Test Report on the Joint Single Footprint Retrieval Algorithm (JoSFRA) Retrieval System for AIRS Onboard the Aqua Platform

Prepared by: **Sun Wong, Tao Wang, and Jacola Roman**

Contributors: Qing Yue, Evan Manning, Eric Fetzer, Bjorn H. Lambrigtsen, Thomas S. Pagano, Joao Teixeira

Jet Propulsion Laboratory, California Institute of Technology

Product Version 2.0
Document Version 1.0
October 2023

Prepared for the National Aeronautics and Space Administration

Table of Contents

1. Overview

This report assesses the general quality of the Level 2 (L2) products from Version 0.08.6 (herein called V1) of the JOint Single Footprint Retrieval Algorithm (JoSFRA) retrieval system for Atmospheric Infrared Sounder (AIRS) onboard the NASA Aqua satellite (referred to in this report as JoSFRA). This first release of JoSFRA is an early research version intended to solicit user comments.

JoSFRA is a single-footprint IR-only retrieval system for AIRS, which applies an optimal estimation algorithm for simultaneous retrieval of atmospheric temperature, water vapor, surface temperature, cloud-top temperature, effective cloud optical depth, and effective cloud particle radius. Cloud scattering and absorption are in the radiative transfer forward model. AIRS single-footprint thermal infrared L1B radiance data are used directly. Cloud a-priori data comes from coincident MODIS cloud data. The single footprint retrieval improves the horizontal resolution but as microwave data are not used, the retrieval is not performed at altitudes below thick clouds.

The current version of JoSFRA used forecast fields from the National Center for Environmental Prediction (NCEP) as its temperature and water vapor profiles a priori. In retrieval, those NCEP forecast fields are spatially and temporally interpolated to the AIRS field of view (FOV). The retrieval is performed in two steps. In "Step One", simultaneous retrievals of temperature and CO_2 profiles, surface temperature, cloud properties, and emissivity (over land) are performed, mostly on the 14 μm CO_2 band and window regions. Meanwhile, ozone and water vapor profiles are also retrieved as "interferent" gases (The 9.6 μm band of ozone is not used in the retrieval). In "Step Two", the water vapor profile a priori reverts to the NCEP forecast, only the water vapor profile is retrieved from subset channels of water vapor, using the retrievals from Step One as fixed parameters. Accordingly, JoSFRA provides quality control (QC) flags for two steps, with "*qc_flag_step_one*" representing QC for temperature, O_3, CO_2, cloud properties, etc., and "*qc_flag_step_two*" representing QC for specific humidity. More details can be found in the Algorithm Theoretical Basis Document ATBD for JoSFRA [*Irion et al*., 2023].

The AIRS instrument suite was launched on the NASA Aqua platform in 2002 and provides global coverage of the Earth's atmosphere supporting both climate research and weather forecasting. The instrument suite consists of a space-based hyperspectral infrared sounder and two microwave instruments, the Advanced Microwave Sounding Unit (AMSU-A) and the Humidity Sounder for Brazil (HSB). JoSFRA only applies the radiance measurements from the AIRS instrument. AIRS is a cross-track scanning instrument with 2378 infrared channels and 4 visible/near-infrared channels, providing high vertical resolution. This vertical resolution, however, degrades in the presence of clouds.

This report aims to test the performance of JoSFRA L2 V2 retrievals. The data can be downloaded from the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC). The goal is to assess the general quality of several core products of the JoSFRA L2 retrievals through the following analyses:

1) evaluating the retrieval yields and information contents for atmospheric temperature (T) and specific humidity (q or H_2O) profiles;

2) quantifying the biases and root mean square errors (RMSE) of T and H2O retrievals by pixel-scale comparisons collocated to the radiosonde measurements from the Integrated Global Radiosonde Archive to the corresponding AIRS Version 7 (V7) retrievals, and to the 5th

generation global atmospheric reanalysis from the European Centre for Medium-Range Weather Forecasts (ERA5, both land and ocean);

3) the skills of the final retrieval over various climate regimes and scene types including the skill of improvement from the first guess (NCEP CFSR reanalysis) as well as the single footprint skill; and

4) preliminary tests of the JoSFRA retrievals of chemical constituents: O_3 and CO_2.

2. Data and Methodology

2.1 JoSFRA V1 Retrieval Products

This test version of JoSFRA, applied to AIRS onboard Aqua, is v0.08.6, herein V1. More details on the retrieval algorithm can be found in the JoSFRA Retrieval Algorithm Theoretical Basis Document (ATBD, Irion 2023). For reference, AIRS v7 IR Only and IR+MW results will also be shown.

JoSFRA final retrievals of T and H_2O profiles are reported at 46 and 28 basis layers, respectively. The H_2O profiles are specific humidity defined as the mass fraction of water vapor in moist air in units of kg/kg.

JoSFRA has one quality control indicator (qc_flag) associated with both the q and T profiles.
The qc_flag is defined as follows:
QC = 0: Good quality from top-of-atmosphere to surface
QC = 1: Good quality from top-of-atmosphere to a certain pressure level, qc_pres or qc_pres_h2o_vap
QC = 2: Do not use
QC = 3: Retrieval failure
For more information, please see the ATBD [*Irion et al.* 2023].

T and H2O analyses are aggregated by spatial coverage as well as the associated first guess of cloud top temperature and solar zenith angle from the JoSFRA L2 product. This will provide insight into the quality of JoSFRA retrievals in different climate regimes and conditions. All variables presented here are reported on the 130 Along-Track Field Of View (FOV) and 90 Cross-Track FOV spatial dimensions for each granule. The detailed file format and definitions of variables in the JoSFRA files can be found in the product user guide [*Thrastarson et al*. 2023].

2.2 Testing metrics used in the report

The retrievals are evaluated using three metrics: the retrieval yield, retrieval bias, and root mean square differences (RMSE), which will be elaborated below.

- Retrieval yield profiles are calculated as the percent of retrievals falling into each category of quality control flag at different pressure levels.
- Retrieval biases and RMSE are calculated using the equation below where both the truth and the retrievals are filtered with quality flags.

For temperature:

$$bias \equiv mean[x_{data} - x_{truth}]$$

$$RMSE \equiv \sqrt{mean\{[x_{data} - x_{truth}]^2\}}$$

For water vapor:

$$bias \equiv mean[x_{data} - x_{truth}]/mean[x_{truth}] \times 100\%$$

$$RMSE \equiv \sqrt{mean\{[x_{data} - x_{truth}]^2\}}/mean[x_{truth}] \times 100\%$$

- A skill score test was calculated to assess whether the JoSFRA V1 retrieval outperforms the first guess (NCEP Forecast) – when compared to the common reference dataset (ERA5/AIRS V7). Skill score is a forecasting metric that allows a user to test whether one forecast has more skill than another [Murphy, 1988; WMO, 2012]. The formula for the mean squared error skill score is as follows:

$$Skill\ Score = 1 - \frac{MSE_{data}}{MSE_{FG}} = 1 - \frac{mean((T_{data} - T_{reference})^2)}{mean((T_{FG} - T_{reference})^2)}$$

Please note that this metric depends highly on what is considered the reference.

If the skill score is greater than 0, JoSFRA has more skill than the FG (the MSE of the JoSFRA is smaller than the MSE of FG/NCEP forecast); and vice versa. The value of the skill score represents how much the MSE was reduced or increased. For example, a skill score value of 0.25 means JoSFRA reduced the MSE observed in the FG by 25%.

- Since JoSFRA is a single-foot print retrieval algorithm, T and q profiles are retrieved at the resolution of one AIRS pixel (~15 km) compared to the traditional AIRS retrievals on 3x3 pixels at 45 km resolution. We have performed a test on the single footprint skill when we compare JoSFRA retrievals to the IGRA radiosonde measurements. For each 3x3 group of the AIRS footprints, we collocate the one closest and farthest to the matched radiosonde profile and calculate the corresponding uncertainty information (biases and RMSEs). We also compute the mean and median of the 3x3 profiles. The collocated AIRS retrievals should be compared to the mean and median of these 3x3 profiles, while the single-footprint skills are realized as the RMSEs of the closest profiles to the matched radiosonde measurements should be smaller than those of the farthest profiles.

2.3 ERA5 Reanalysis

The ERA5 is the 5th generation global atmospheric reanalysis from the European Centre for Medium-Range Weather Forecasts (ECMWF), replacing the ERA-Interim reanalysis which stopped being produced on August 31st, 2019. Produced using a 4D-Var data assimilation, the ERA5 has a much higher spatial and temporal resolution than its predecessor, the ERA-Interim. In addition, newly reprocessed datasets along with recent instruments have been assimilated into the ERA5 that could not be ingested into the ERA-Interim [*Hennermann*, 2020].

The ERA5 has output at various temporal resolutions. This study made use of hourly output of temperature and specific humidity profiles with 37 pressure levels on a latitude-longitude grid of 0.25° x 0.25°. These gridded products were extracted from the Copernicus Climate Change Service (C3S) [Copernicus 2017] Climate Data Store for two days: January 13th and July 13th of 2011.

Gridded ERA5 (30 km grid or 0.25° x 0.25°) hourly output was collocated to JoSFRA globally for January 13th and July 13th of 2011, using the nearest neighbor approach with the following requirements:

- < 1-hour temporal differences
- < 1° latitude/longitude radius

2.4 IGRA Comparison

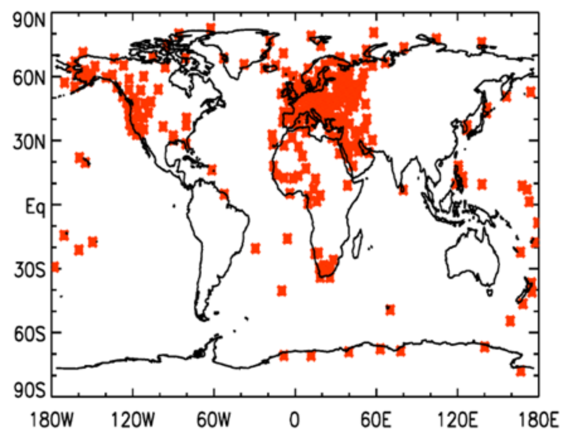

Figure 1. The location of IGRA radiosonde measurements can be matched to the Aqua-platform with collocation criteria of 100km in space and 2 hours in time.

IGRA measurements have been used previously to validate the AIRS Version 6 L2 T and q products (Wong et al. 2015). A similar method is used here. Since measurements from radiosondes are sporadic in space and time, and mostly confined to land, all samples within certain latitude bands are used for validation. The collocation is carried out with a temporal tolerance of within 2 hours, and a spatial tolerance of 100 km from the sonde sites. Statistics are reported in three latitude bands based on the locations of the IGRA sites: the tropics (30°S-30°N), Northern Hemispheric middle latitudes (30°-60°N), and polar region (60°-90°N) for 11th-15th in January and July of 2011. The first guesses, AIRS V7, and the JoSFRA retrievals are filtered all by JoSFRA quality control when the QC $\leq$ 1. Figure 1 shows the location of the radiosondes that can be matched to the Aqua platform overpasses.

3. Results

3.1 Yields and Degree of Freedom for T and H_2O Retrievals

Current JoSFRA level-2 product report quality control (QC) flag for temperature as "*qc_flag_step_one*" and quality control flag for specific humidity as "*qc_flag_step_two*". If the QC flag is 0, the data can be used to surface; if the QC flag is 1, data should be used with caution at pressures greater than "*qc_pres*" for T and "*qc_pres_h2o_vap*" for H_2O, which indicate the lowest pressure level with reliable T/H_2O retrievals. Therefore, we consider all temperature profiles with *qc_flag_step_one* equal to 0/1/2 as retrievals of best/good/bad quality, respectively; similar rules apply to H_2O profiles with *qc_flag_step_two*.

Fig. 1 below demonstrates the global frequencies of each QC flag for both temperature and H_2O during January 13, 2016. Globally, 30-40% of retrievals are of the best quality (QC=0). For temperature, ~32% is considered good retrieval (QC=1), whereas for H_2O this decreases to ~16%, which implies the challenges in retrieving H_2O. This is more emphasized for QC=2, which indicates retrievals of bad quality – it is 10% for temperature but ~34% for H_2O. Note that about

20% of retrievals report QC=3, which means a retrieval failure, i.e., the retrieval encountered an unrecoverable error in the solver (e.g., attempted inversion of a singular matrix.

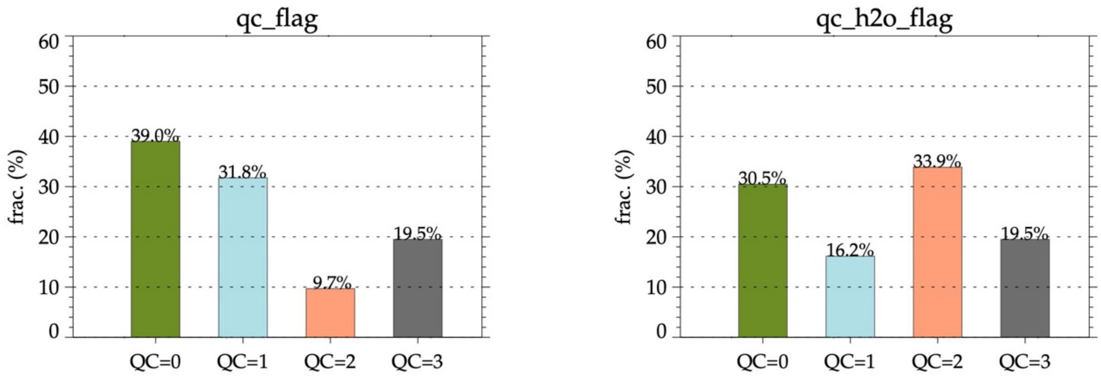

Figure 1. Statistics of quality control (QC) flags for temperature (left, "*qc_flag_step_one*") and for H_2O (right, "*qc_flag_step_two*") during January 13, 2011.

Fig. 2 below shows the map distribution of temperature retrievals passed certain QC filtering on January 13, 2011. For each QC screening, locations with >80% fraction indicate that in those regions, over 80% of temperature retrievals, of all observations, were flagged with that QC flag. Here, QC=0 refers to the best quality retrievals, QC=1 refers to good quality retrievals, and QC=2 refers to the bad retrievals that are not suggested to be used. For temperature, most of the retrievals are considered good and best, but the quality might not be intuitive in the polar cold regions because JoSFRA marks over 80% of the retrievals to have the best quality (QC=0), but usually infrared retrievals are greatly challenged by those cold, icy surfaces given a small thermal contrast between the surface and the air.

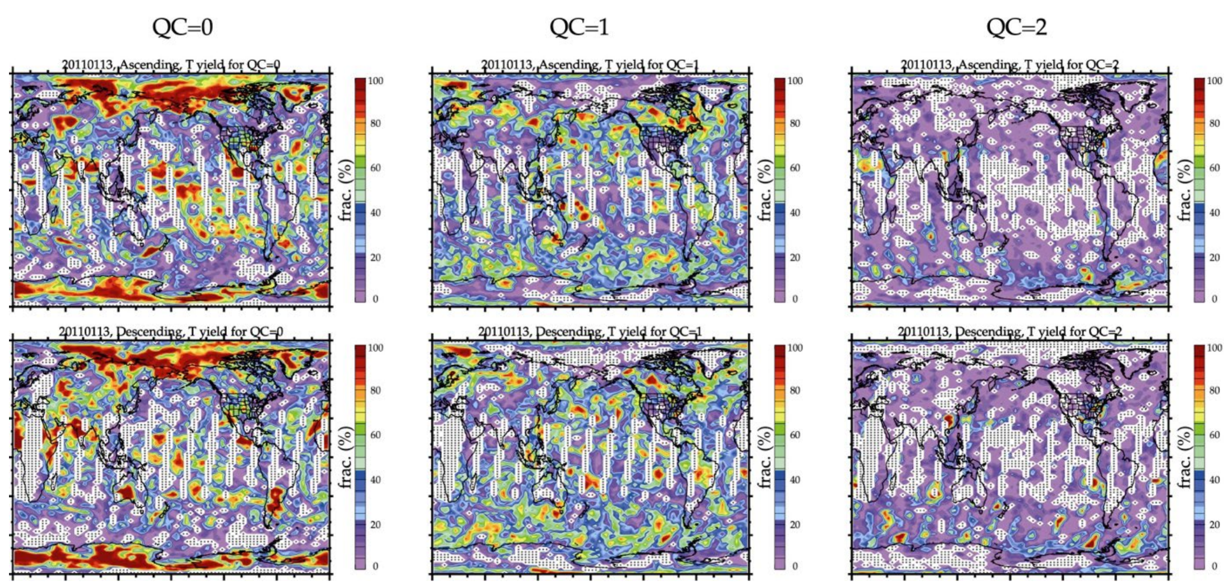

Figure 2. Map distribution of temperature retrievals passed certain QC filtering (each column) for ascending (top row) and descending (bottom row) granules on January 13, 2011. Here, QC=0 and 1 are considered the best and good retrievals, respectively; and QC=2 refers to the retrievals not suggested to be used. For example, locations with a fraction >80% for QC=0 indicate that over 80% of retrievals there were flagged as QC=0 compared to total observations.

