read et al*., 2007] instead of pressure. In fact, further algorithm improvements in the future may help to improve the sensitivity of AIRS water vapor retrievals. Thus, understanding the edge of AIRS sensitivity at the UT/LS is important. To compare with the past validation results, we chose several spatial and temporal windows that have been used in previous analyses. In Figures 8 and 9, we used three combinations of spatial and temporal windows to compare the satellite data: (1) ±3 h and ±100 km (dotted red line, same windows as *Divakarla et al*. [2006]), (2) ±12 h and ±67.5 km (dashed blue line, similar to *Gettelman et al*. [2004] within 1 calendar day and ±0.5°), and (3) a smaller window of ±1 h and ±22.5 km (solid black line). This small window size has not been reported before in previous validations from the surface to the UT/LS. *Lamquin et al*. [2012] applied a similar combination of small spatial (≤22.5 km) and temporal (≤30 min) windows, but were restricted to the UT/LS region. Using the small window, we can improve the assessments of the AIRS/AMSU-A data by showing results when they were closely collocated with in situ measurements.

[30] In general, as the spatial and temporal window sizes increase, both $|d\mathrm{H_2O}\mathrm{perc}|$ and $|d\mathrm{Temp}|$ values become larger, and this trend happens at almost every pressure level. The exception happens in the boundary layer (≥900 mbar) and the UT/LS (≤200 mbar), where there were fewer samples. Aside from the differences in the amount of data, the different sensitivities to comparison windows could also be a result of the different variance scaling characteristics in these pressure ranges [*Kahn and Teixeira*, 2009]. When using the same spatial and temporal window sizes (3 h and 100 km), our results show similar $|d\mathrm{H_2O}\mathrm{perc}|$ and $|d\mathrm{Temp}|$ results to *Divakarla et al*. [2006], i.e., ~20–60% for $|d\mathrm{H_2O}\mathrm{perc}|$ and ~2.5–1 K for $|d\mathrm{Temp}|$ from the surface to the UT/LS. The $|d\mathrm{Temp}|$ at the surface (~2.5 K in Figure 8c) is slightly higher than that of *Divakarla et al*. [2006] (~1.7 K), which could be a result of the small number of qualified comparisons (e.g., 11 GV observations) at the surface in our data. However, using the smaller window of 1 h and 22.5 km, the water

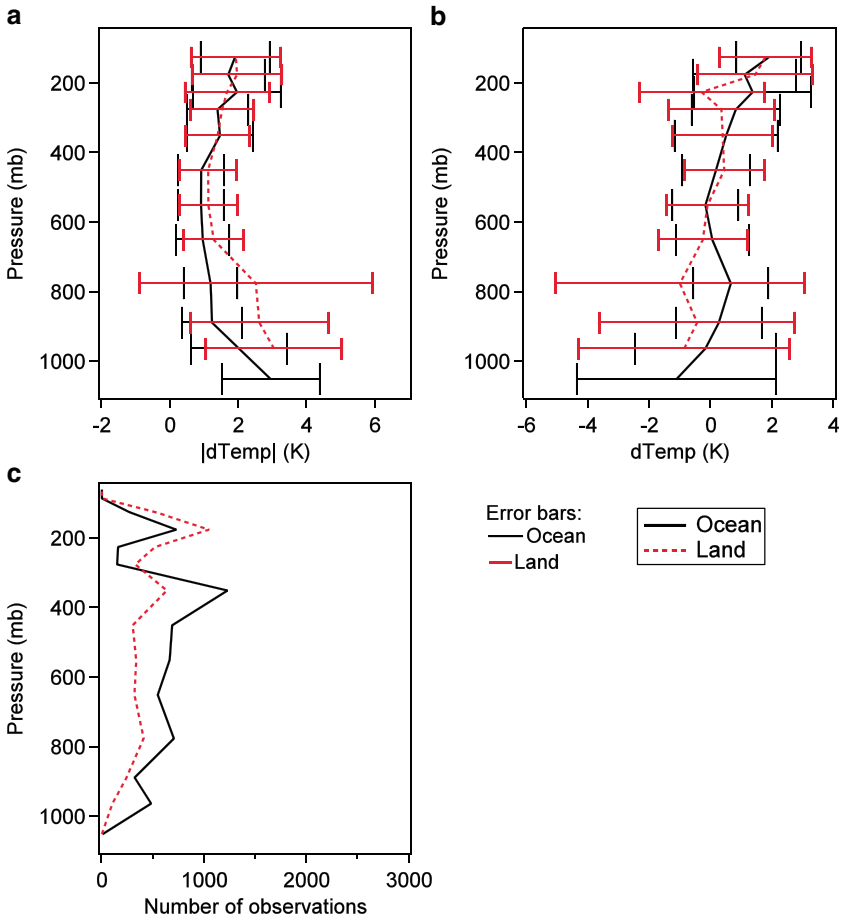

Figure 13. Similar to Figure 12 but for |*d*Temp| and *d*Temp values over land and ocean.

vapor and temperature comparison results can be improved by up to ~10% in |dH$_2$Operc| and ~0.5 K in |*d*Temp| compared with using the 3 h and 100 km window. In addition, using the 12 h and 67.5 km window (similar to the window size of *Gettelman et al.* [2004]) in our data resulted in up to ~10% increase in |dH$_2$Operc| and ~0.3 K increase in |*d*Temp| than using the window of *Divakarla et al.* [2006]. These results imply that the intercomparisons between different validation studies need to account for the uncertainties contributed from the various window sizes.

4.4. Sensitivities of the Comparison Results to Spatial and Temporal Differences Using Annular Comparison Window

[31] The comparisons based on Method 1 show that both spatial and temporal window sizes influence the comparison results. To further quantify these influences, we used Method 2—the annular window—to calculate the |dH$_2$Operc|, dH$_2$Operc, |*d*Temp|, and *d*Temp values for the satellite data in each "band" of space and time around the GV observations. To illustrate the sensitivities to different time window sizes, we fixed the distance window at 0–67.5 km and compared the AIRS/AMSU-A data within 0–3, 3–6, and 6–12 h of the GV data. Similarly, we fixed the time window at 0–6 h and made the comparison at 0–67.5, 67.5–135, and 135–270 km. The results show that |dH$_2$Operc| and |*d*Temp| values increase with the spatial and temporal window sizes, even though the number of GV observations selected by each window is almost the same (Figures 10 and 11). There is a higher sensitivity of |dH$_2$Operc| to the increasing time window sizes at ~200–700 mbar than in the boundary layer or above 200 mbar, and the maximum sensitivity is around 550 mbar. The maximum sensitivity at 550 mbar agrees with the previous finding that the variance scaling generally has the highest value around 400–600 mbar at 24°S, 90°W and 24°N, 90°W [*Kahn and Teixeira*, 2009]. Similarly, |*d*Temp| has high sensitivity to increasing time window at ~200–700 mbar. Comparing Figures 10a and 10d, the increase in |dH$_2$Operc| from 0–3 h to 3–6 h is similar to that from 0–67.5 km to 67.5–135 km, while the |dH$_2$Operc| increase from 67.5–135 km to 135–270 km is much larger than that from 3–6 h to 6–12 h. For temperature, the increase in |*d*Temp| for each spatial and temporal window band is similar. The vertical and horizontal variabilities of the sensitivities of |dH$_2$Operc| and |*d*Temp| to spatial and temporal windows might be a result of different variance scaling characteristics at difference regions, as well as the diurnal cycle or the vertical and horizontal wind speed distributions, and further investigation is needed to determine the exact cause.

[32] On the contrary, the changes of dH$_2$Operc and *d*Temp values are not as obvious compared with those of |dH$_2$Operc| and |*d*Temp|. The less obvious trend for dH$_2$Operc and *d*Temp may be a result of the random distribution of satellite data around the GV observation as the averaging window sizes expand. That is to say, the AIRS/AMSU-A data value may become either larger or smaller than the aircraft data

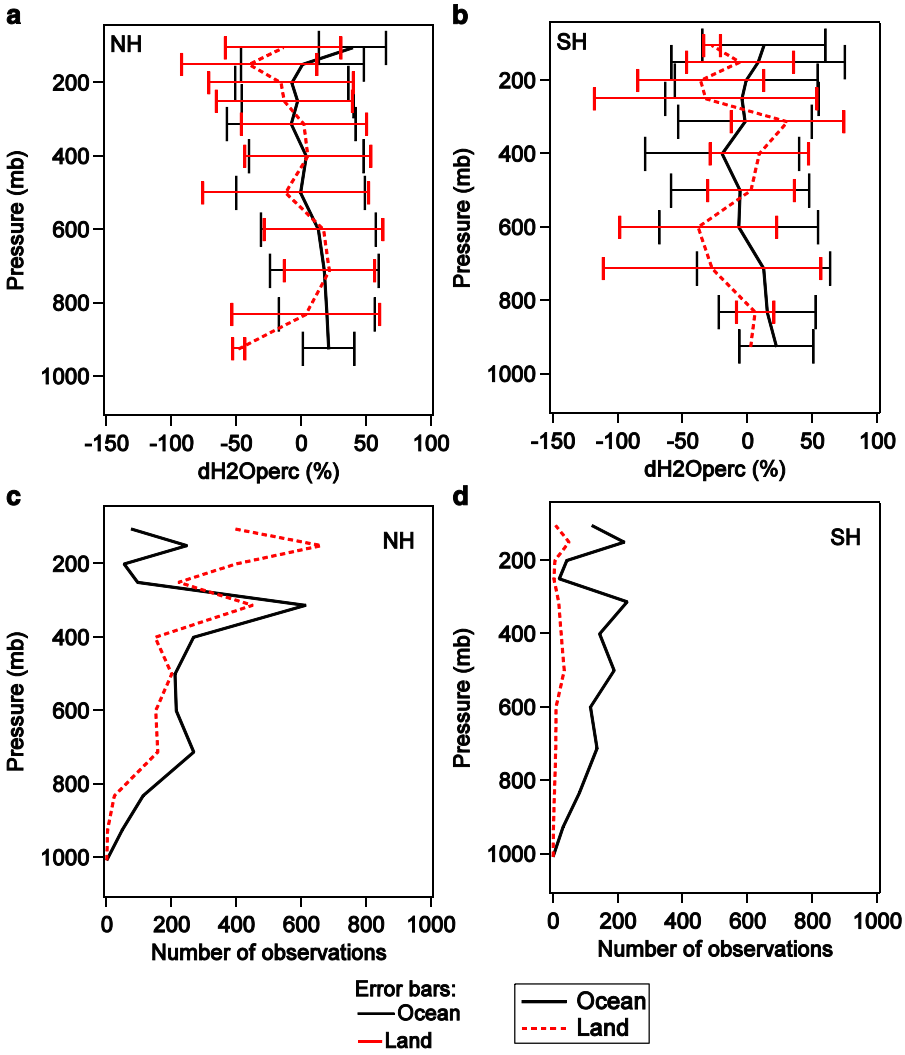

Figure 14. Comparisons of dH$_2$Operc values between the NH and SH. Solid black line represents the comparison results over land, while dotted red line is for over ocean. The vertical profiles of dH$_2$Operc in the (a) NH and (b) SH. The number of GV observations used in the comparisons at each pressure level in the (c) NH and (d) SH.

value when the distance and time differences increase. Therefore, even though the disagreement between each AIRS/AMSU-A datum and the aircraft observation becomes larger as the comparison window expands, the mean of these differences (including both positive and negative values) does not change very much.

4.5. Ocean Versus Land Comparisons for Water Vapor and Temperature

[33] The land fraction in the satellite FOV has an influence on the quality of AIRS/AMSU-A data as illustrated in previous work [*Divakarla et al.*, 2006]. Here we separated all the comparisons into land (landFrac $\geq$ 0.5) and ocean (landFrac $<$ 0.5). To ensure enough comparison samples in each category, we broaden the window size to 3 h and 67.5 km. The vertical profiles of |dH$_2$Operc|, dH$_2$Operc, |dTemp|, and dTemp of land and ocean are illustrated in Figures 12 and 13, respectively. The number of GV observations used in the |dH$_2$Operc| calculation was the same as that in dH$_2$Operc calculation (Figure 12c), since both calculations were subject to the same selection windows and quality control criteria. Similarly, the number of GV observations is the same for both |dTemp| and dTemp calculations (Figure 13c). At 300–800 mbar, the |dH$_2$Operc| values over land are ~0–5% higher than those over ocean (Figure 12a). This result suggests an improvement in the land retrieval in Version 5, since previously an ~10% disagreement in water vapor root mean square values were reported between land and ocean validations using Version 4 data [*Divakarla et al.*, 2006]. Yet in the boundary layer ($\geq$800 mbar) and the UT/LS ($\leq$300 mbar), the |dH$_2$Operc| values over land still have ~10% disagreement with those over ocean.

[34] The dH$_2$Operc vertical profile in Figure 12b shows a larger disagreement (~20–50%) between land and ocean observations compared with the |dH$_2$Operc| analyses in Figure 12a in the boundary layer ($\geq$800 mbar) and the UT/LS ($\leq$200 mbar). Based on the dH$_2$Operc analyses, the land (and ocean) retrievals are drier (and moister) than the aircraft observations by up to ~ −30% (~20%) of dH$_2$Operc in the boundary layer and UT/LS. Yet we note that the large differences in the boundary layer could partly result from the lack