Unlike in previous versions the yield bears large differences in different seasons, in this version, we see quite similar yields when examining the July data (not shown). Also, because this version changed to use cloud emissivity to derive cloud optical depths, to bypass the issue of no MODIS clouds to be used during nighttime, we do not see a large contrast between ascending and descending granule results. Therefore, instead of showing similar and redundant results for different dates and for ascending/descending granules, we only show results from a single day (January 13, 2011) for the ascending granules.

Fig. 3 below shows the same but for specific humidity, which shows that the specific humidity retrieval is greatly challenged, largely due to the existence of clouds. This is especially true in the tropical western Pacific, where over 80% of the retrievals were flagged QC=2. Besides clouds, the icy cold surface also affects the quality of specific humidity retrievals.

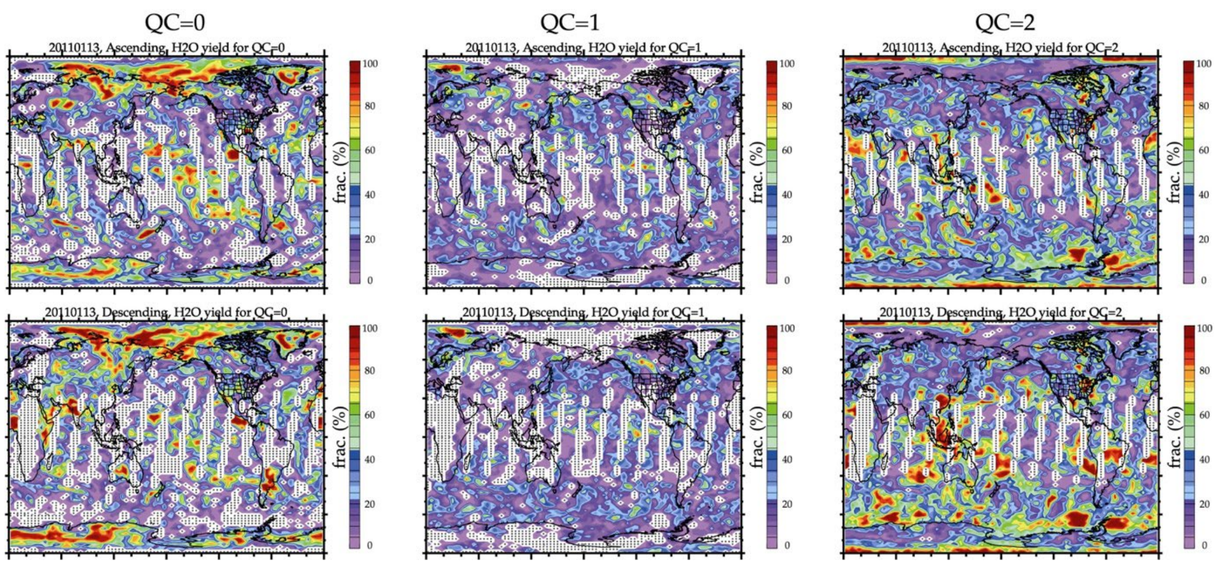

Figure 3. Same as in Fig. 1, but for specific humidity retrievals.

Fig. 4 below shows the degree of freedom (DoF) for temperature and specific humidity retrievals on January 13, 2011, compared to CALIOP mean cloud fractions for the entire month (due to sparse density). Note the different scales for DoF panels. Both temperature and specific humidity retrievals demonstrate overall higher DoFs in the tropics and lower DoFs in middle to higher latitudes. In the tropics, regions of lower DoF indicate less information content from the infrared observations, which is usually caused by the existence of clouds, as indicated by the cloud fractions from CALIOP. In the higher altitudes, low DoF is more likely to be caused by less thermal contrasts.

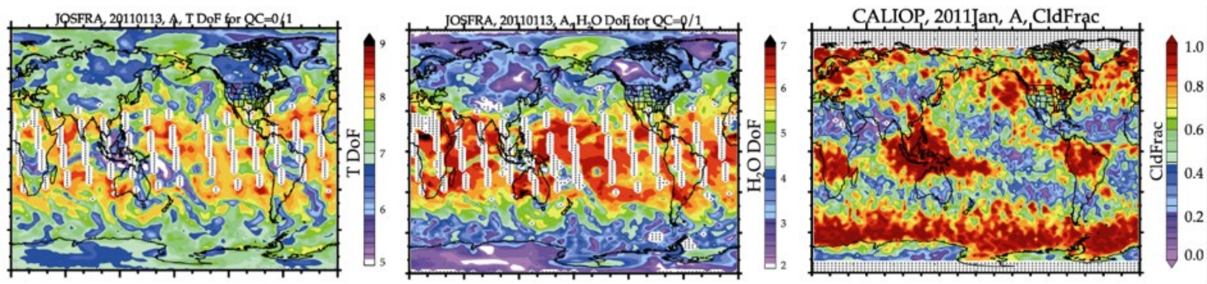

Figure 4. JoSFRA degree of freedom (DoF) for temperature (left) and specific humidity (middle) for January 13, 2011, compared to the January mean cloud fraction inferred from CALIOP (right). Only cases with QC=0 or QC=1 were shown.

3.2 Comparisons to AIRS V7

From now on, we only show results with QC=0 and QC=1. **Fig. 5** below shows an example of JoSFRA temperature retrievals (left column) compared to the same day's data from AIRS version 7 temperature retrievals (middle column) for January 13 and July 13, 2011, respectively. In most of the troposphere and lower stratosphere, JoSFRA temperature retrievals compare well to AIRS, with only ± 0.2-0.5 K differences in general. One exception is over the north polar in January 13, where the differences could be more than 2-3 K. The right column compares JoSFRA with its first guess from NCEP NSFR reanalysis, where we see that JoSFRA retrieval made the most changes near the tropopause (dashed lines) and over the stratosphere (50-1 hPa).

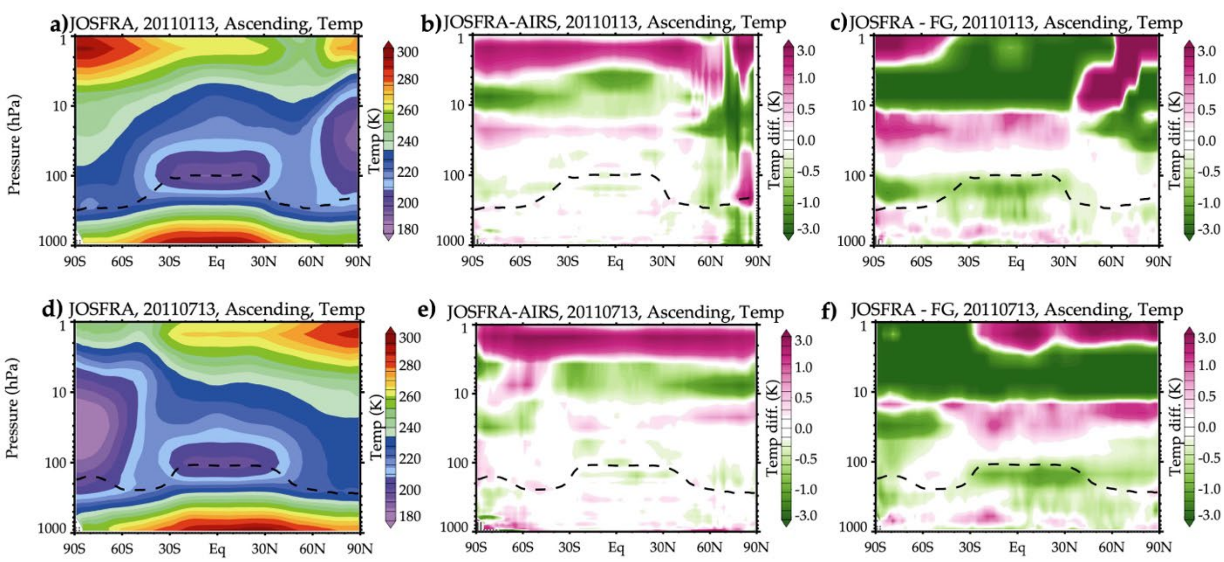

Figure 5. JoSFRA retrieval of temperatures (left column); the difference between JoSFRA and AIRS v7 temperatures (middle column); and the difference between JoSFRA retrieval and its first guess NCEP NSFR reanalysis (right column) for January 13 (upper row) and July 13 (lower row) of 2011. Black dashed lines mark the lapse-rate tropopause. Only results from ascending granules are shown; descending granules show very similar results.

We also perform a skill score test (Section 2.2) to assess whether the final retrieval of temperatures has more skill than its first guess (NCEP CFSR reanalysis) – when comparing to the reference AIRS V7 datasets.

Fig. 6 shows the skill score of temperature retrievals for both ascending and descending granules of January 13 and July 13, 2011, respectively. Here, warm color means the retrieval compares better to AIRS than its first guess does; and vice versa. The first impression is that JoSFRA retrieval of temperatures compares better to AIRS temperatures at the middle to upper stratosphere 2–20 hPa. In 20–40 hPa, the first guess seems to have a smaller MSE. JoSFRA shows more skill near the tropical tropopause (dashed line). In the troposphere, the competition of MSE between the retrieval and the first guess does not show clear preferences.

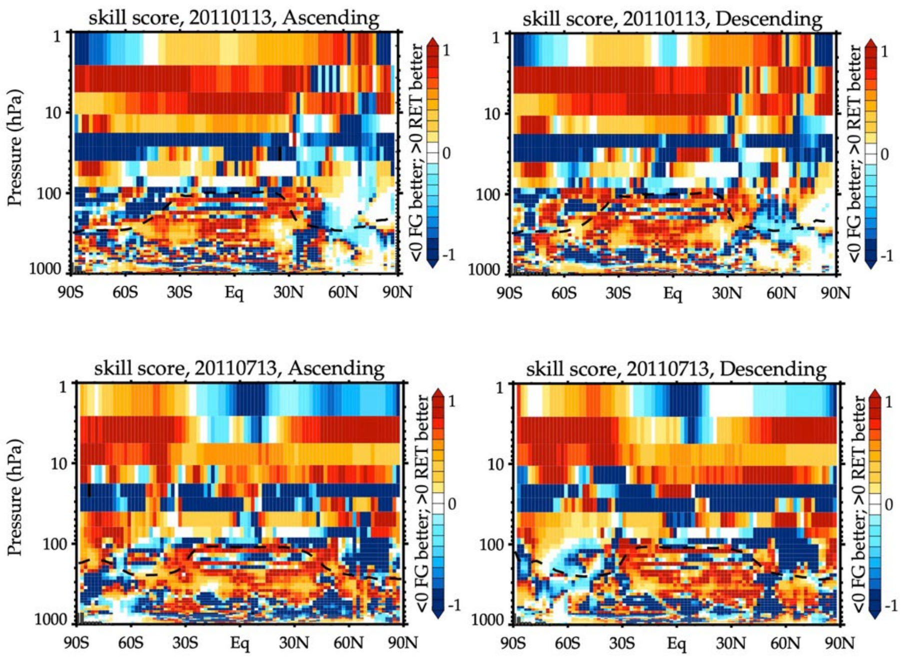

Figure 6. Temperature retrieval skill scores for ascending (left column) and descending (right column) granules on January 13 (upper row) and July 13 (lower row) of 2011. Warm colors suggest that the retrievals have smaller MSE; cold colors indicate smaller MSE in the first guess.

Now that we know where JoSFRA temperature vertical structures compare well to that of AIRS, we can take a closer look at the JoSFRA temperature retrievals on the map. **Fig. 7** shows JoSFRA temperatures compared to AIRS and ERA5 daily mean at four different pressure levels. We see that JoSFRA temperatures broadly agree with both AIRS and ERA5, with a similar transition from warmer tropical lower troposphere (900 hPa) to colder tropical tropopause (~ 100 hPa). Moreover, compared to AIRS, JoSFRA shows advantages in northern higher latitudes with quality retrievals, whereas AIRS has no successful retrievals.

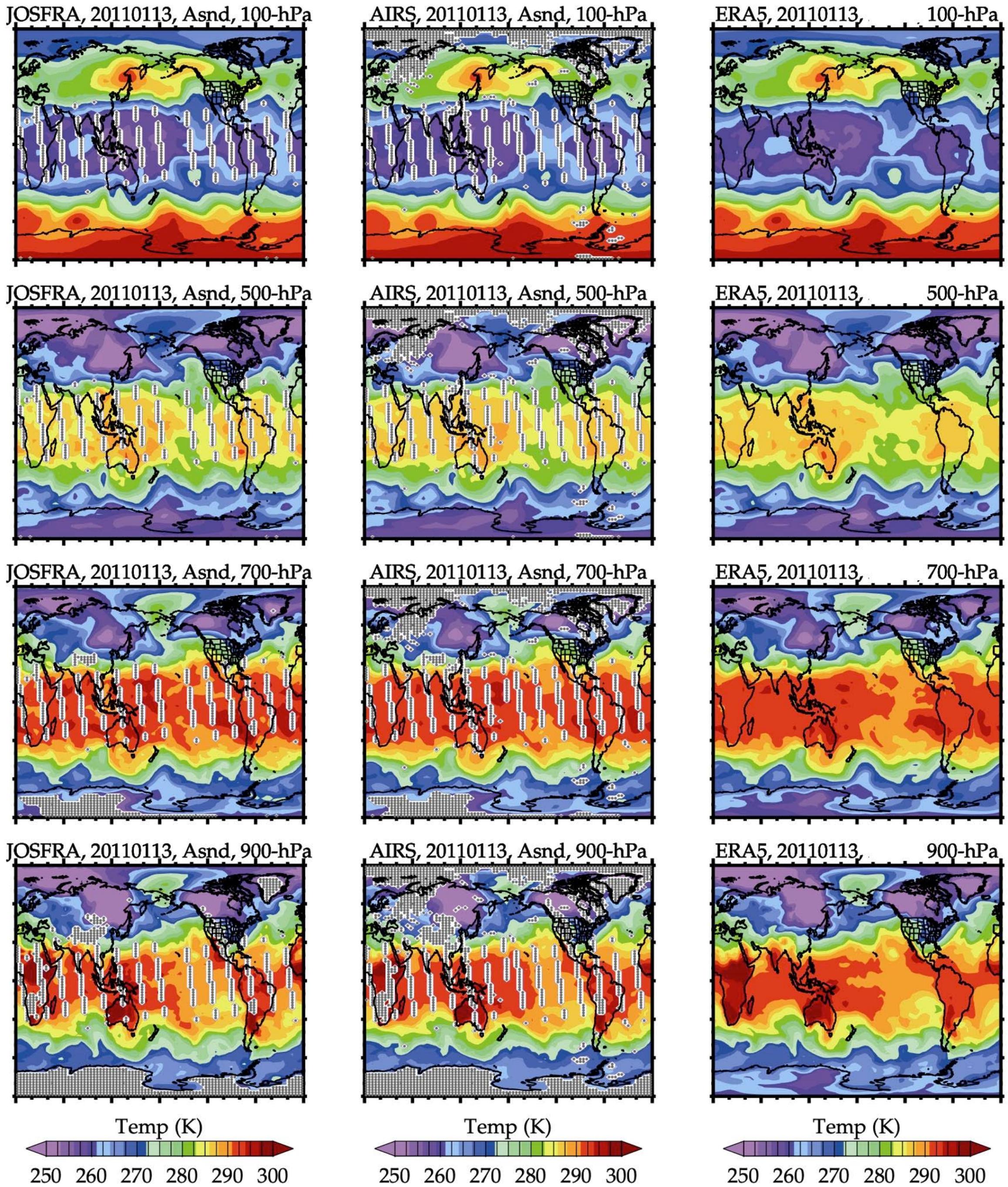

Figure 7. JoSFRA retrieval of temperatures (left column) compared to AIRS (middle) and ERA5 daily mean (right) temperature at 900, 700, 500, and 100 hPa, on January 13 of 2011. Here, we only showed ascending granules of JoSFRA and AIRS retrievals.