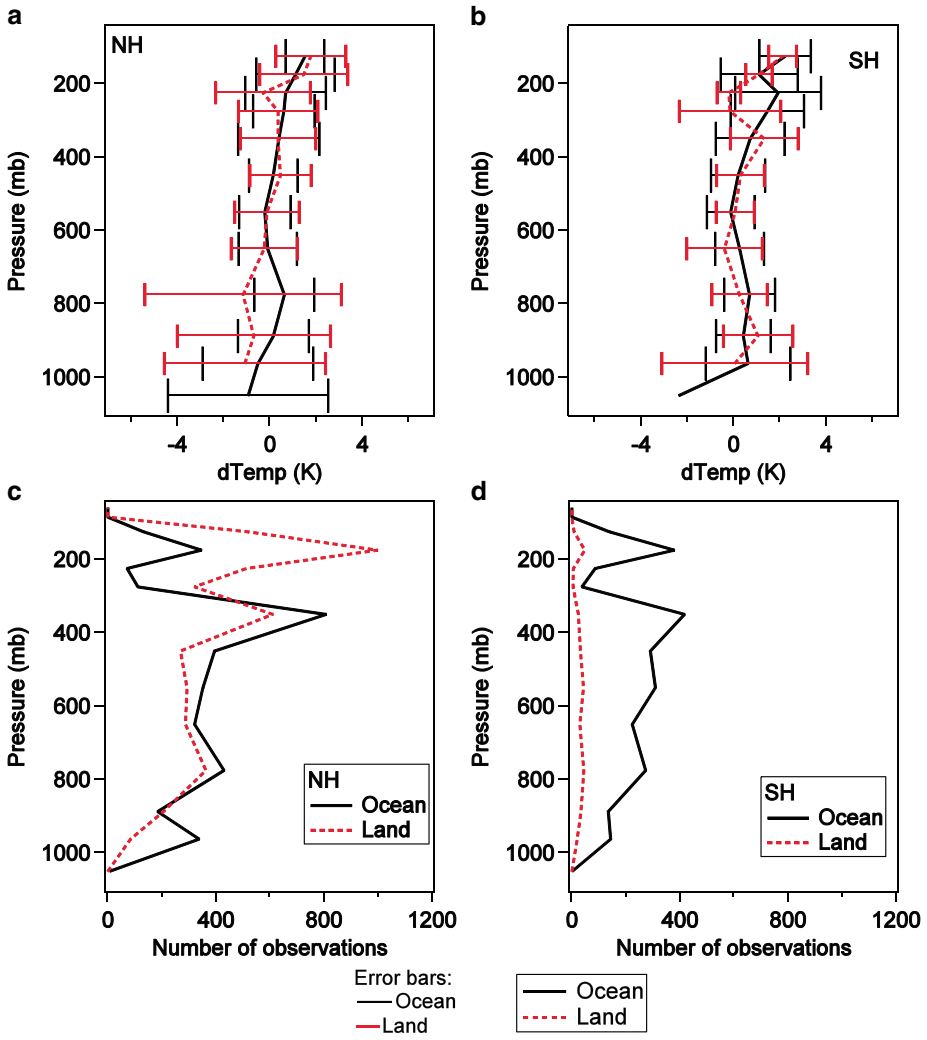

Figure 15. Similar to Figure 14, but for *d*Temp values between the NH and SH.

of sampling. Several factors can contribute to the large differences between the over-land and over-ocean measurements of AIRS/AMSU-A, including the existence of clouds and the intrinsic atmospheric variabilities.

[35] An improved temperature retrieval over land is also shown in Version 5 data compared with Version 4 data at pressure ≤700 mbar. The |*d*Temp| disagreement between land and ocean observations of Version 5 is ~0.3 K at pressure ≤700 mbar (Figure 13a), while the previous results of Version 4 showed that the disagreement in the temperature root mean square values between land and ocean is ~0.5 K. However, large temperature disagreements (~1–1.5 K in |*d*Temp|) between land and ocean measurements still exist at the lower altitudes (≥700 mbar). The *d*Temp profile in Figure 13b shows larger disagreements between land and ocean than |*d*Temp| analyses at pressure ≤300 mbar and ≥700 mbar. Compared with the aircraft observations, the land (and ocean) retrievals have colder (and warmer) temperature of ~ −0.5 to −1 K (0–0.5 K) in the boundary layer, while at the UT/LS, both land and ocean retrievals have warmer temperature by ~0–2 K (0.5–2 K). Overall, the improvements in Version 5 water vapor and temperature retrievals at 800–200 mbar show potential applications of AIRS/AMSU-A data in studying relative humidity, precipitation, and cloud properties over land.

4.6. Water Vapor and Temperature Comparison Between the NH and SH

[36] The dH$_2$Operc and *d*Temp comparison results for the NH and SH are shown in Figures 14 and 15, respectively. Similar to the land and ocean analyses, the NH and SH comparisons were based on the 3 h and 67.5 km window to ensure enough comparison samples. The solid black line and the dotted red line represent the land and ocean observations in each hemisphere, respectively. The aircraft did not sample many over-land regions in the SH, mostly over Australia and New Zealand. Therefore, the large dH$_2$Operc value over land in the SH (Figure 14b) might be due to the lack of sampling. On the other hand, the comparison results over ocean show no significant differences in dH$_2$Operc and *d*Temp between the two hemispheres.

4.7. Matrices of Comparison Results by Various Spatial and Temporal Selection Windows

[37] To illustrate the influences of various spatial and temporal windows, we use contour plots to show the sensitivities of water vapor and temperature comparisons to various *d*Dist

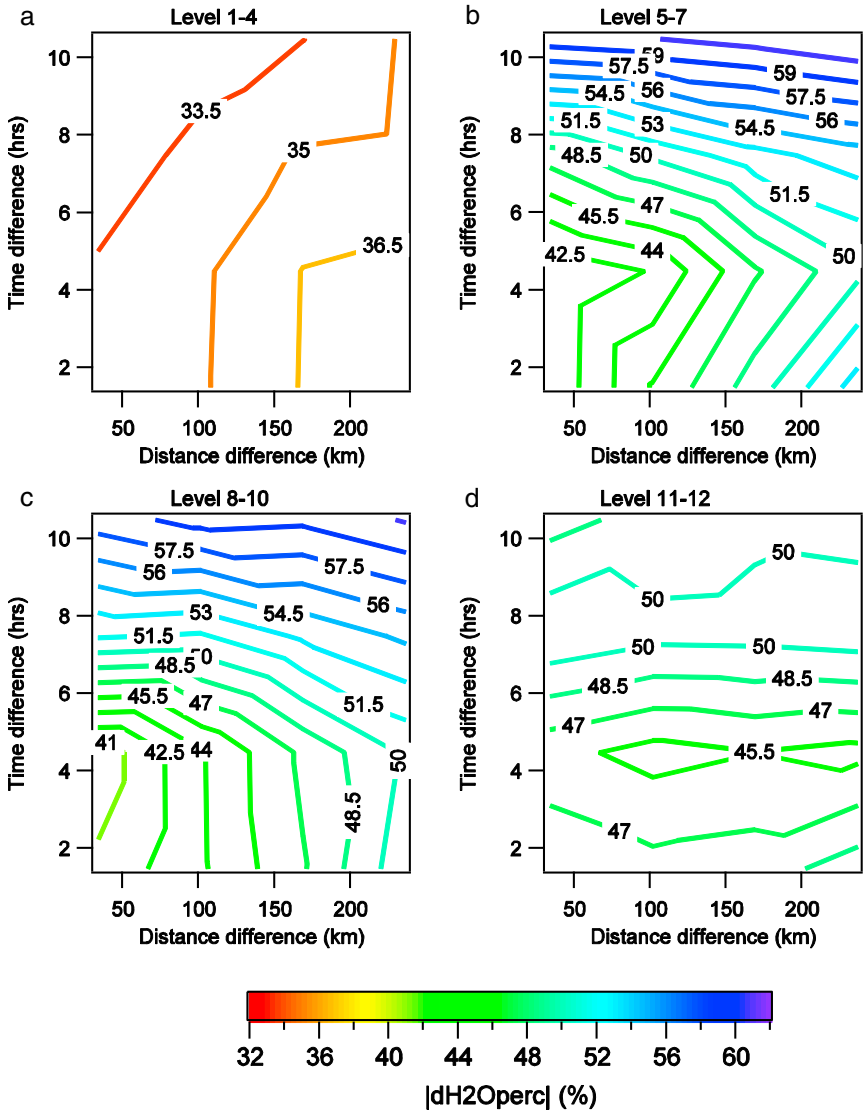

Figure 16. Matrices of water vapor comparison results ($|d\mathrm{H_2Operc}|$) based on various spatial and temporal windows. *d*Time increases from 0 to 12 h by 3 h steps (*y* axis), and *d*Dist increases from 0 to 270 km by 67.5 km steps (*x* axis). The center values of each spatial or temporal window bands were used for plotting the contours (e.g., 1.5 h, 4.5 h, 7.5 h, and 10.5 h; 33.75 km, 101.25 km, 168.75 km, and 236.25 km). The comparisons were separated into four pressure ranges: (a) Levels 1–4 of 1100–700 mbar, (b) Levels 5–7 of 700–400 mbar, (c) Levels 8–10 of 400–200 mbar, and (d) Levels 11–12 of 200–100 mbar. Each gradient is 1/20 of the full color scale.

and *d*Time values (Figures 16 and 17). The purpose of demonstrating the sensitivity of comparison results to *d*Dist and *d*Time is to help constrain the uncertainties coming from the arbitrary selections of spatial and temporal windows in previous validation studies. Here we test *d*Dist from 0 to 270 km (i.e., 12 × 22.5 km) and *d*Time from 0 to 12 h (i.e., 12 × 1 h). The selection of *d*Dist and *d*Time covers most of the temporal and distance window sizes used in previous studies [*Gettelman et al.*, 2004; *Divakarla et al.*, 2006; *Lamquin et al.*, 2012]. The water vapor comparisons were binned into four pressure ranges: (1) Levels 1–4 of 1100–700 mbar, (2) Levels 5–7 of 700–400 mbar, (3) Levels 8–10 of 400–200 mbar, and (4) Levels 11–12 of 200–100 mbar (Figure 16). We note that the pressure bins for temperature are slightly different from those for water vapor, which are (1) Levels 1–4 of 1100–850 mbar, (2) Levels 5–7 of 700–500 mbar, (3) Levels 8–10 of 400–250 mbar, and (4) Levels 11–12 of 200–150 mbar (Figure 17). We use different pressure bins for water vapor and temperature because temperature is retrieved exactly on the labeled pressure level, while water vapor is the mean value between the labeled pressure level and its upper level.

[38] The $|d\mathrm{H_2Operc}|$ values in the contour plots show an increasing trend from the surface to the UT/LS, which is consistent with the analyses in Figure 8a. The gradient of the contour lines is 1/20 of the full color scale for all contour plots in this study. The $|d\mathrm{H_2Operc}|$ values are larger at the higher altitudes because a slight absolute change in water

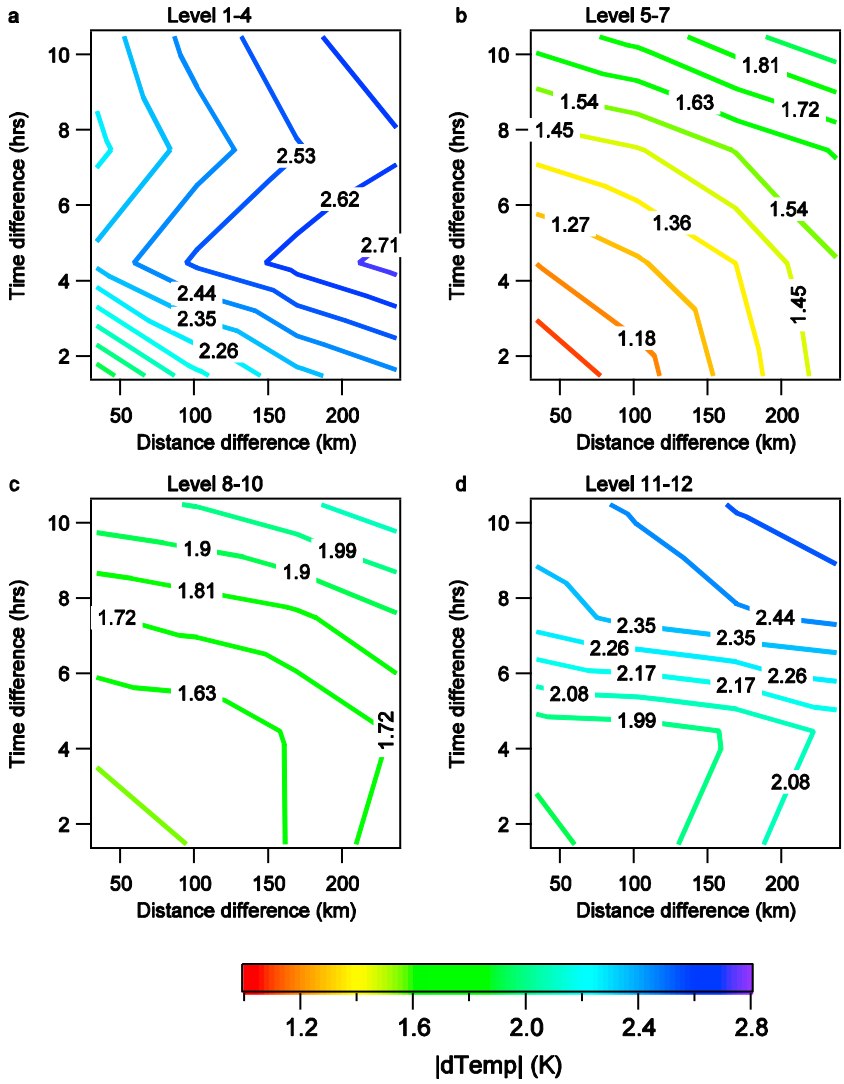

Figure 17. Similar to Figure 16, matrices of temperature comparison results (|*d*Temp|) based on various spatial and temporal windows. Four pressure ranges were analyzed: (a) Levels 1–4 of 1100–850 mbar, (b) Levels 5–7 of 700–500 mbar, (c) Levels 8–10 of 400–250 mbar, and (d) Levels 11–12 of 200–150 mbar. We caution that the pressure ranges for temperature are discontinuous and different from those for water vapor because the AIRS/AMSU-A temperature data are reported as the value on the labeled pressure level and the water vapor data are reported as the mean value between the labeled pressure level and its upper level.

vapor concentration at the UT/LS can lead to a large percentage change. The lower tropospheric |$d$$H_2$Operc| values also have a smaller gradient (fewer contour lines) when compared with the higher altitudes, indicating that the influences of *d*Dist and *d*Time are less obvious at the lower troposphere. Similar analyses on the Pr value of water vapor correlation show that Pr values decrease from ~0.9 to ~0.7 from the surface to the UT/LS, suggesting a weaker correlation at the UT/LS. Compared with water vapor, the Pr values of temperature show a narrower range from ~0.9 to 1.0.