Fig. 8 below shows specific humidity compared to AIRS for the same dates. We see a 10-30% large bias in the lower to middle troposphere 400-900 hPa compared to both AIRS and to JoSFRA's first guess, which indicates that AIRS specific humidity should be close to NCEP CFSR reanalysis in those regions. In the upper 100-200 hPa close to the tropopause, JoSFRA shows a 30-50% high bias compared to AIRS, but a 30-50% low bias compared to its first guess.

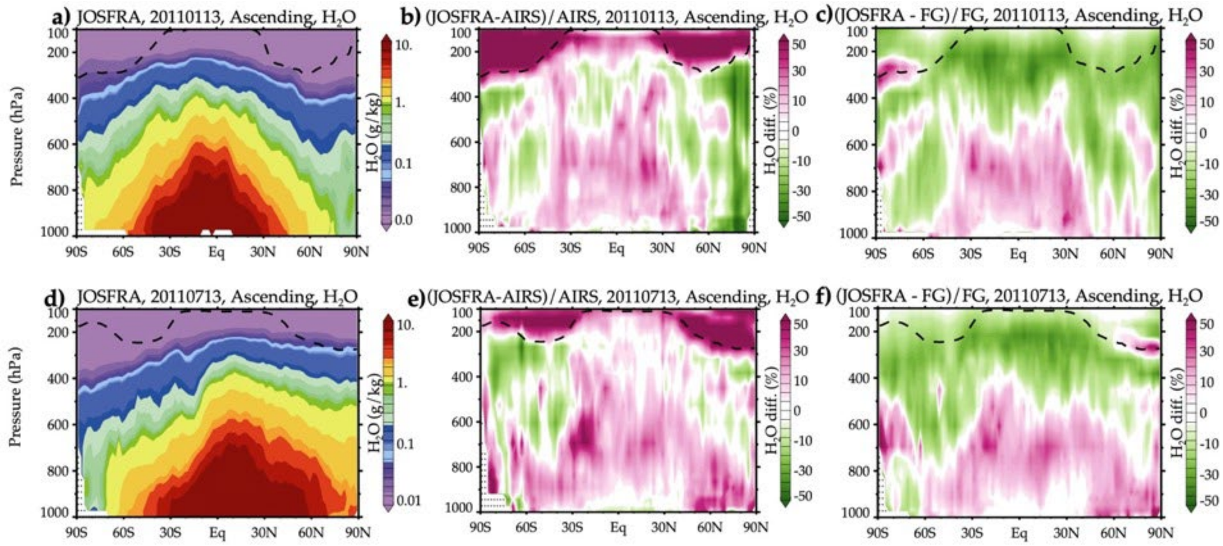

Figure 8. Same as Fig. 5, but for specific humidity.

The skill score for H_2O uses the same formula as shown in Eq. (1). **Fig. 9** shows the skill score of JoSFRA specific humidity retrievals for both ascending and descending granules of January 13 and July 13, 2011, respectively. Here, warm color means the retrieval compares better to AIRS than its first guess does; and vice versa. The first impression is that compared to JoSFRA's first guess, JoSFRA retrieval of specific humidity is closer to that from AIRS at upper troposphere 400 hPa and above levels. Below 400-500 hPa level, JoSFRA's specific humidity retrievals seem less skillful than its first guess.

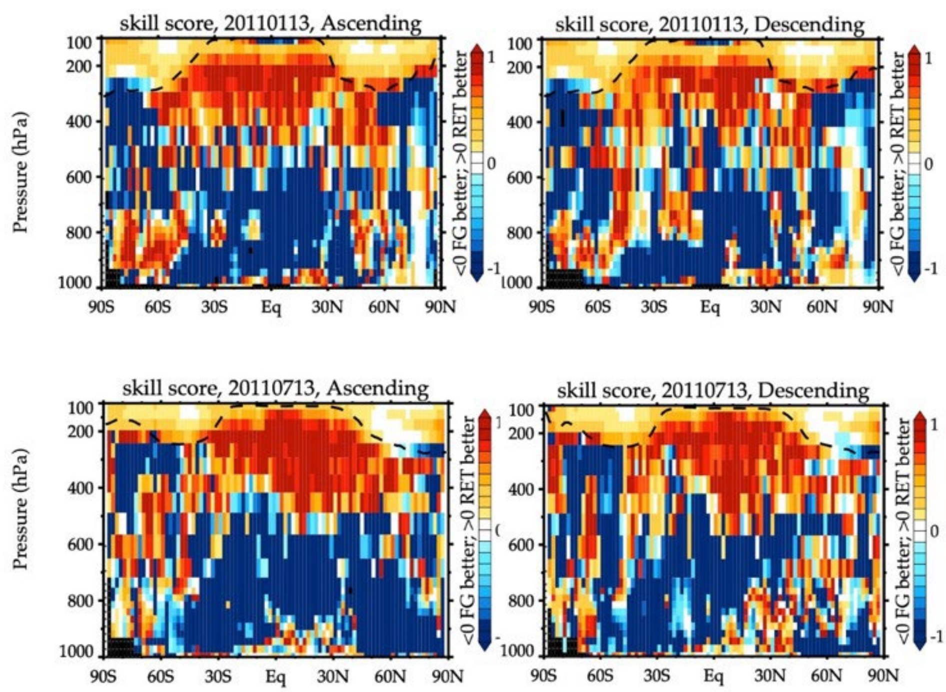

Figure 9. Same as Fig. 6, but for skill score of specific humidity retrievals. Warm colors suggest that the retrievals have smaller MSE; cold colors indicate smaller MSE in the first guess.

Fig. 10 compares JoSFRA specific humidity to AIRS and ERA5 daily mean at four different pressure levels (900, 700, 500, and 300 hPa). We see that JoSFRA specific humidity broadly agree with both AIRS and ERA5, with wetter air in the tropics compared to drier air in the mid- to high latitudes at all levels. Similarly, we see great advantages of JoSFRA over AIRS in that JoSFRA still has lots of successful retrievals in the Northern high latitudes, whereas AIRS loses a fairly number of retrievals that failed in QC.

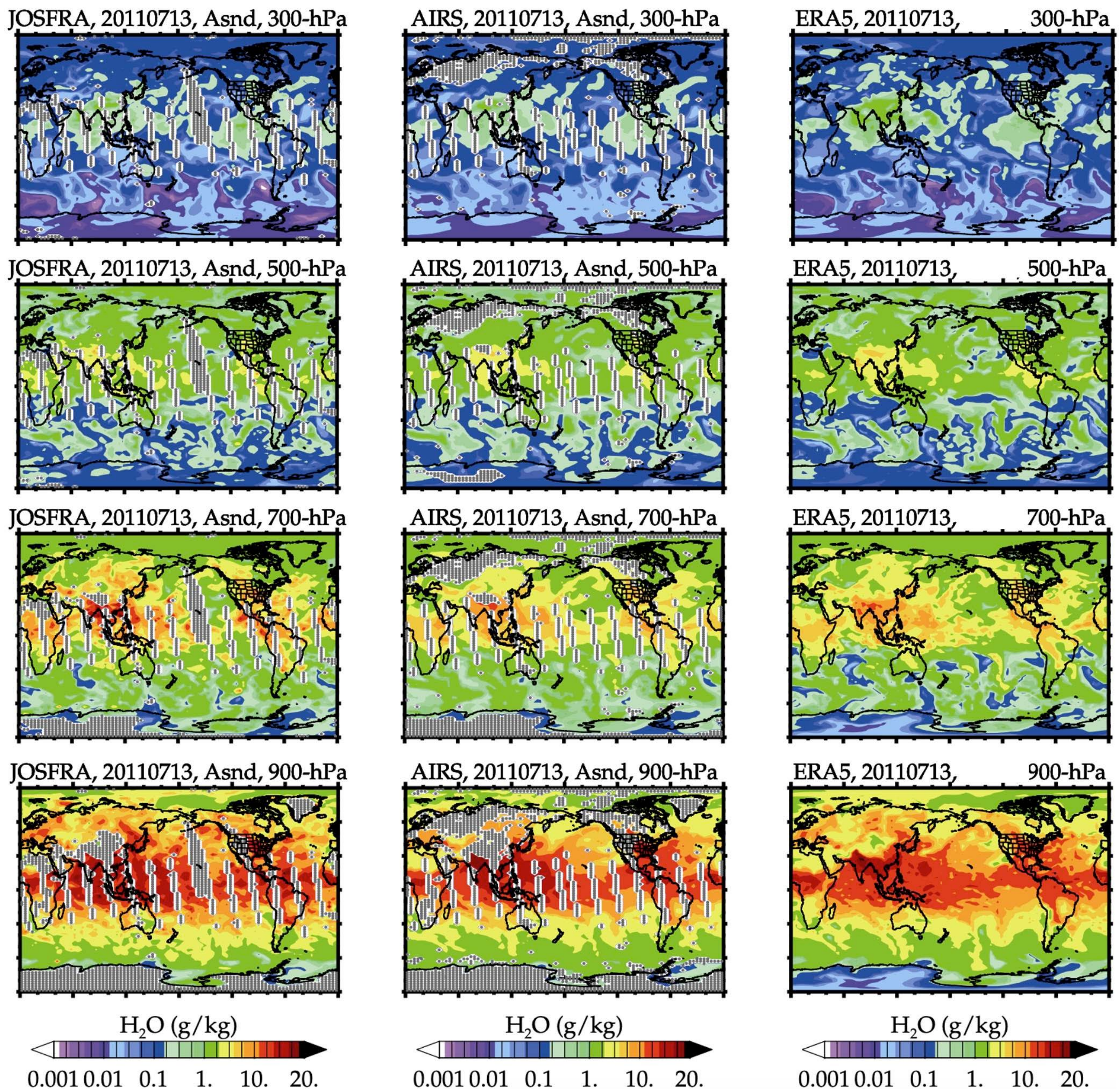

Figure 10. JoSFRA retrieval of specific humidity (left column) compared to AIRS (middle) and ERA5 daily mean (right) temperature at 900, 700, 500, and 300 hPa, on January 13 of 2011. Here, we only showed ascending granules of JoSFRA and AIRS retrievals.

3.3 Comparisons to ERA5

Results presented here highlight the performance of JoSFRA v1 compared to the ERA5 reanalysis. Results will be aggregated by spatial domain, climate regimes and quality control flags.

3.3.1 Global Mean Retrieval Performance

Fig. 11 shows the global mean bias (left), RMSE (2^{nd} column), skill score (3^{rd} column) and yield (right) for temperature by quality control (rows) and month (rows). The final retrieval is the solid line, and the first guess is the dashed line. AIRS v7 IR Only is in green, AIRS v7 IR+MW is in maroon and JoSFRA v1 (v0.08.6) is in blue. JoSFRA has a warm bias less than 0.5 Kelvin in the troposphere and a small cold bias in the stratosphere with respect to the ERA5. In the troposphere, the JoSFRA final retrieval bias is slightly larger than its' first guess and slightly larger in magnitude to AIRS v7. The JoSFRA retrieval has no skill (values less than 0) suggesting the first guess has a smaller MSE compared to the final retrieval. The RMSE for JoSFRA is largest near the surface (close to 2 Kelvin) and smallest in the upper troposphere (close to 1 Kelvin). Overall, about 25% of all retrievals are rejected in JoSFRA and around 50% meet the 'good' quality control (QC=0 or 1). The JoSFRA yields are smaller than the AIRS yields by about 25%.

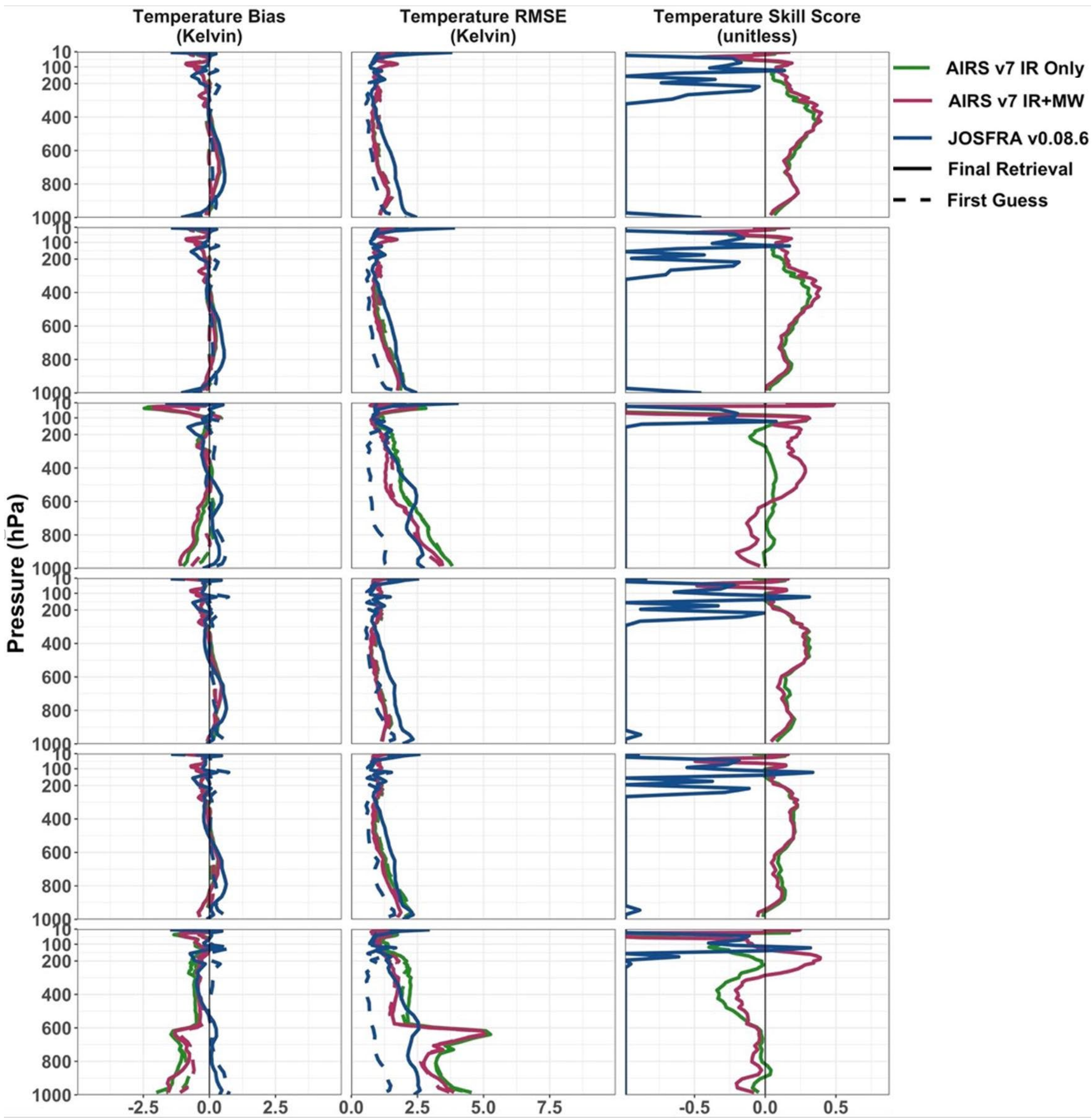

Figure 11. Global mean temperature bias (left), RMSE (2nd column), Skill Score (3rd column), and yield (right) by QC flag and month (rows).