[39] The sensitivity of |$d$$H_2$Operc| is different between ≤4 h and ≥4 h. For comparisons ≤4 h, |$d$$H_2$Operc| is largely influenced by *d*Dist from 1100 to 200 mbar (Figures 16a–16c), and only at 200–100 mbar the influence of *d*Time becomes more dominant (Figure 16d). For comparisons ≥4 h, |$d$$H_2$Operc| is largely influenced by *d*Dist at 1100–700 mbar (Figure 16a), yet *d*Time shows more dominant influence for ≤700 mbar (Figures 16b–16d). Compared with |$d$$H_2$Operc|, |*d*Temp| shows a similar trend at ≥4 h, that is, |*d*Temp| is dominantly influenced by *d*Dist and *d*Time at 1100–700 mbar and ≤700 mbar, respectively (Figure 17). However, |*d*Temp| shows a different trend than |$d$$H_2$Operc| at ≤4 h, that is, |*d*Temp| is equally influenced by *d*Dist and *d*Time from 1100 to 500 mbar (Figures 17a and 17b), while at 400–150 mbar, the influence of *d*Dist becomes more dominant (Figures 17c and 17d). The higher sensitivity of |$d$$H_2$Operc| to *d*Dist than |*d*Temp| at 1100–700 mbar may be a result of the larger variability in water vapor concentrations with respect to topographical variations.

[40] To test the latitudinal influences on the sensitivities of |$d$$H_2$Operc| and |*d*Temp| to spatial and temporal windows, we selected the pressure levels of 5–10 (i.e., 700–200 mbar for water vapor, and 700–250 mbar for temperature) and separated the comparisons into four latitudinal regions as

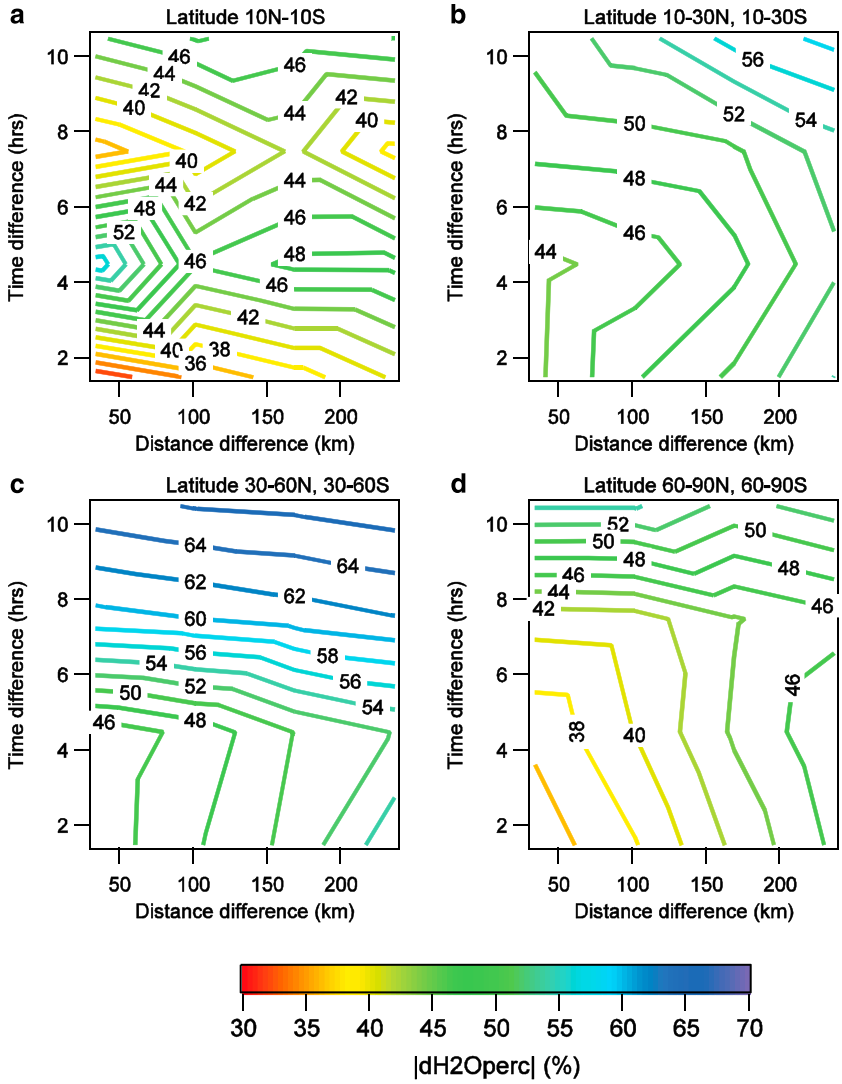

Figure 18. Latitudinal influences on the water vapor comparison sensitivities to *d*Dist and *d*Time. Four latitudinal regions were analyzed: (a) 10°N–10°S, (b) 10°N–30°N and 10°S–30°S, (c) 30°N–60°N and 30°S–60°S, and (d) 60°N–90°N and 60°S–90°S. The $|d\mathrm{H_2Operc}|$ analyses are within water vapor Levels 5–10, which is 700–200 mbar.

shown in Figure 18 and 19: (1) 10°N–10°S, (2) 10° N–30°N and 10°S–30°S, (3) 30°N–60°N and 30° S–60°S, and (4) 60°N–90°N and 60°S–90°S. For both water vapor and temperature comparisons, the sensitivity to *d*Dist relatively increases with respect to the sensitivity to *d*Time, which is consistent with the previous finding that the spatial variabilities of water vapor and temperature increase with latitudes [*Kahn et al.*, 2011]. Water vapor comparisons show better results at the deep tropics (10°N–10°S) and the polar regions (60°N–90°N, 60°S–90°S) than the subtropics and mid-latitudes (10°S–60°S, 10°N–60°N), while temperature shows better results at the tropics (30°N–30°S) than the extratropics (30°N–90°N, 30°S–90°S). Therefore, the major difference between the water vapor and temperature sensitivities across latitudes is at the polar regions, where water vapor (temperature) retrievals show smaller (larger) disagreements with in situ observations compared with the other latitudes.

5. Summary and Implications

5.1. Implications for Validation Studies

[41] In this study, we used a comprehensive data set including in situ measurements over land and ocean in both hemispheres to compare with the water vapor and temperature retrievals of AIRS/AMSU-A. Based on our data set, the $|d\mathrm{H_2Operc}|$ and $|d\mathrm{Temp}|$ values decrease by ~5–10% and ~0.2 K, respectively, when the window sizes decrease from 3 h and 100 km [*Divakarla et al.*, 2006] to 1 h and 22.5 km. Similarly, $|d\mathrm{H_2Operc}|$ and $|d\mathrm{Temp}|$ decrease by ~10–20% and ~0.5 K, respectively, when the window sizes narrow from 12 h and 67.5 km [*Gettelman et al.*, 2004] to 1 h and 22.5 km.

[42] The annular window further quantifies the sensitivities to each band of *d*Dist and *d*Time. The *d*Dist has a larger influence on water vapor and temperature at lower altitudes, which is consistent with previous findings of larger spatial variabilities of temperature and water vapor at lower altitudes [*Kahn et al.*, 2011]. In contrast, at the higher altitudes, *d*Time

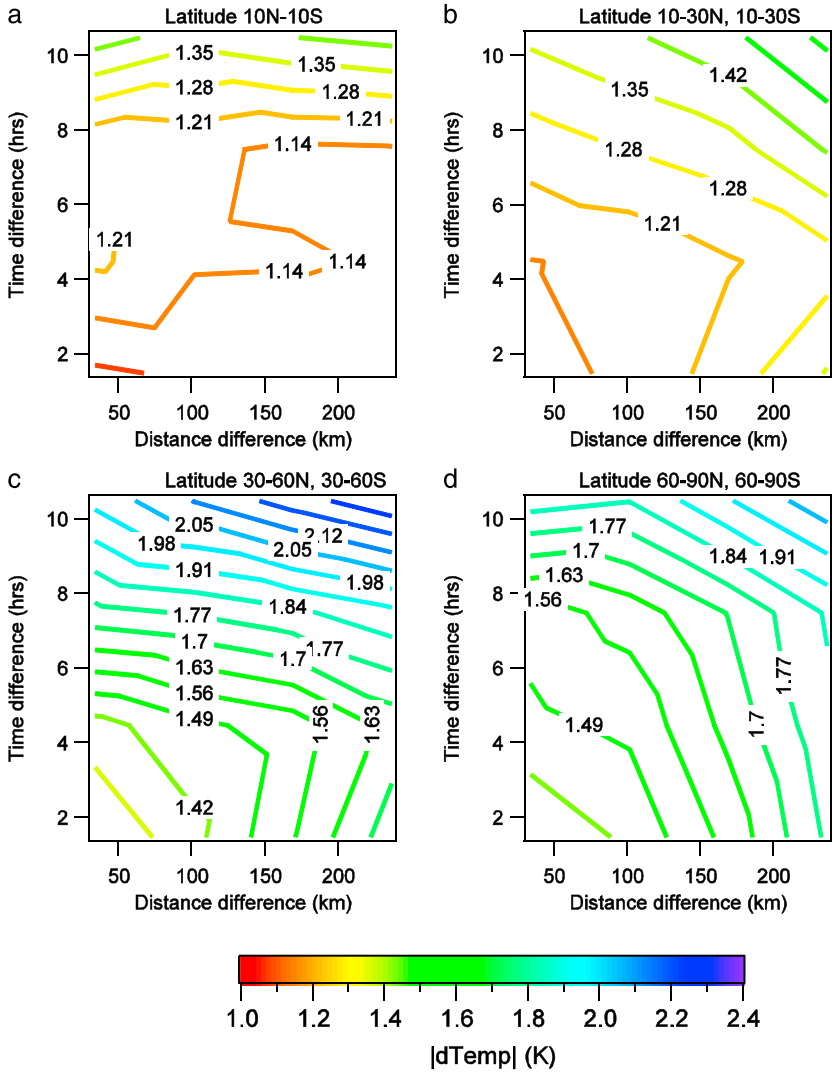

Figure 19. Similar to Figure 18 but for contour plots of |*d*Temp| at four latitudinal regions. All temperature contour plots are within the temperature Levels 5–10, i.e., 700–250 mbar.

has a larger impact on water vapor comparison, which may be a result of the decreasing influences of topography. The H_2O comparison results are most sensitive to *d*Dist at the tropics, which agrees with the peak of the scaling exponent of H_2O at the tropics in the previous study [*Kahn et al.*, 2011]. In addition, the temperature comparison results are most sensitive to *d*Dist at the midlatitudes, which also agrees with the peak of the scaling exponent of temperature at the midlatitudes [*Kahn et al.*, 2011].

[43] The gradient of the $|d\mathrm{H_2Operc}|$ and |*d*Temp| contour plots at various *d*Dist and *d*Time can be used to constrain the influences of window sizes. Overall, for every 22.5 km increase in *d*Dist or every 1 h increase in *d*Time, there is an increase of ~2% in $|d\mathrm{H_2Operc}|$ and ~0.1 K in |*d*Temp|. Both water vapor and temperature comparisons are suggested to be done within 4 h and 100 km in order to keep the errors in $|d\mathrm{H_2Operc}|$ and |*d*Temp| smaller than ~10%. In particular, smaller temporal windows are suggested for water vapor comparisons at the UT/LS because of the higher sensitivities of $|d\mathrm{H_2Operc}|$ to *d*Time at higher altitudes. In addition, smaller spatial windows are suggested for comparisons in extratropical regions because of the higher sensitivities to *d*Dist at higher latitudes.

5.2. Implications for Relative Humidity Studies

[44] Land and ocean comparisons show improvements in Version 5 land retrievals at 800–300 mbar compared with Version 4 [*Divakarla et al.*, 2006]. The improvements are ~5% and 0.2 K in $|d\mathrm{H_2Operc}|$ and |*d*Temp|, respectively. However, the Version 5 data still have large disagreements with the in situ observations in the boundary layer (≥800 mbar) and the UT/LS (≤300 mbar). Therefore, we suggest that future studies should exercise caution regarding the uncertainties in the water vapor and temperature retrievals in these regions. We note that due to the lack of sampling in the boundary layer, we cannot rule out the influence of sampling on the large disagreements between AIRS and in situ observations in the boundary layer. Thus, the following discussions on relative humidity differences between satellite retrievals and in situ observations include the contributions from spatial and temporal variabilities as well as the sampling issues.

[45] For the boundary layer, the land (ocean) retrievals are colder and drier (warmer and moister) compared with the in situ observations (Figures 12b and 13b). Assuming that the surface air temperature is 15°C, the differences

of ~ −0.5 to −1 K in temperature (K) for over-land retrievals can lead to ~3–7% differences in relative humidity estimation. In addition, the ~ −30–0% differences in water vapor concentration (g/kg) lead to up to ~ −30% differences in relative humidity. Similarly, for over-ocean retrievals, the ~ −0.2 to 0.7 K temperature differences and the ~20% water vapor concentration differences lead to ~ −4–1% and ~20% differences in relative humidity estimation at the surface, respectively. Therefore, the relative humidity estimated by satellite data over land (ocean) could have differences of ~ −27–7% (~16–21%) in the boundary layer, combining the contributions from water vapor and temperature differences. These relative humidity differences in the boundary layer will be larger when temperature is lower, such as in the polar regions, and smaller when temperature is higher, such as in the tropics.