Fig. 12. is the same as **Fig. 11**, but for specific humidity. JoSFRA has a wet bias less than 20% between the surface and 400 hPa and a dry bias between 400 and 200 hPa. Above 200 hPa, JoSFRA has a wet bias. The magnitude of the bias is generally greater than the magnitude observed in AIRS v7. Between 400 and 200 hPa the bias of the first guess is quite wet and the JoSFRA retrieval pulls that first guess drier. Indeed, JoSFRA had significant skill in this altitude range, reducing the overall MSE from the first guess by 30%. The JoSFRA RMSE is less than 75% in the troposphere. About 30% to 40% of all the retrievals are rejected in JoSFRA, compared to less than 25% in AIRS v7. There is no seasonal difference in the results.

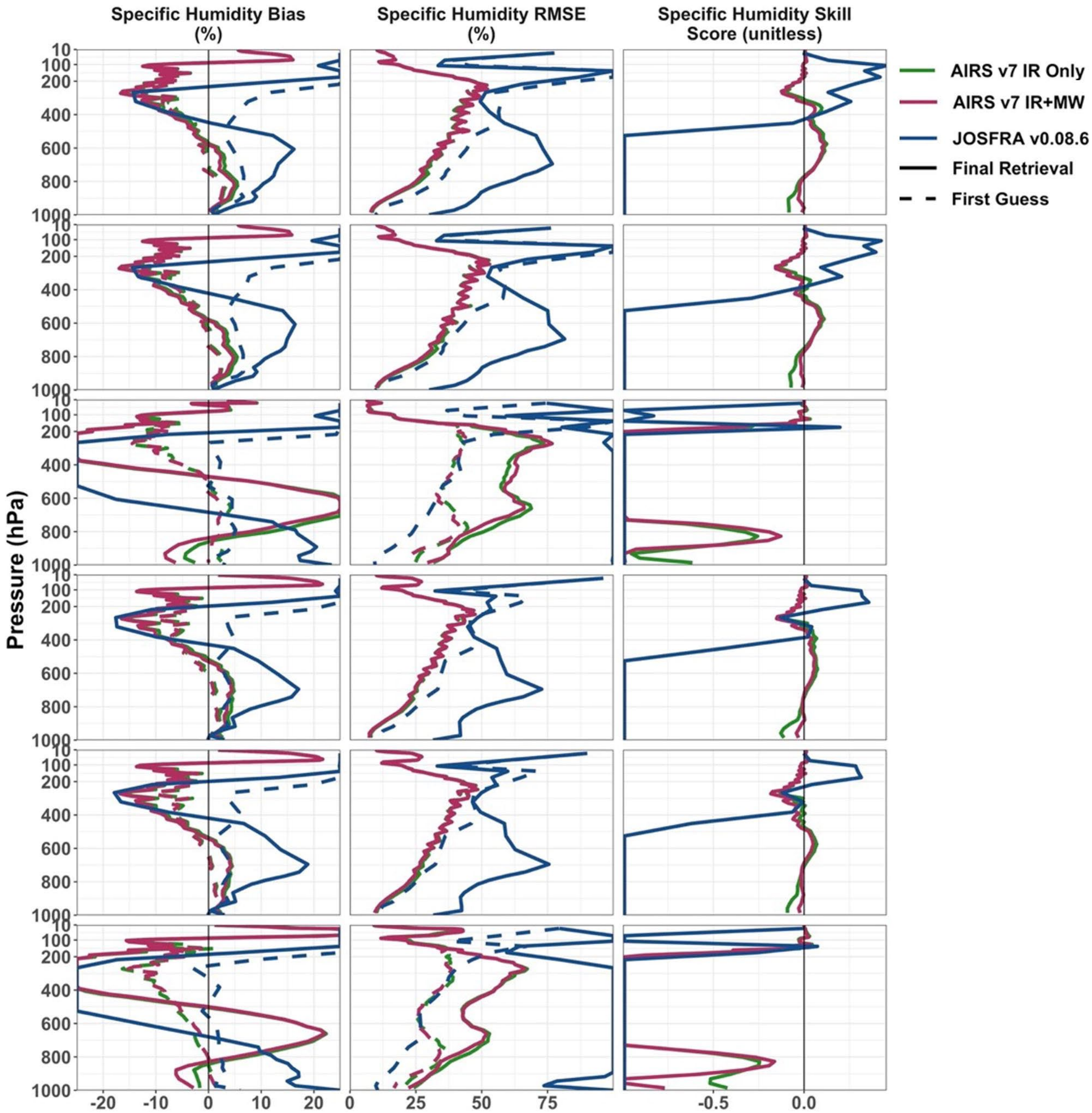

Figure 12. Global mean specific humidity bias (left), RMSE (2nd column), Skill Score (3rd column), and yield (right) by QC flag and month (rows).

3.3.2 Zonal Mean Retrieval Performance

Fig. 13. shows the zonal (aggregated into 5º latitude bins) mean temperature bias (left) and RMSE (right) for JoSFRA and AIRS v7 IR-Only aggregated by the quality control (columns), the month (rows) and the first guess and final retrieval (rows). Regardless of the latitude or height, JoSFRA generally has a warm bias that is typically less than 1 Kelvin and is smallest near the tropopause. The latitudinal dependence is quite different compared to AIRS v7 which has a cool bias throughout the stratosphere and a warm bias in most of the troposphere. The JoSFRA RMSE is typically less than 2 Kelvin in the troposphere and is generally higher than the RMSE observed

in the first guess or AIRS v7. The JoSFRA RMSE is largest when QC=2 suggesting the quality control flags are filtering out bad retrievals.

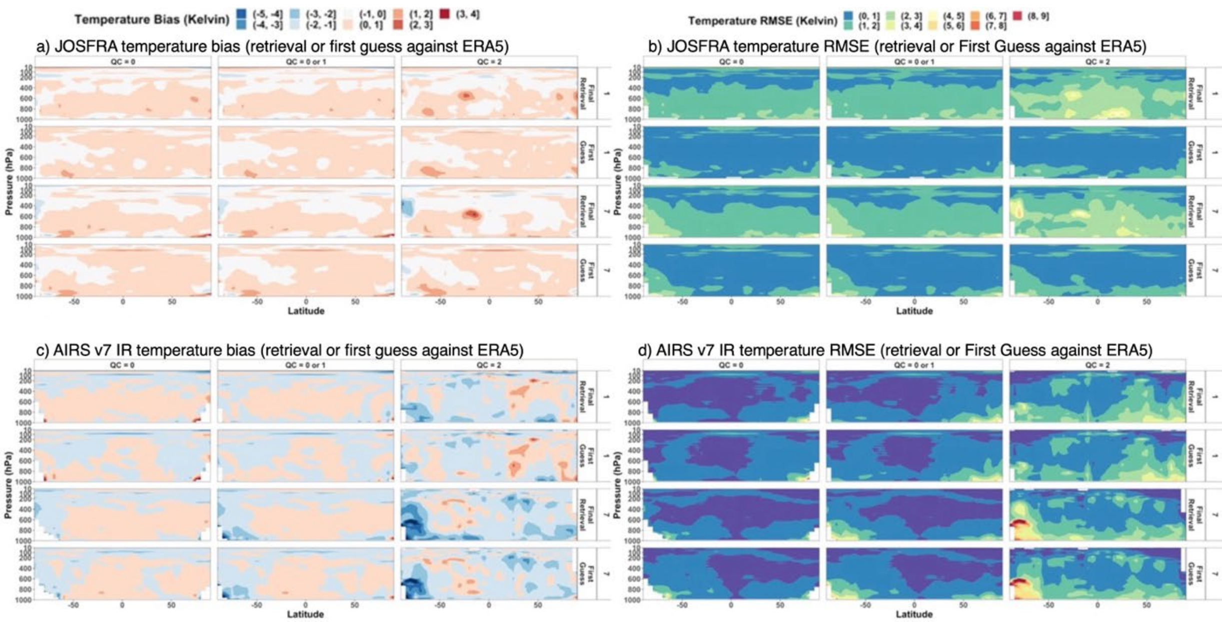

Figure 13. Zonal mean (5-degree latitude bins) temperature bias (left) and RMSE (right) for JoSFRA (top row) and AIRS v7 IR-Only (lower row) aggregated by QC flag (column) and month (row). Here, bias and RMSE for both retrievals and first guess are calculated against collocated ERA5.

Fig. 14 shows the temperature skill score, with positive numbers (red) indicating retrievals improve from the first guess, and negative numbers (blue) indicating retrievals deteriorate from the first guess. Unlike AIRS v7, which improves significantly off its' first guess (a neural net), JoSFRA has little skill with respect to its first guess (NCEP forecast). Furthermore, JoSFRA does not stick to its' first guess, which would have resulted in a skill score of 0, but instead, moves farther away from the first guess (typically too warm). However, near the tropopause in the tropics and extratropics, JoSFRA reduces the overall MSE by 20-40% with respect to its first guess. There is also some skill near the surface in the northern high latitudes during July. A large portion of the retrievals are rejected in the southern mid-latitudes (around 50 to 60ºS) during January and July. More than 60% of the retrievals have a QC=0 in the high latitudes.

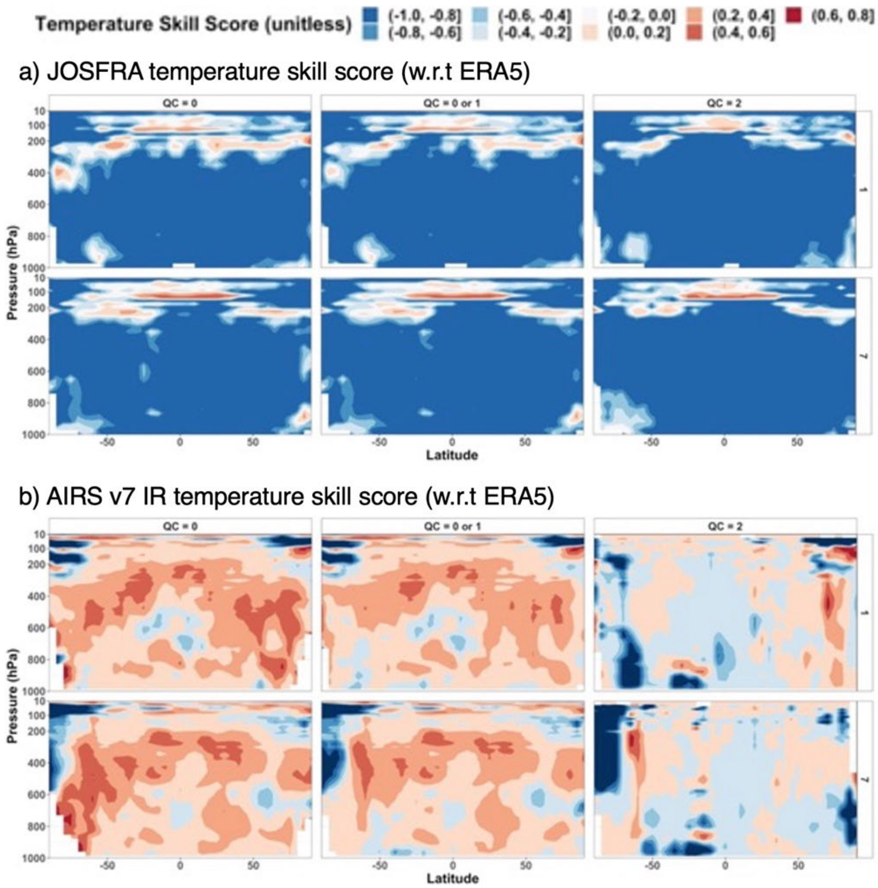

Figure 14. Zonal mean (5-degree latitude bins) temperature skill score for JoSFRA (top row) and AIRS v7 IR-Only (lower row) aggregated by QC flag (column) and month (row).

Fig. 15 is the same as **Fig. 13**. but shows the H_2O bias (left) and RMSE (right). JoSFRA has a 20% wet bias throughout the troposphere. Between 600 hPa and 200 hPa there is a dry bias that exceeds 20%. Throughout the stratosphere, JoSFRA has a strong wet bias that is also observed in the first guess. The JoSFRA RMSE is smallest in the tropics near the surface and largest in the high latitudes below 600 hPa where it typically exceeds 75%. The RMSE of the first guess is usually smaller. One notable exception is between 200 and 100 hPa in the mid to high latitudes. The RMSE is always largest when QC=2 for the final retrieval of JoSFRA; however, the first guess RMSE is smallest when QC=2 suggesting the final retrieval may have moved the first guess too dry (mid to upper troposphere) and too wet (lower tropoposphere), as observed in the bias results.

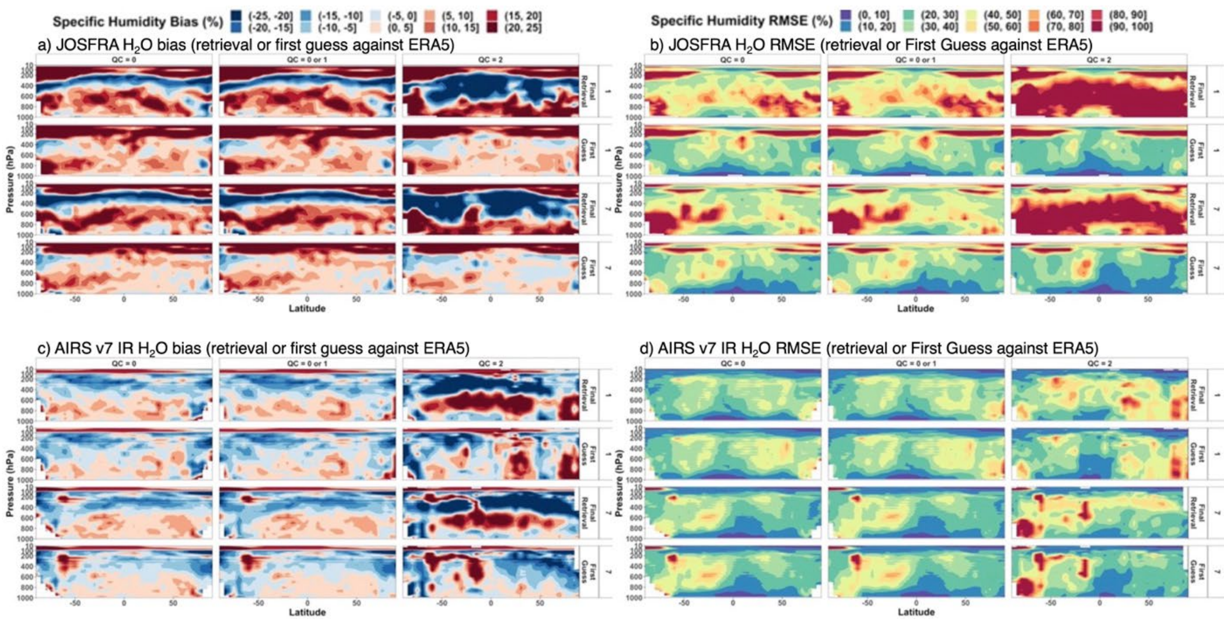

Figure 15. Zonal mean (5-degree latitude bins) H2O bias (left) and RMSE (right) for JoSFRA (top row) and AIRS v7 IR-Only (lower row) aggregated by QC flag (column) and month (row). Here, bias and RMSE for both retrievals and first guess are calculated against collocated ERA5.

Fig. 16 is the same as **Fig. 14**, but compares H_2O skill scores for both JoSFRA and AIRS v7 IR-only when referred to collocated ERA5 results. Unlike AIRS v7, which has some skill compared to its first guess (particularly in the mid to high latitudes), JoSFRA has no skill between the surface and 400 hPa. Above 400 hPa, JoSFRA has skill and reduces the first guess MSE by 20-40% consistently for all QC flags. This was not true for AIRS which saw significantly poorer skill when QC=2. JoSFRA rejects the most retrievals in the southern mid-latitudes (around 50ºS) and retains the most in the high latitudes (northern during January and southern during July).