[46] For the UT/LS regions, the land (ocean) measurements are warmer and drier (warmer and moister) than the in situ data. Assuming temperature is −40°C in the UT/LS, the over-land relative humidity estimated by satellite data could have differences of ~ −60–0% (i.e., ~ −20–0% contributed from temperature differences and ~ −40–0% from water vapor differences). In contrast, the over-ocean relative humidity will have differences of ~ −20–15% (i.e., ~ −20 to −5% contributed from temperature differences and ~0–20% from water vapor differences). The relative humidity over land in the UT/LS could be underestimated by up to ~60% because both water vapor and temperature contribute to the negative differences in relative humidity. With the large sensitivities of the atmospheric radiative forcing to water vapor concentration in the UT/LS [*Solomon et al*., 2010], the underestimation of relative humidity over land could lead to an underestimation of the greenhouse gas effect of water vapor, as well as underestimations of ice supersaturation and cirrus cloud coverage over land, all of which could lead to an underestimation of heating at the surface [*Held and Soden*, 2000; *Fusina et al*., 2007]. Thus, future studies using AIRS/AMSU-A data on relative humidity and precipitation over land should account for these uncertainties when using the water vapor and temperature data in the boundary layer and the UT/LS.

5.3. Implications for Future Work

[47] In this study, the land observations of the aircraft mostly took place over North America. However, it is not clear if the land measurements of AIRS/AMSU-A over other continents (such as Africa, Asia, and South America) would follow the same trend. Therefore, future comparisons are suggested to provide more insights into other continents in order to help fully assess the over-land measurements of AIRS/AMSU-A. In addition, the latitudinal influence on the comparison results shown in this study is mostly derived from observations over the central Pacific Ocean, that is, from the North Pole to Alaska, Hawaii, New Zealand, and then down to 67°S. However, it is unclear whether the latitudinal influence would be similar in other regions, such as from North America to South America, or over the Atlantic Ocean. Furthermore, more investigation is needed to assess the latitudinal influence on water vapor and temperature comparisons based on the contributions from synoptic-scale dynamical systems, mesoscale meteorological systems, and geographical variations. The quantification and comparison between the influences of various dynamical systems will help to constrain the uncertainties in AIRS/AMSU-A validations at different regions and seasons.

[48] Future work on the influences of spatial and temporal window sizes is suggested to analyze the difference between daytime and nighttime observations. The in situ observations in this study mostly happened during the daytime, which limited our assessment of the nighttime retrievals. In addition, the sensitivities of the water vapor and temperature comparisons to spatial and temporal window sizes may vary around small-scale events, such as deep convection. More case studies are needed to assess the comparison uncertainties around these regions. We note that Version 6 data are now publically available, and future efforts may want to examine them more closely with these types of analyses. Since Version 6 data have improved retrievals of surface emissivity and increased yield of product in the troposphere than Version 5 [*Olsen et al*., 2013], we expect improved comparison results of temperature and moisture especially in the boundary layer. However, our results of comparing various spatial and temporal windows are not expected to change qualitatively whether using Version 5 or Version 6, since the temporal and spatial influences on the differences between satellite and in situ observations mostly represent the influences of intrinsic atmospheric variabilities.

[49] **Acknowledgments.** M. D. thanks the support of the NASA Earth and Space Science Fellowship (NESSF graduate fellowship NNX09AO51H). We acknowledge the NSF support for the START08 and HIPPO Global campaigns from ATM-0840732 and AGS-1036275, respectively. We appreciate the support of the START08 and HIPPO Global science teams and the NCAR EOL RAF flight, mechanical, and technical crews. We also thank the two anonymous reviewers for their helpful comments on the manuscript.

References

Chahine, M. T., et al. (2006), AIRS: Improving weather forecasting and providing new data on greenhouse gases, *Bull. Am. Meteorol. Soc.*, *87*(7), 911–926, doi:10.1175/BAMS-87-7-911.

Cusack, S., J. M. Edwards, and R. Kershaw (2007), Estimating the subgrid variance of saturation, and its parametrization for use in a GCM cloud scheme, *Q. J. Roy. Meteorol. Soc.*, *125*(560), 3057–3076, doi:10.1002/qj.49712556013.

Divakarla, M. G., C. D. Barnet, M. D. Goldberg, L. M. McMillin, E. Maddy, W. Wolf, L. Zhou, and X. Liu (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.

Fasullo, J. T., and K. E. Trenberth (2012), A less cloudy future: The role of subtropical subsidence in climate sensitivity, *Science (New York, N.Y.)*, *338*(6108), 792–4, doi:10.1126/science.1227465.

Ferguson, C. R., and E. F. Wood (2010), An evaluation of satellite remote sensing data products for land surface hydrology: Atmospheric Infrared Sounder, *J. Hydrometerol.*, *11*, 1234–1262, doi:10.1175/2010JHM1217.

Fetzer, E. J., et al. (2008), Comparison of upper tropospheric water vapor observations from the Microwave Limb Sounder and Atmospheric Infrared Sounder, *J. Geophys. Res.*, *113*, D22110, doi:10.1029/2008JD010000.

Fusina, F., P. Spichtinger, and U. Lohmann (2007), Impact of ice supersaturated regions and thin cirrus on radiation in the midlatitudes, *J. Geophys. Res.*, *112*(December), D24S14, doi:10.1029/2007JD008449.

Gao, W., F. Zhao, Y. Xu, and X. Feng (2008), Validation of the surface air temperature products retrieved from the Atmospheric Infrared Sounder over China, *IEEE Trans. Geosci. Remote Sens.*, *46*(6), 1783–1789, doi:10.1109/TGRS.2008.916640.

Gettelman, A., et al. (2004), Validation of Aqua satellite data in the upper troposphere and lower stratosphere with in situ aircraft instruments, *Geophys. Res. Lett.*, *31*, L22107, doi:10.1029/2004GL020730.

Gettelman, A., W. D. Collins, E. J. Fetzer, A. Eldering, F. W. Irion, P. B. Duffy, and G. Bala (2006a), Climatology of upper-tropospheric relative humidity from the Atmospheric Infrared Sounder and implications for climate, *J. Climate*, *19*(23), 6104–6121, doi:10.1175/JCLI3956.1.

Gettelman, A., E. J. Fetzer, A. Eldering, and F. W. Irion (2006b), The global distribution of supersaturation in the upper troposphere from the Atmospheric Infrared Sounder, *J. Climate*, *19*(23), 6089–6103, doi:10.1175/JCLI3955.1.

Goldberg, M. D., Y. Qu, L. M. McMillin, W. Wolf, L. Zhou, and M. G. Divakarla (2003), AIRS near-real-time products and algorithms in support of operational numerical weather prediction, *IEEE Trans. Geosci. Remote Sens.*, *41*(2), 379–389, doi:10.1109/TGRS.2002.808307.

Held, I. M., and B. J. Soden (2000), Water vapor feedback and global warming, *Annu. Rev. Energy Environ.*, *25*(1), 441–475, doi:10.1146/annurev.energy.25.1.441.

Jiang, J. H., et al. (2012), Evaluation of cloud and water vapor simulations in CMIP5 climate models using NASA "A-Train" satellite observations, *J. Geophys. Res.*, *117*, D14105, doi:10.1029/2011JD017237.

Jones, L. A., C. R. Ferguson, J. S. Kimball, K. Zhang, S. T. K. Chan, K. C. McDonald, E. Njoku, and E. Wood (2010), Satellite microwave remote sensing of daily land surface air temperature minima and maxima from AMSR-E, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens*, *3*(1), 111–123, doi:10.1109/JSTARS.2010.2041530.

Kahn, B. H., and J. Teixeira (2009), A global climatology of temperature and water vapor variance scaling from the Atmospheric Infrared Sounder, *J. Climate*, *22*(20), 5558–5576, doi:10.1175/2009JCLI2934.1.

Kahn, B. H., A. Eldering, A. J. Braverman, E. J. Fetzer, J. H. Jiang, E. Fishbein, and D. L. Wu (2007), Toward the characterization of upper tropospheric clouds using Atmospheric Infrared Sounder and Microwave Limb Sounder observations, *J. Geophys. Res.*, *112*, D05202, doi:10.1029/2006JD007336.

Kahn, B. H., A. Gettelman, E. J. Fetzer, A. Eldering, and C. K. Liang (2009), Cloudy and clear-sky relative humidity in the upper troposphere observed by the A-train, *J. Geophys. Res.*, *114*, D00H02, doi:10.1029/2009JD011738.

Kahn, B. H., et al. (2011), Temperature and water vapor variance scaling in global models: Comparisons to satellite and aircraft data, *J. Atmos. Sci.*, *68*(9), 2156–2168, doi:10.1175/2011JAS3737.1.

Lamquin, N., C. J. Stubenrauch, K. Gierens, U. Burkhardt, and H. Smit (2012), A global climatology of upper-tropospheric ice supersaturation occurrence inferred from the Atmospheric Infrared Sounder calibrated by MOZAIC, *Atmos. Chem. Phys.*, *12*(1), 381–405, doi:10.5194/acp-12-381-2012.

Liang, C. K., A. Eldering, F. W. Irion, W. G. Read, E. J. Fetzer, B. H. Kahn, and K.-N. Liou (2010), Characterization of merged AIRS and MLS water vapor sensitivity through integration of averaging kernels and retrievals, *Atmos. Meas. Tech. Discuss.*, *3*(4), 2833–2859, doi:10.5194/amtd-3-2833-2010.

McMillin, L. M., J. Zhao, M. K. Rama Varma Raja, S. I. Gutman, and J. G. Yoe (2007), Radiosonde humidity corrections and potential Atmospheric Infrared Sounder moisture accuracy, *J. Geophys. Res.*, *112*, D13S90, doi:10.1029/2005JD006109.

Miloshevich, L. M., H. Vömel, A. Paukkunen, A. J. Heymsfield, and S. J. Oltmans (2001), Characterization and correction of relative humidity measurements from Vaisala RS80-A radiosondes at cold temperatures, *J. Atmos. Oceanic Tech.*, *18*, 135–156.

Olsen, E. T., E. Fishbein, T. Hearty, S.-Y. Lee, F. W. Irion, B. H. Kahn, and E. Manning (2007a), AIRS Version 5 Release Level 2 Standard Product QuickStart.

Olsen, E. T., E. Fishbein, S. Granger, S.-Y. Lee, E. Manning, and M. Weiler (2007b), AIRS/AMSU/HSB Version 5 Data Release User Guide.

Olsen, E. T., et al. (2013), AIRS/AMSU/HSB Version 6 changes from Version 5.

Pagano, T., M. T. Chahine, H. H. Aumann, B. Lambrigtsen, E. J. Fetzer, E. Olsen, E. Fishbein, C. Thompson, and S. Y. Lee (2004), The atmopheric infrared sounder data products for weather prediction and climate studies, *35th COSPAR Scientific Assembly Held 18 25 July 2004 in Paris France*, 2147.

Pan, L. L., et al. (2010), The stratosphere–troposphere analyses of Regional Transport 2008 Experiment, *Bull. Am. Meteorol. Soc.*, *91*(3), 327–342, doi:10.1175/2009BAMS2865.1.

Peixoto, J. P., and A. H. Oort (1992), *Physics of Climate*, 154, Springe-Verlag, New York, Inc.

Pierce, D. W., T. P. Barnett, E. J. Fetzer, and P. J. Gleckler (2006), Three-dimensional tropospheric water vapor in coupled climate models compared with observations from the AIRS satellite system, *Geophys. Res. Lett.*, *33*, L21701, doi:10.1029/2006GL027060.

Read, W. G., et al. (2007), Aura Microwave Limb Sounder upper tropospheric and lower stratospheric H_2O and relative humidity with respect to ice validation, *J. Geophys. Res.*, *112*, D24S35, doi:10.1029/2007JD008752.

Solomon, S., K. H. Rosenlof, R. W. Portmann, J. S. Daniel, S. M. Davis, T. J. Sanford, and G.-K. Plattner (2010), Contributions of stratospheric water vapor to decadal changes in the rate of global warming, *Science (New York, N.Y.)*, *327*(5970), 1219–23, doi:10.1126/science.1182488.

Susskind, J. (2007), Improved atmospheric soundings and error estimates from analysis of AIRS/AMSU data, *SPIE—Int. Soc. Opt. Eng.*, *6684*, 66840 M–66840 M–12, doi:10.1117/12.734336.

Susskind, J., C. D. Barnet, and J. M. Blaisdell (2003), Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds.

Susskind, J., C. D. Barnet, J. M. Blaisdell, L. Iredell, F. Keita, L. Kouvaris, G. Molnar, and M. T. Chahine (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*, D09S17, doi:10.1029/2005JD006272.

Tian, B., E. J. Fetzer, B. H. Kahn, J. Teixeira, E. Manning, and T. Hearty (2013), Evaluating CMIP5 models using AIRS tropospheric air temperature and specific humidity climatology, *J. Geophys. Res.*, doi:10.1002/jgrd.50117.

Tobin, D. C., H. E. Revercomb, R. O. Knuteson, B. M. Lesht, L. L. Strow, S. E. Hannon, W. F. Feltz, L. A. Moy, E. J. Fetzer, and T. S. Cress (2006), Atmospheric radiation measurement site atmospheric state best estimates for Atmospheric Infrared Sounder temperature and water vapor retrieval validation, *J. Geophys. Res.*, *111*, D09S14, doi:10.1029/2005JD006103.

Verver, G., M. Fujiwara, P. Dolmans, C. Becker, P. Fortuin, and L. Miloshevich (2006), Performance of the Vaisala RS80 A/H and RS90 Humicap sensors and the Meteolabor "Snow White" chilled-mirror hygrometer in Paramaribo, Suriname, *J. Atmos. Oceanic Tech.*, *23*, 1506–1518, doi:10.1175/JTECH1941.1.

Wofsy, S. C., et al. (2011), HIAPER Pole-to-Pole Observations (HIPPO): Fine-grained, global-scale measurements of climatically important atmospheric gases and aerosols, *Philos. Transact. A Math. Phys. Eng. Sci.*, *369*(1943), 2073–86, doi:10.1098/rsta.2010.0313.

Wu, L. (2009), Comparison of atmospheric infrared sounder temperature and relative humidity profiles with NASA African Monsoon Multidisciplinary Analyses (NAMMA) dropsonde observations, *J. Geophys. Res.*, *114*, D19205, doi:10.1029/2009JD012083.

Zondlo, M. A., M. E. Paige, S. M. Massick, and J. A. Silver (2010), Vertical cavity laser hygrometer for the National Science Foundation Gulfstream-V aircraft, *J. Geophys. Res.*, *115*, D20309, doi:10.1029/2010JD014445.