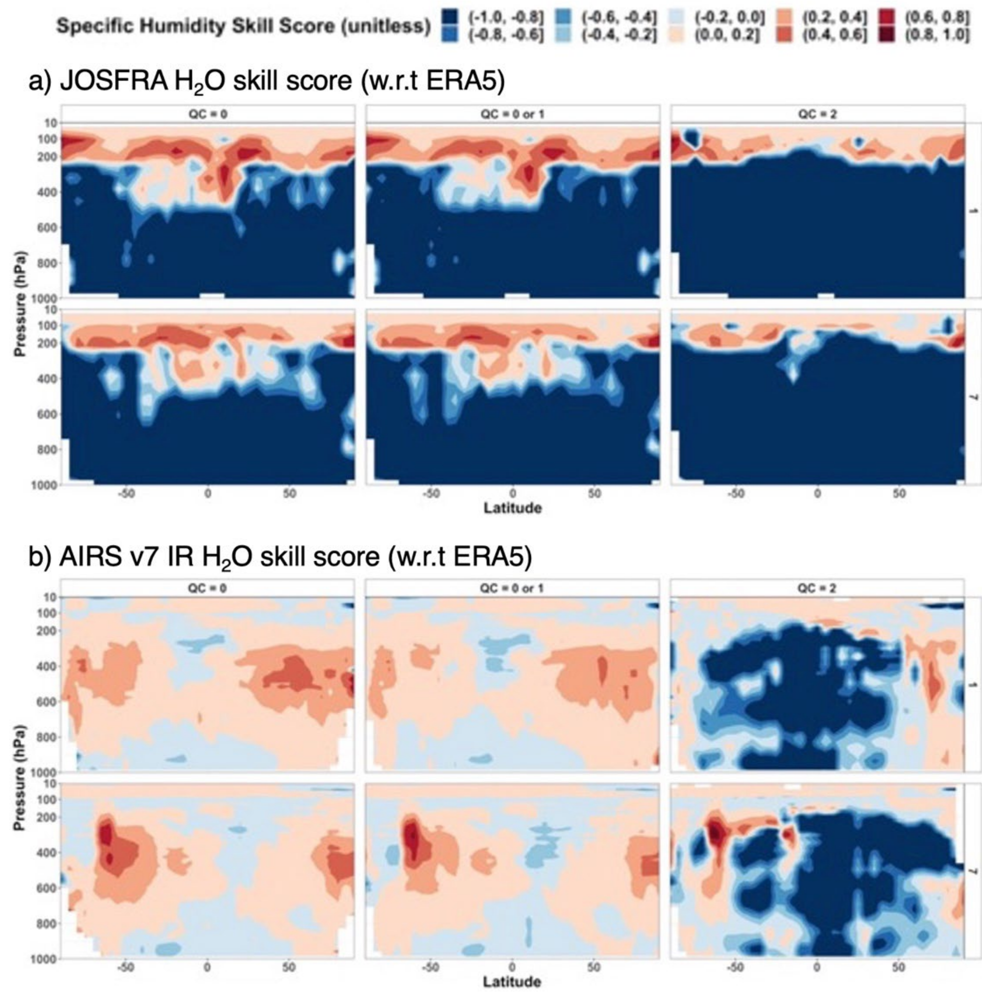

Figure 16. Zonal mean (5-degree latitude bins) H_2O skill score for JoSFRA (top row) and AIRS v7 IR-Only (lower row) aggregated by QC flag (column) and month (row).

3.3.3 Global Maps Retrieval Performance

Fig. 17 shows the temperature a) bias, b) RMSE, and c) Skill Score for JoSFRA when QC = 0 or 1 at 250 hPa (top row), 500 hPa (2nd row), 700 hPa (3rd row), and 850 hPa (bottom) during January of 2011. The magnitude of the bias gets larger and warmer as you go from the top of the atmosphere to the surface. Similarly, the RMSE grows with increasing pressure while the skill score decreases. The yield is smallest at the surface which might impact the metrics. At 250 hPa in January, a strong cold bias is observed over the western Pacific and Australia.

In January JoSFRA has high skill over the Weddell Sea, off the coast of Antarctica. However, this region corresponds to a lower yield which might skew the metrics. At 500 hPa in January, JoSFRA has a strong warm bias over Greenland and a strong cold bias over the western portion of the United States and Eastern Russia. At 700 hPa during January, JoSFRA has a strong warm bias over Greenland and over Gulf of Guinea, off the coast of Guinea and Ghana, while a strong cold bias is observed over the far eastern part of Antarctica. In January JoSFRA also shows low RMSE east of the Caribbean that extends northeast into the Atlantic Ocean, which corresponds to a high skill score. At 850 hPa, JoSFRA has strong cold bias over the Artic Ocean in January as well as a high RMSE and high yield.

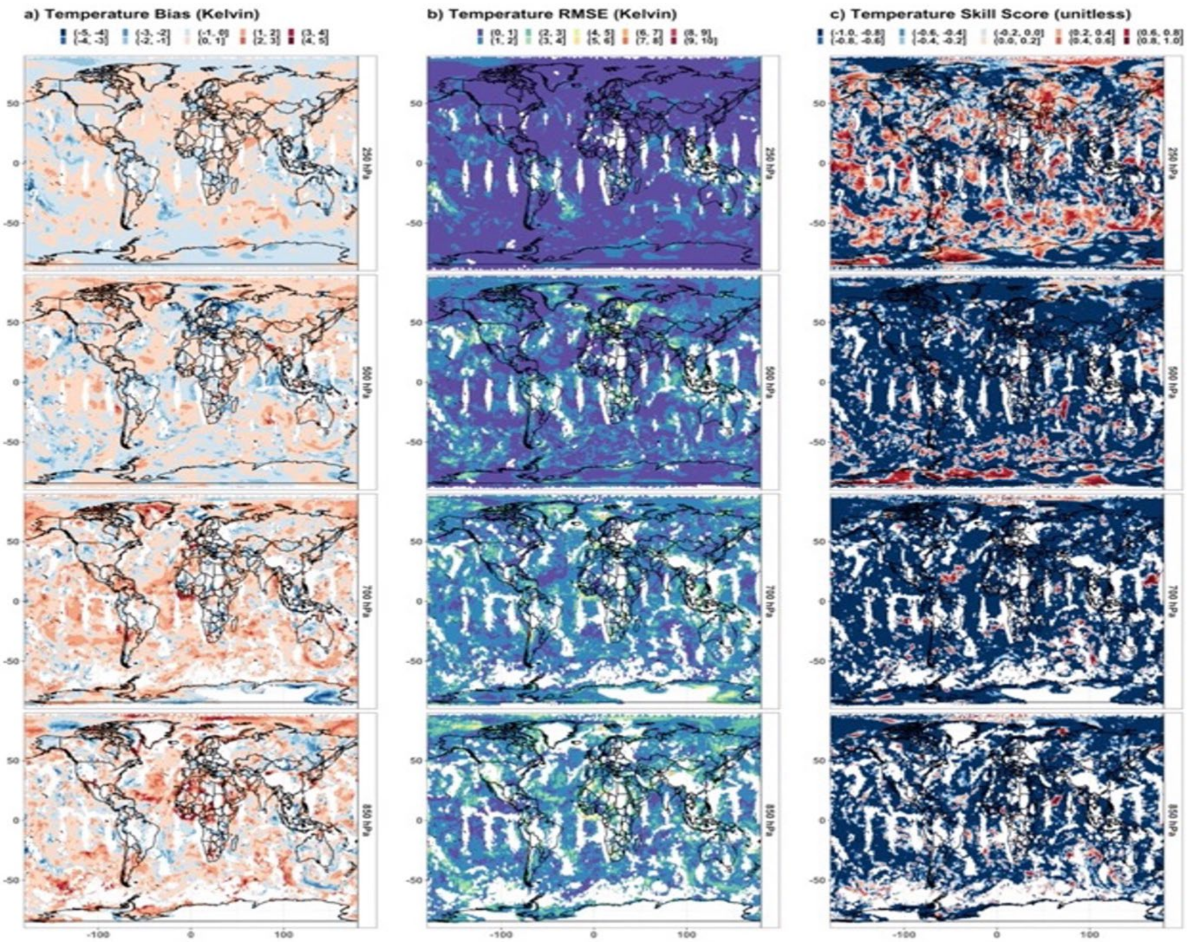

Figure 17. Global map of temperature bias (left), RMSE (2nd column), and skill score (3rd column) for JoSFRA when QC=0 or 1 at 250 hPa (top row), 500 hPa (2nd row), 700 hPa (3rd row), and 850 hPa (bottom row). This is data for January of 2011.

Fig. 18 shows similar but for July for 2011. In July, a strong warm bias is seen over Greenland. These regions also corresponded to high RMSE values (exceeding 2 Kelvin). Meanwhile, JoSFRA has a strong warm bias over the eastern part of Antarctica a strong cold bias over the rest of Antarctica. These regions are also dominated by high RMSEs. During July at 700 hPa, there is a strong warm bias extending off the coast of the Baja Peninsula and a region off the coast of Venezuela that extends northeast across the Atlantic. These regions correspond to regions of high RMSE and low skill. There is a strong warm bias over the Northern portion of Africa with corresponding high RMSEs. Meanwhile, a band of low RMSEs that extends eastward from the Artic Ocean shows high skill.

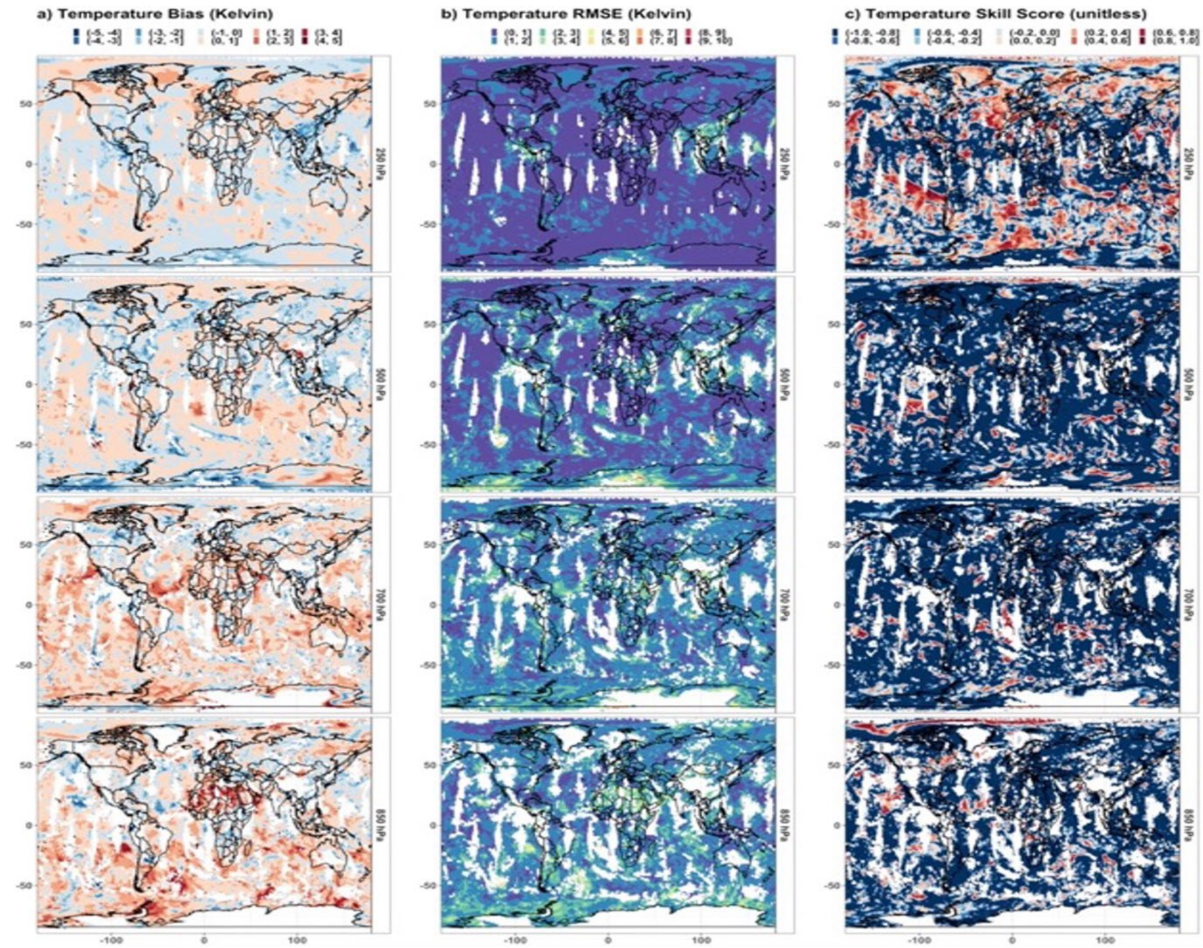

Figure 18. Same as Fig. 17, but for July of 2011.

Fig. 19 and **Fig. 20** are the same as **Fig. 17** and **Fig. 18,** but for specific humidity. At 250 hPa JoSFRA shows a strong seasonal cycle in the bias and RMSE. During January, there is a strong wet bias in the southern high latitudes while in July this strong wet bias is observed in the north. At 250 hPa in January, JoSFRA has high skill over Alaska and Northwestern Canada, as well as over North central Russia. At 500 hPa during January, there is high RMSE in the Caribbean and over the east central portion of Africa while in July there are regions of high RMSE over Southwestern Brazil. These regions correspond to low yields which might suggest that the sample size could explain some of the large RMSE values. During January at 700 hPa JoSFRA has high RMSE over a region east of the Caribbean that extends northeast into the Atlantic. In July at 700 hPa, high RMSE is consistently observed around 30ºS where we also see strong wet biases. At 850 hPa, high RMSEs are consistently correlated to strong wet biases. During July, JoSFRA has high skill scores north of Alaska in the Arctic Ocean. This region has low RMSEs, high yields, and a strong dry bias. This suggests that the first guess might poorly represent the region and JoSFRA improves off initial first guess.

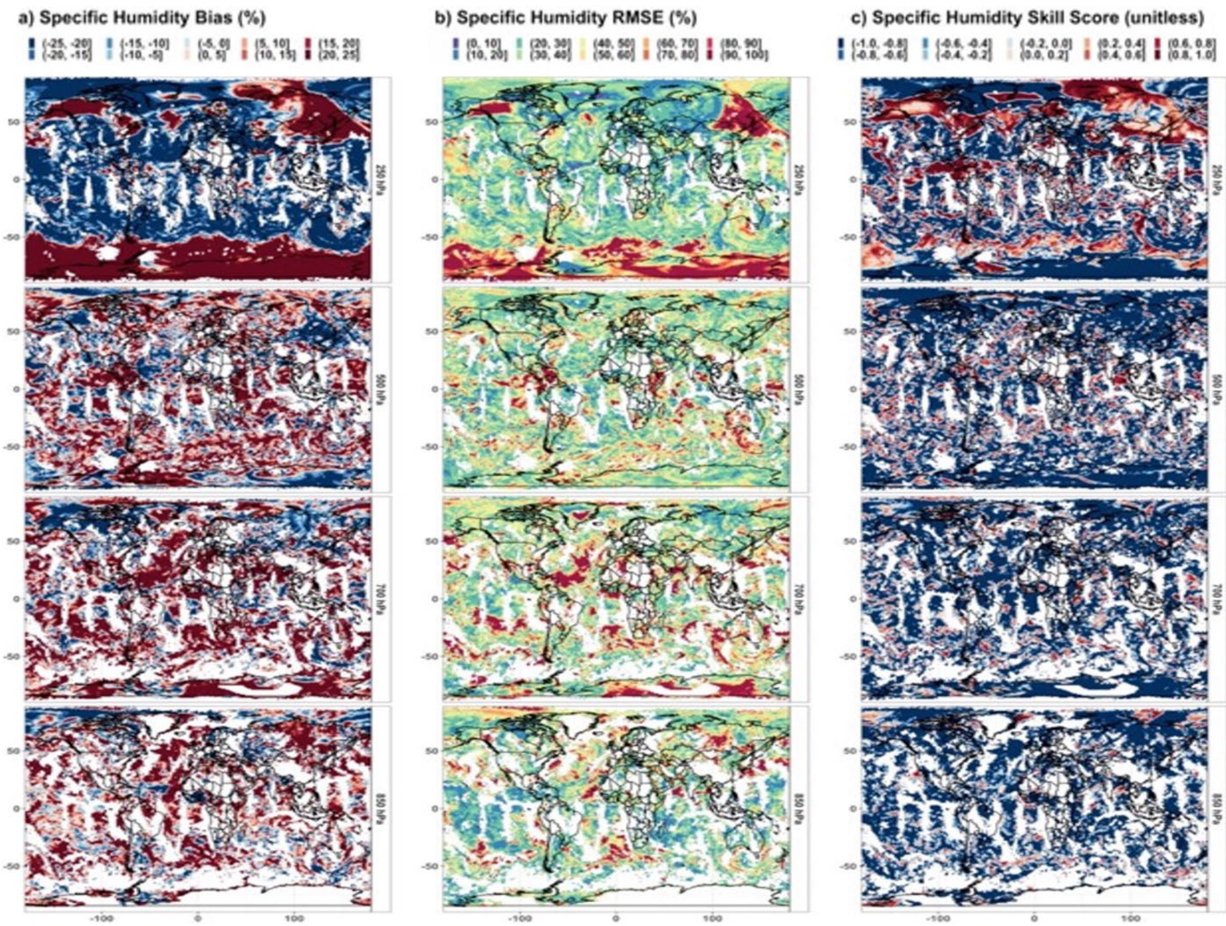

Figure 19. Global map of H_2O bias (left), RMSE (2nd column), and skill score (3rd column) for JoSFRA when QC=0 or 1 at 250 hPa (top row), 500 hPa (2nd row), 700 hPa (3rd row), and 850 hPa (bottom row). This is data for January of 2011.

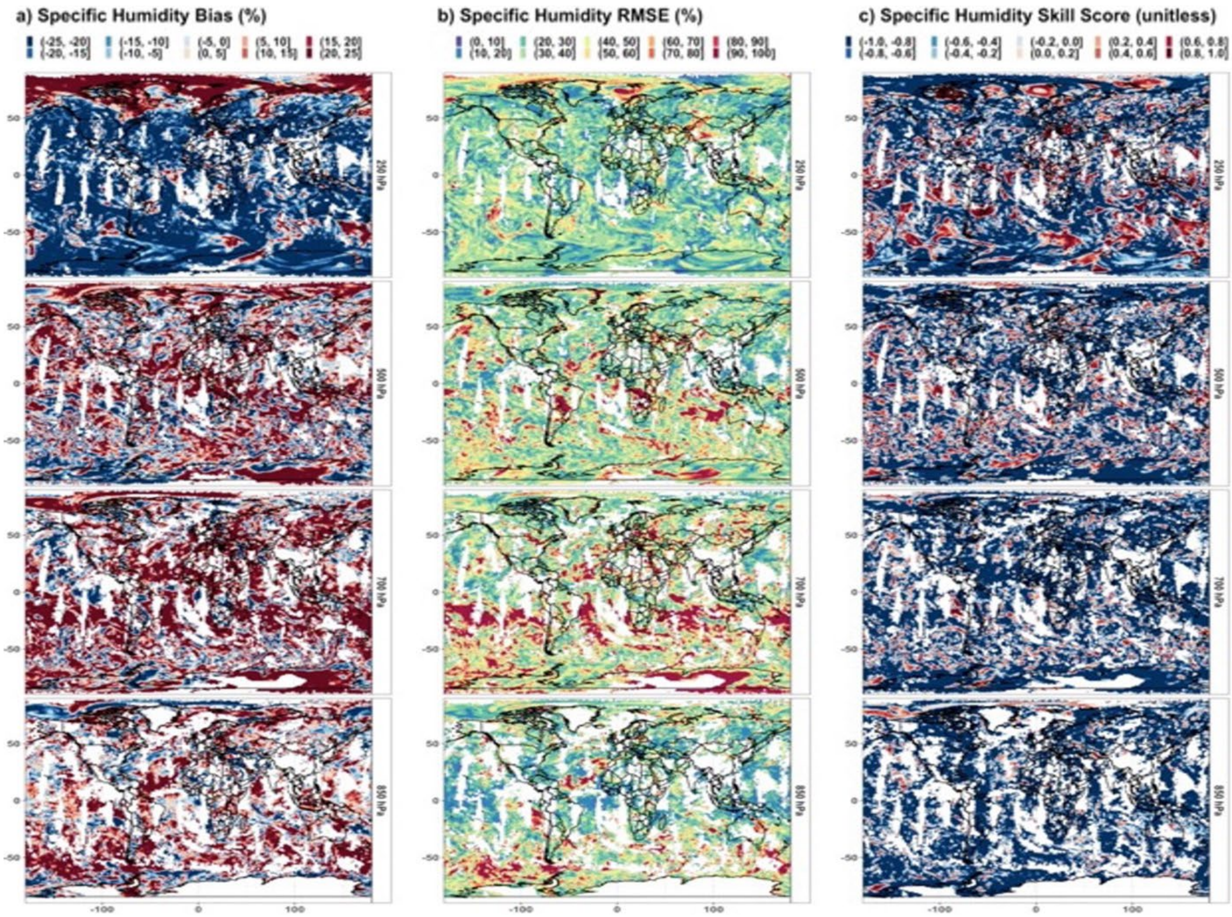

Figure 20. Same as Fig. 19, but for July of 2011.

3.4 Validation against the IGRA radiosondes

3.4.1 T and q Profile Patterns and Variability

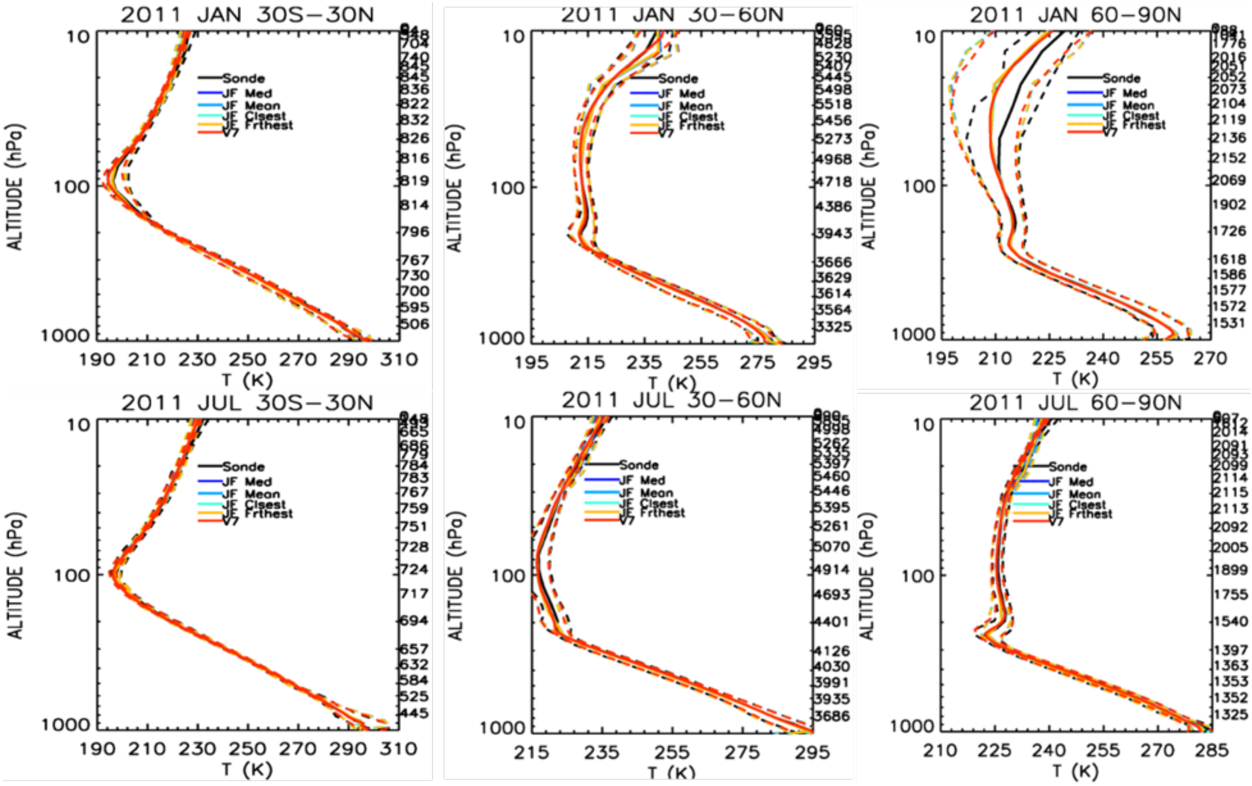

Figure 21. Medians of collocated temperature profiles (in K) averaged over the tropics (left, 30°S-30°N), Northern Hemispheric mid-latitudes (middle, 30°-60°N), and Northern Polar region (right, 60°-90°N) for January 11-15 (top) and July 11-15 (bottom) in 2011. The medians of IGRA radiosonde measurements are in black (Sonde), of medians of JoSFRA 3×3 footprint retrievals in dark blue (JF Med), of mean of JoSFRA 3×3 footprint retrievals in light blue (JF Mean), of the closet JoSFRA retrievals among the 3×3 footprints to the radiosondes in turquoise (JF Clsest), of the farthest JoSFRA retrievals among the 3×3 footprints to the radiosondes in yellow (JF, Frthest), and of AIRS V7 retrievals on 3×3 footprints in red. The dashed lines of the corresponding colors are the 25 and 75 percentiles. The numbers on the right axes are the collocated sample sizes.

Fig. 21 shows that JoSFRA and AIRS-V7 both well capture the patterns of vertical temperature profiles of radiosonde measurements. The northern polar middle stratosphere is subjected to large temperature variability because of planetary wave activity as indicated by both the retrievals and the radiosonde temperature profiles, with the retrievals having colder extremes of polar stratospheric temperatures. Both JoSFRA and AIRS-V7 have similar variability. There is no clear distinction between JoSFRA single-footprint and 3×3 footprint mean/median temperature profiles (or variability).

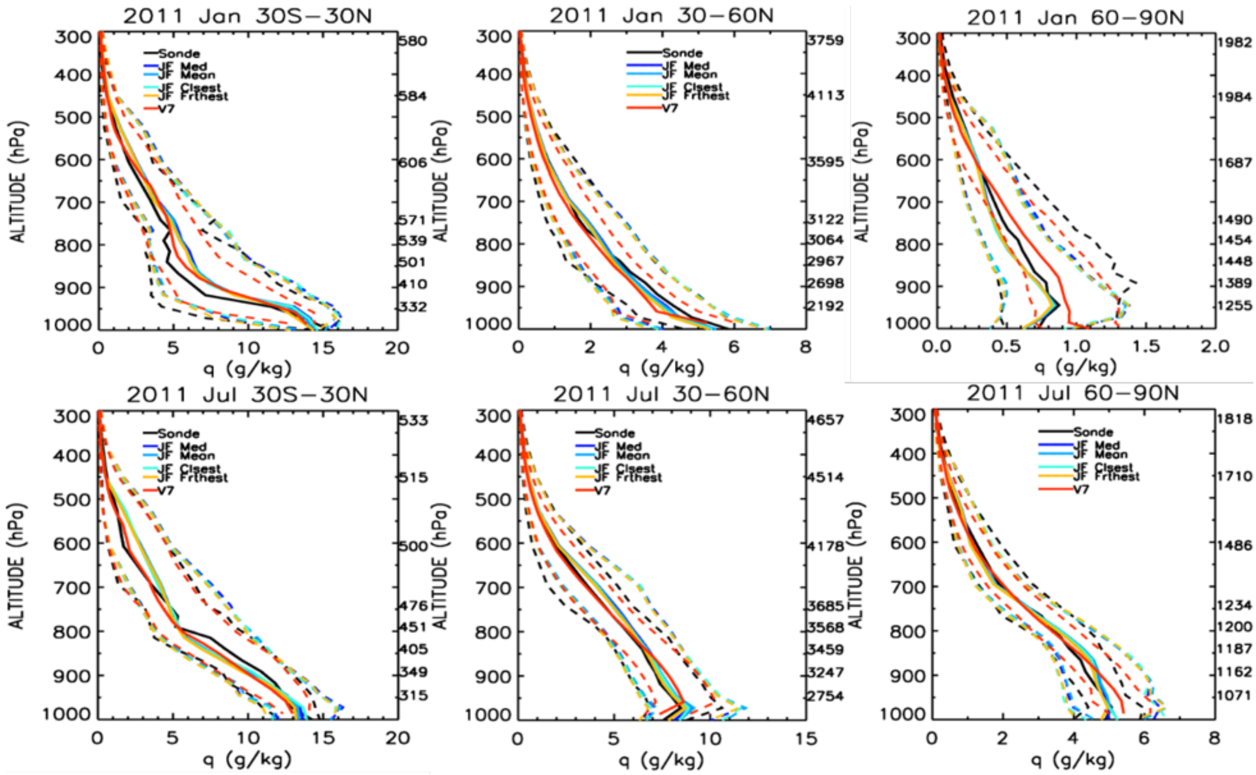

Figure 22. Similar to **Fig. 21**, but for specific humidity profiles (in g/kg).

Fig. 22 indicates that both JoSFRA and AIRS-V7 well capture the patterns of median vertical specific humidity profiles, with JoSFRA giving better near-surface moisture inversion during polar winter than AIRS-V7. Both JoSFRA and AIRS-V7 have reasonable variability, with AIRS-V7 tending to have narrower variations. There is no distinction between JoSFRA single footprint and 3×3 footprint mean/median specific humidity profiles (or variability).

3.4.2 Biases of T and q Retrievals

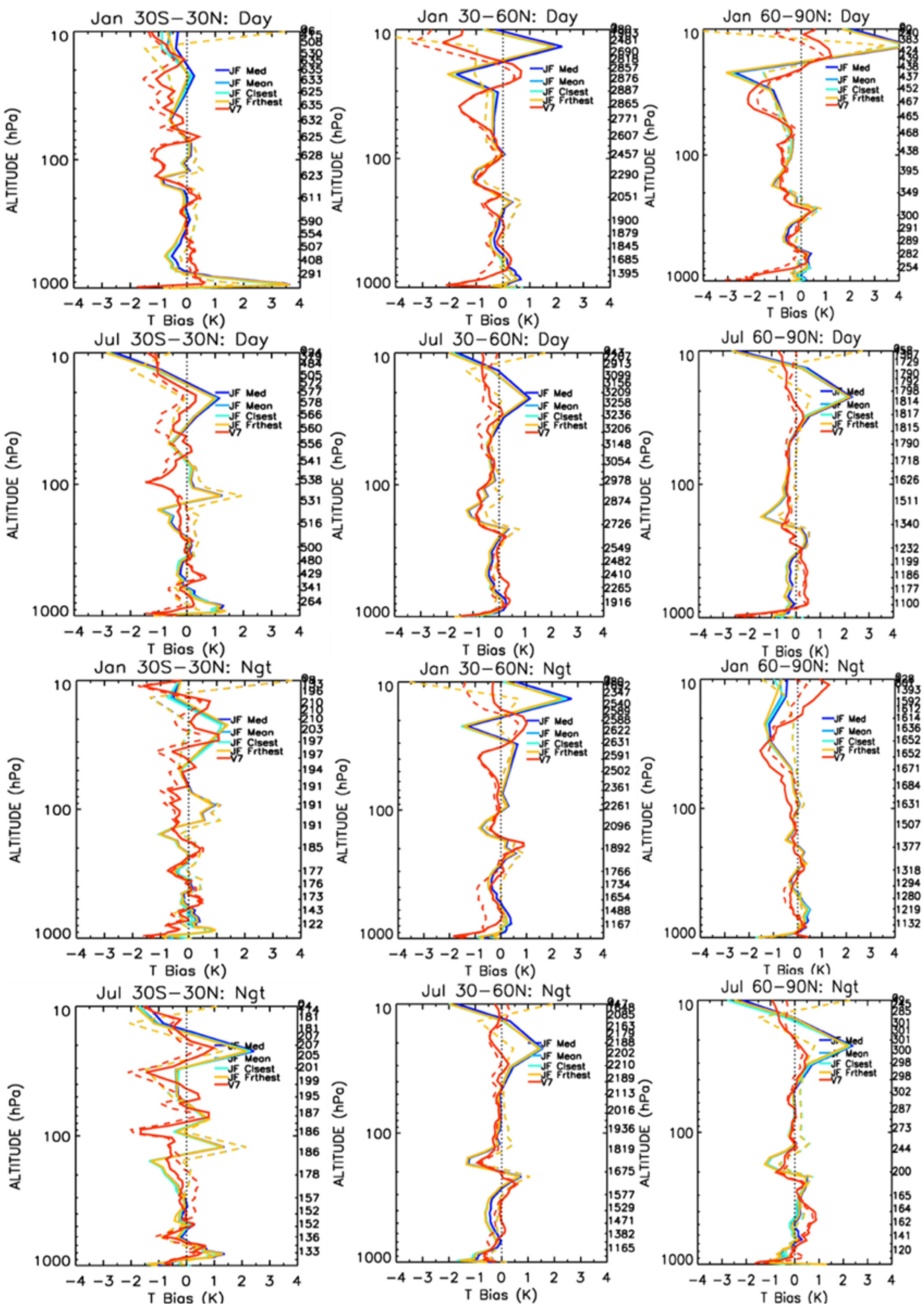

Figure 23. Temperature biases of JoSFRA retrievals (in K) from the collocated radiosonde measurements over the tropics (left, 30°S-30°N), Northern Hemispheric mid-latitudes (middle, 30°-60°N), and Nothern Polar region (right, 60°-90°N) for January 11-15 (1st and 3rd rows) and July 11-15 (2nd and 4th rows) in 2011. The upper two rows are for daytime retrievals and the bottom two rows are for nighttime retrievals. Color labels for different retrievals are the same as **Fig. 22** and the sample sizes are shown as the numbers on the right axes. The dashed lines are the corresponding first guesses of the retrievals.

In 50-700 hPa, JoSFRA temperature biases are within ±1 K (**Fig. 23**), with occasional situations reaching ±1.5 K near the tropopause. Near the surface, JoSFRA does not have the systematic cold biases seen in AIRS-V7. However, JoSFRA has a large warm bias during daytime January cases in the tropics. The distinction between the JoSFRA single-footprint and 3×3 mean/median biases is not obvious.

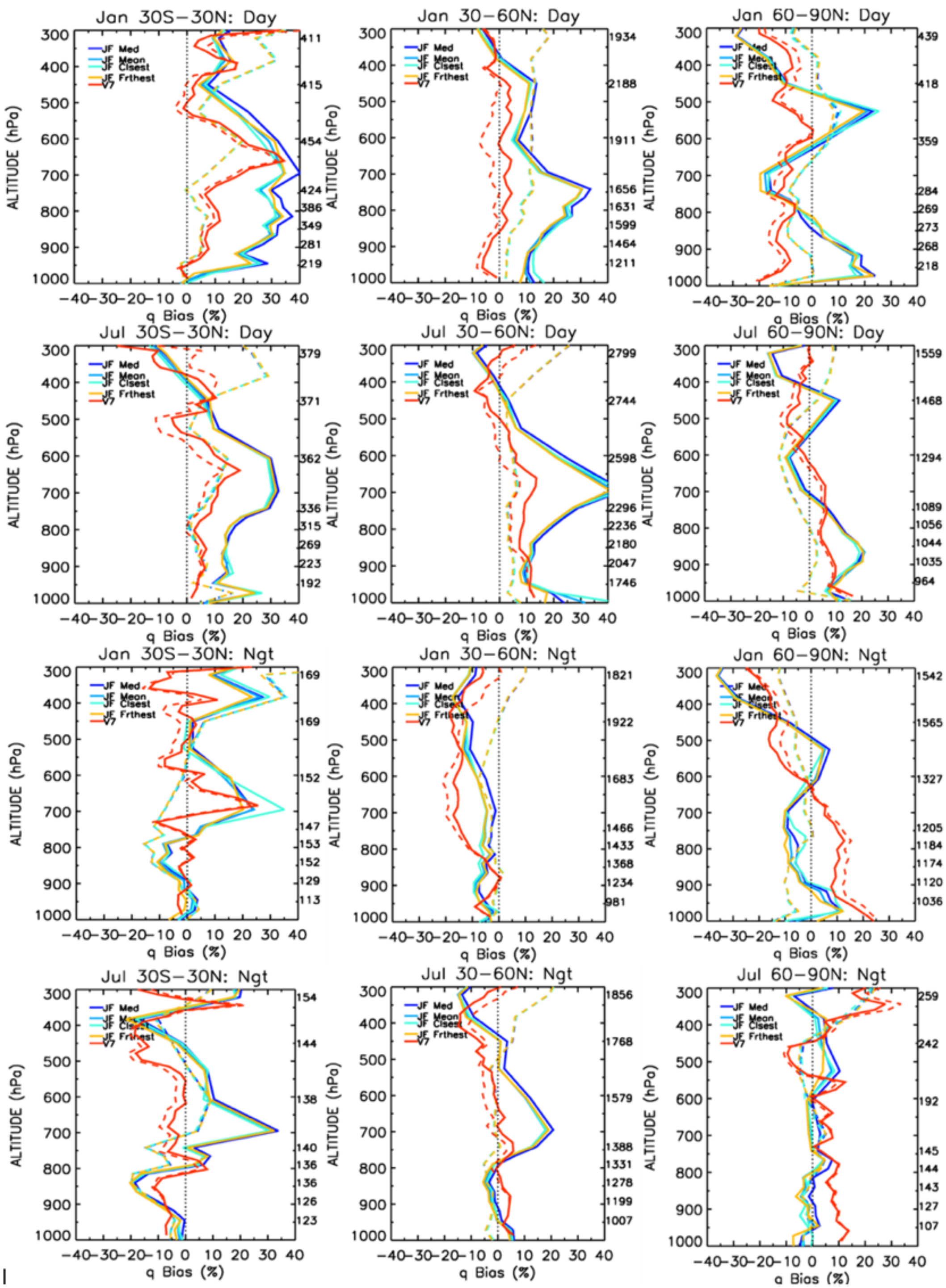

Figure 24. Similar to **Fig. 23**, but for relative biases of JoSFRA retrieved specific humidity.

JoSFRA has larger relative biases of specific humidity (20-40%) than AIRS-V7 (within ±20%, with one exceptional case reaching 25-30% at 650 hPa during January in the tropics.), and in many cases JoSFRA enhances the biases of its first guesses (**Fig. 24**). JoSFRA tends to have a much moister tropical and mid-latitude troposphere in 500-800 hPa.

3.4.3 Cloud-Dependent RMSEs and Single-Footprint Skills of T and q Retrievals

Stratifying error information according to cloud conditions reduces sample sizes of the matched-up profiles. Therefore, we perform cloud-dependent validation for middle latitudes (30-60N) only because of its original larger sample size. **Fig. 25** shows that JoSFRA temperature RMSEs are comparable to AIRS-V7 temperature RMSEs when the ECFs are small (0-0.1). In this nearly clear condition, JoSFRA has temperature RMSEs ~1 K in the middle troposphere, increasing to 1.5-2 K in the upper troposphere and boundary layer. Increasing ECFs enhances JoSFRA's RMSEs below 300 hPa to 2 K or more. JoSFRA tends to enhance its RMSEs from its first guesses while AIRS-V7 RMSEs are close to its first guesses, and JoSFRA performance near the surface during cloudy conditions is in general better than AIRS-V7 because of the better first guesses used by JoSFRA. The patterns of RMSEs described are independent of season (January vs. July) or day and night.

Temperature has, in general, less spatial variability than specific humidity; therefore, it is not an ideal variable to test the skill of single-footprint retrieval. However, one can still see in **Fig. 2**5 that in some cases the JoSFRA footprints closest to the radiosondes exhibit smaller RMSEs than the footprints farthest to the radiosondes (the turquoise lines have smaller RMSEs than the yellow lines), particularly in the troposphere when there are moderately thick clouds (ECFs around 0.1-0.9). The means/medians of 3×3 footprint retrievals usually have smaller RMSEs than the single-footprint ones, indicating that random noises do significantly contribute uncertainty to single-footprint retrievals.

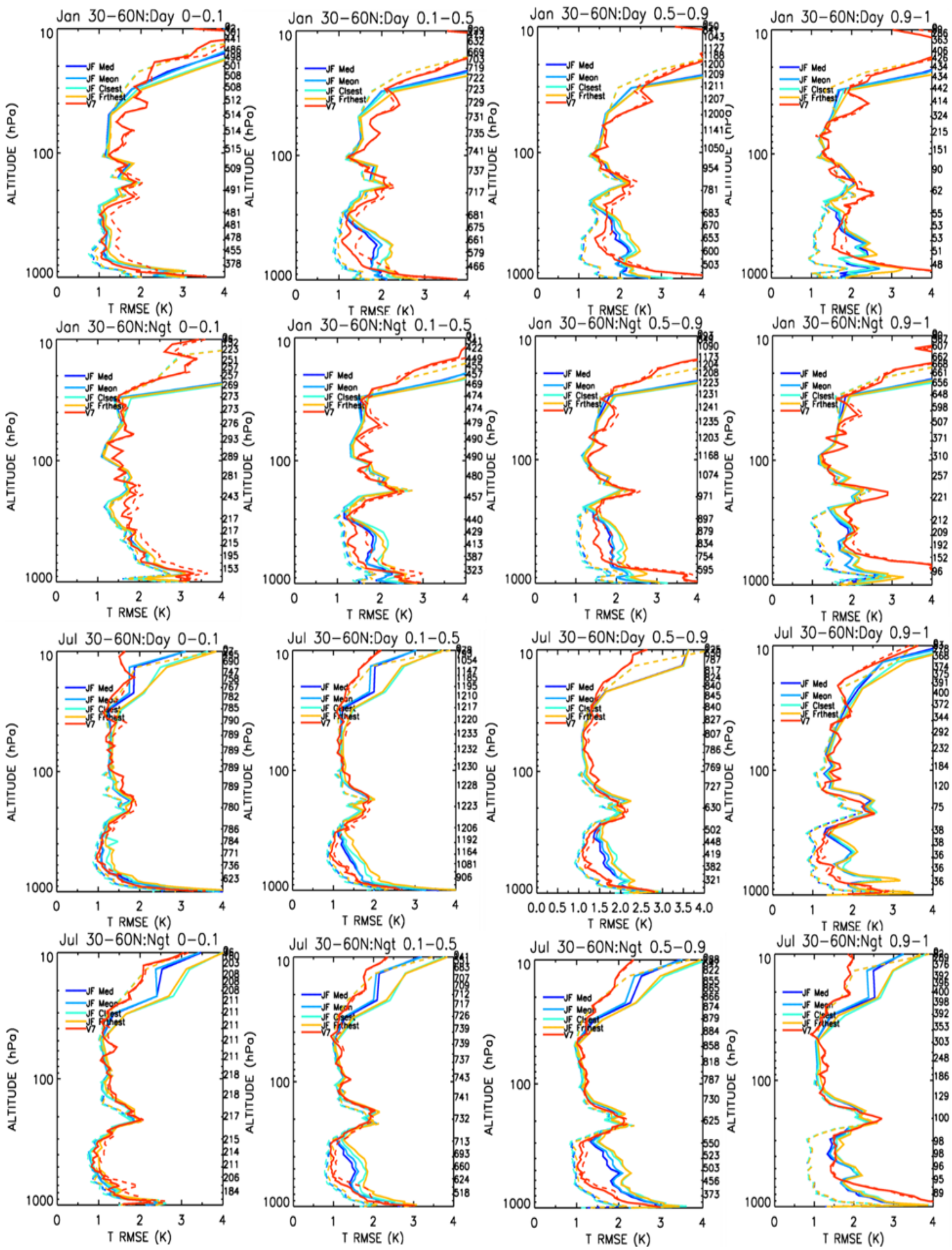

Figure 25. Temperature RMSEs of JoSFRA retrievals (in K) from the collocated radiosonde measurements as functions of collocated AIRS V7 effective cloud fractions (ECFs, ranges shown in the titles of each column) over the Northern Hemispheric mid-latitudes (middle, 30°-60°N) for January 11-15 (the upper two rows) and July 11-15 (the bottom two rows) in 2011. The 1st column are averages of profiles with ECFs of 0-0.1, the 2nd column with ECFs of 0.1-0.5, the third column with ECFs of 0.5-0.9, and the fourth column with ECFs of 0.9-1. The 1st and 3rd rows are

for daytime retrievals and the 2nd and 4th rows are for nighttime retrievals. Color labels for different retrievals are the same as **Fig. 22** and the sample sizes are shown as the numbers on the right axes. The dashed lines are the RMSEs of the corresponding first guesses of the retrievals.

Fig. 26 shows that JoSFRA has larger RMSEs (50% to >100%) than AIRS-V7 (20-80%) for specific humidity retrievals in the troposphere. JoSFRA enhances its first guesses' RMSEs and has occasional sudden spikes in RMSEs in the vertical profiles. Such enhancements and discontinuous jumps in RMSEs of JoSFRA increase with ECFs, and their occurrence seems to be random; however, once it occurs, the jump seems to exist across all ECFs. The origin of such discontinuous jumps is not clear.

Specific humidity has larger spatial variability and provides a good test for single-footprint retrieval skills. In "clear-sky" conditions (ECFs in the range 0-0.1), JoSFRA has consistent single-footprint skill above 700 hPa (**Fig. 25**, the turquoise lines have smaller RMSEs than the yellow lines). Such consistent single-footprint skills decrease in higher ECF ranges and are completely eliminated when the discontinuous jumps in RMSEs occur.

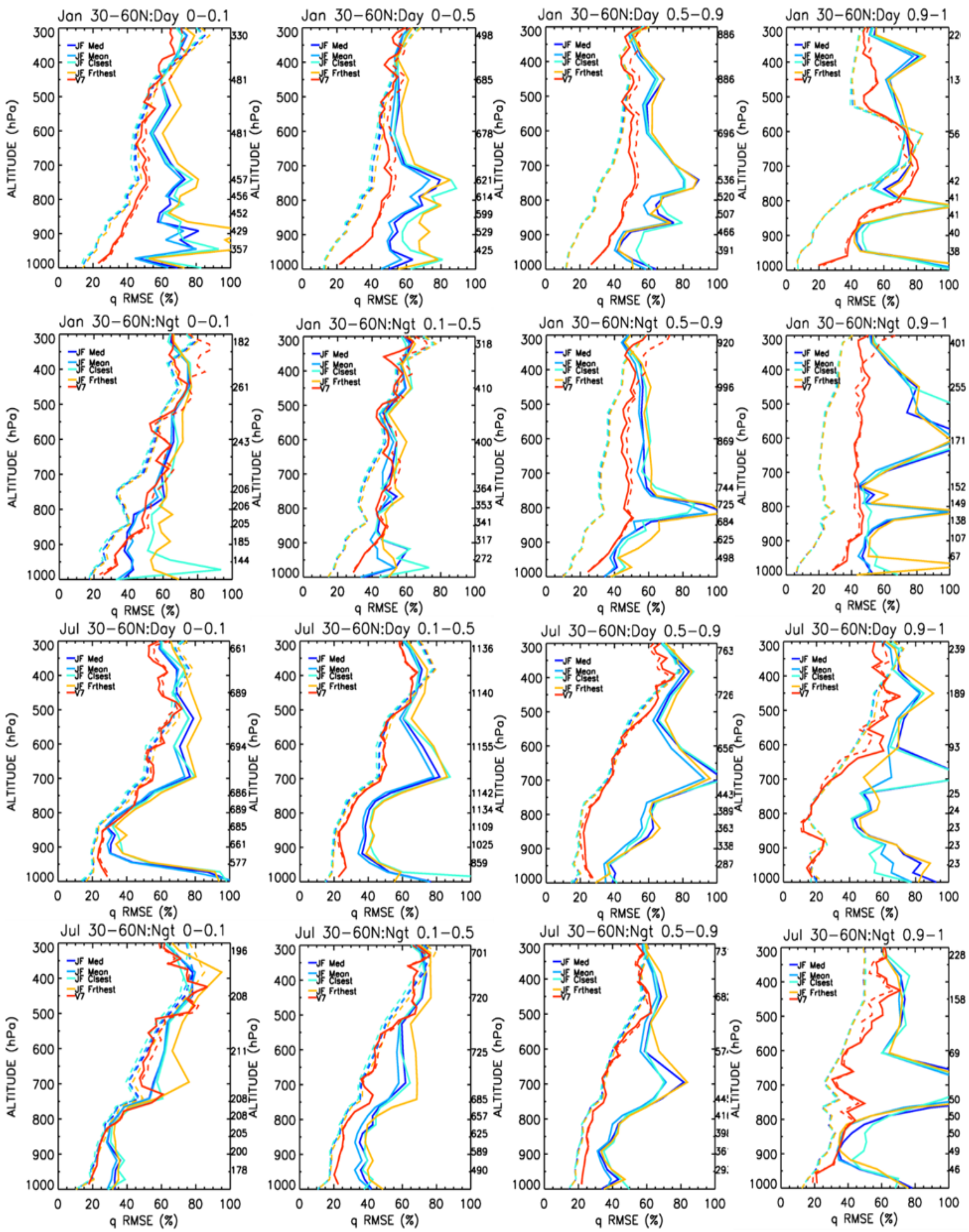

Figure 26. Similar to **Fig. 25**, but for relative RMSEs of specific humidity.

3.5 Testing for Composition Retrievals (O_3 and CO_2)

The current version of JoSFRA only has two composition products: ozone (O_3) profile and carbon dioxide (CO_2) profiles. Because composition products are not the focus of current version, here we only do a simple sanity check on the two species. For JoSFRA O_3 profile, we compare them against the AIRS v7 O_3 profiles. The AIRS O_3 profile has been carefully validated against ozonesonde [*Wang et al.*, 2022], in which we see that AIRS v7 O_3 has a 10-40% high bias in the troposphere, and a ±10% bias in the stratosphere.

Fig. 27 shows the degree of freedom (DoF) for O_3 and CO_2 in JoSFRA retrievals. We see that the DoF (information contents) for O_3 and CO_2 are less than 1, which is much smaller compared to DoF for temperature and H_2O. For both species, more information content is shown in the tropics to lower mid-latitudes, and less is in higher mid-latitude to polar regions. In the tropics, a few spots with low information content seem to correlate with the cloud fraction inferred from CALIOP data, indicating that the existence of clouds greatly degraded the retrievability of O_3 and CO_2.

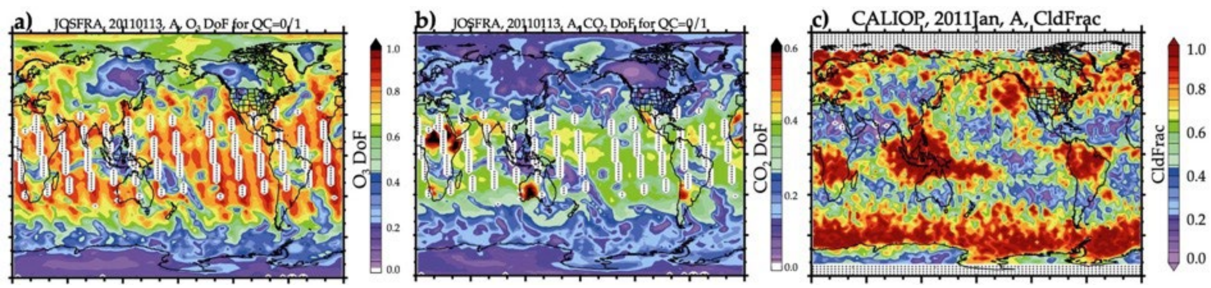

Figure 27. JoSFRA degree of freedom (DoF) for O_3 (left) and CO_2 (middle) in January 13, 2011, compared to the January mean cloud fraction inferred from CALIOP (right). Only cases with QC=0 or QC=1 were shown.

As a sanity check, we compare JoSFRA O_3 profiles to AIRS v7, as shown in **Fig. 28** below. In general, JoSFRA O_3 retrieval seems reasonable in that it produces the highest O_3 concentration near the 10-hPa level. The enhanced O_3 production due to photolysis at 10 hPa (~ 30 km) shifts from south during January towards north during July, following the seasonal variations of photolysis rates. Compared to AIRS, JoSFRA shows a 30-50% high bias near the tropopause levels, whereas AIRS O_3 is about 10% high compared to the ozonesonde [*Wang et al.*, 2022]. Therefore, JoSFRA O_3 could be 40-60% higher here if compared to the ozonesonde. In the stratosphere winter hemisphere, we also see JoSFRA O_3 being ~ 10% higher (60–90ºN in January and 60–90ºS in July). In the troposphere, JoSFRA O_3 is 20-40% lower than AIRS, with the fact that AIRS tropospheric O_3 is 10-35% higher compared to ozonesonde [*Wang et al.*, 2022], this means that JoSFRA O_3 is probably better than AIRS, and the bias could be small if compared to ozonesonde. This is further verified when compared to JoSFRA's first guess from NCEP CFSR reanalysis (right column), in which we see that the differences between retrieval and first guess are minimal, indicating that most of the tropospheric O_3 retrievals were relaxed to their first guesss, which was based on observationally based climatology analogous to that developed for Version 8 TOMS and SBUV [*McPeters et al.*, 2007].

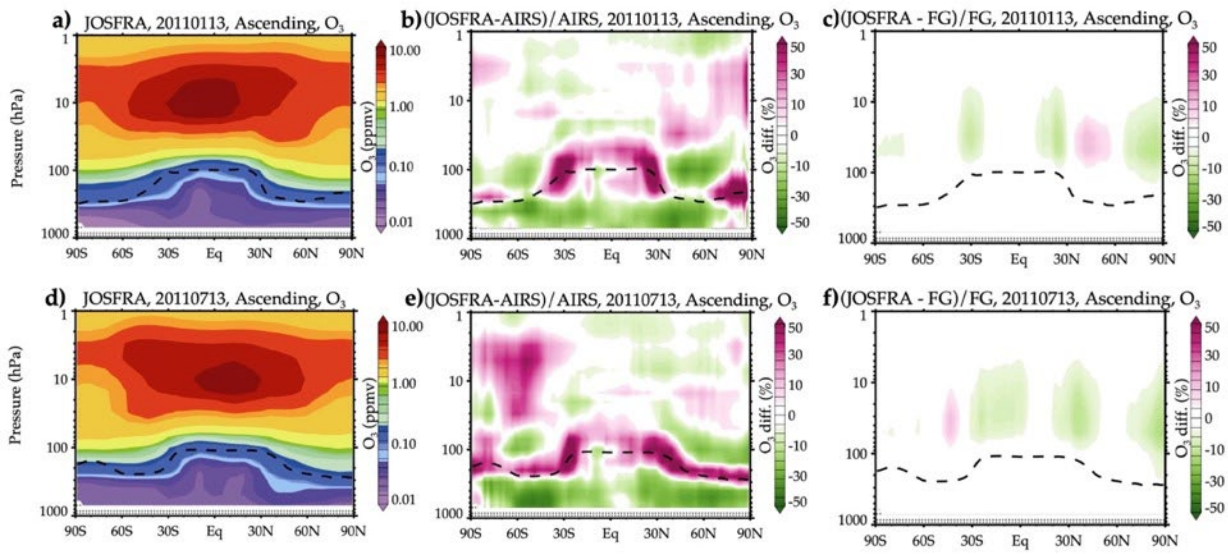

Figure 28. JoSFRA daily gridded zonal mean O3 profiles (left column) compared to O3 profiles from AIRS (middle column) and from JoSFRA's first guess (right column), for January 13 (upper row) and July 13 (lower role) of 2011. Black dashed lines are the lapse-rate tropopause. Only cases with QC=0 or QC=1 were shown.

The fact that JoSFRA's tropospheric O_3 retrieval is relaxed to its first guess can also be seen in the skill score plots, as shown in **Fig. 29** below. Here, we see that the skill score is almost uniformly zero in the troposphere and around the tropopause, suggesting that $MSE_{retrieval} = MSE_{firstGuess}$, so the retrieval is very close to the first guess, as suggested by the right column in **Fig. 28**. In the stratosphere, the score can be either positive (retrieval has smaller MSE) or negative (first guess has smaller MSE).

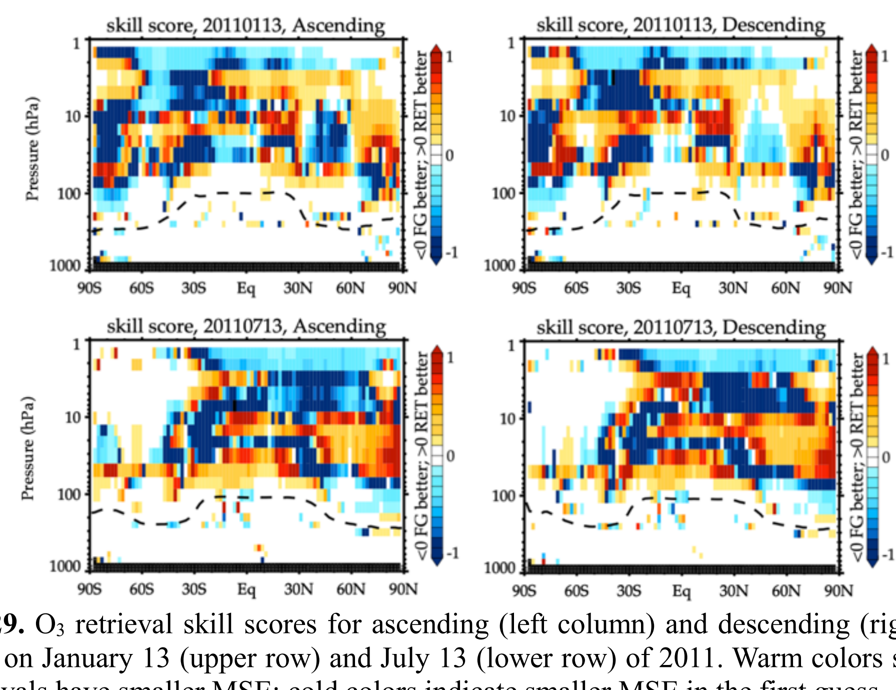

Figure 29. O_3 retrieval skill scores for ascending (left column) and descending (right column) granules on January 13 (upper row) and July 13 (lower row) of 2011. Warm colors suggest that the retrievals have smaller MSE; cold colors indicate smaller MSE in the first guess.

In evaluating JoSFRA CO_2 retrieval, we simply compare it to the CarbonTracker-CO2 product by the NOAA Earth System Research Laboratories (ESRL). The CarbonTracker-CO_2 is a CO_2 measurement and modeling system developed by NOAA to keep track of sources and sinks of CO_2 around the world (https://www.esrl.noaa.gov/gmd/ccgg/carbontracker/). To be specific, we use the latest version (CT2022) of 3-hourly, global 3D CO_2 mole fraction fields available in NetCDF format (https://gml.noaa.gov/aftp//products/carbontracker/co2/CT2022/molefractions/co2_total/), which is an update from the previous formal version 2019B [*Jacobson et al.*, 2020].

Fig. 30 shows an example of the JoSFRA CO_2 (panel a) compared to its first guess (panels b-c) and to the CarbonTracker CO_2 (panels d-e) for January 13, 2011. The comparisons look similar for the July data, so we did not show it here. **Fig. 30** suggests that JoSFRA stratospheric CO_2 retrieval is largely relaxed to its first guess. In January, JoSFRA reproduces the northern hemispheric CO_2 emission hotspot near the surface, but this is attributable to the first guess given the similarity of the two (panels a and b). JoSFRA shows a quite high CO_2 background in the southern hemisphere, which is also caused by the high CO_2 in the first guess.

Another thing worth mentioning is that JoSFRA shows elevated CO_2 near the tropical tropopause, which looks unreal given that there are no known chemical sources of CO_2 at that altitude. Overall, the current version of retrieval relies largely on its first guess, which has room to improve in the next version because currently, the first guess shows a smaller vertical gradient of tropical CO_2 overall, which indicates a weaker vertical upwelling compared to CarbonTracker-CO_2.

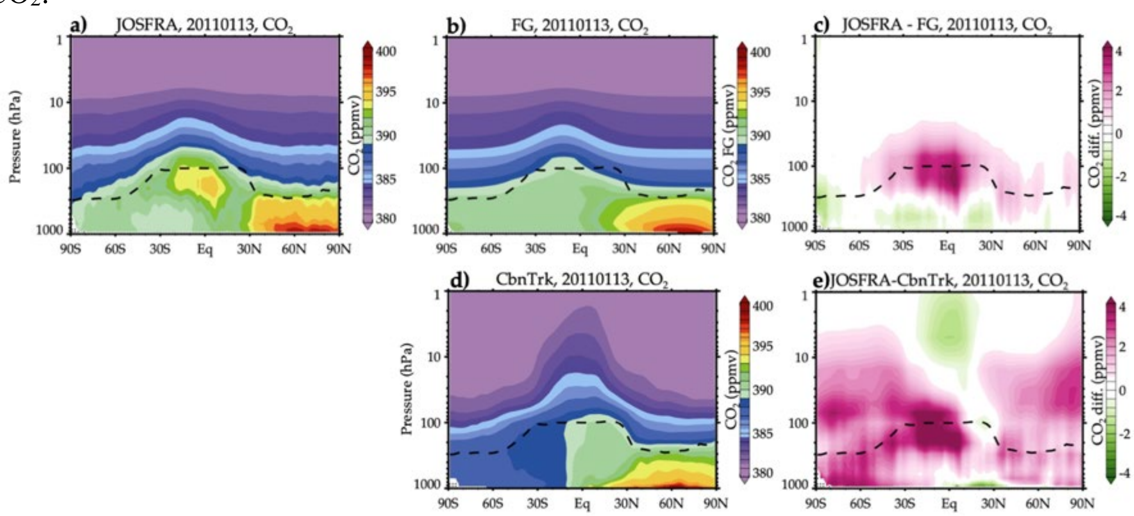

Figure 30. (a) JoSFRA daily gridded zonal mean CO_2 profiles compared to (b) JoSFRA's first guess profiles and (c) the differences; compared to (d) CarbonTracker-CO_2 and (e) the differences for January 13 of 2011. Black dashed lines are the lapse-rate tropopause. The July data looks similar, so we did not show it here.

4. Summary

a. Yields

- Overall, JoSFRA has 30-40% of retrievals flagged as of the best quality (QC=0).
- For temperature retrievals, about 70% of JoSFRA retrievals can be used for science study (QC=0 or QC=1) .
- For specific humidity retrievals, JoSFRA has about 45-50% of profiles that are suitable for scienfic study. About 30% of specific humidity retrievals are flagged as bad quality, indicating

that the retrieval is greatly challenged due to the existence of clouds, especially over the tropical Western Pacific.

- The retrieval yield for specific humidity is between 40-60% for most of the troposphere, and it degrades when near the surface.
- Temperature and specific yields are consistent in different seasons.

b. Temperature

When compared to ERA5:

- JoSFRA had a warm bias in the troposphere that was less than 0.5 K.
- Skill score < 0 except near the tropopause.
- Small day/night differences

When compared to IGRA radiosondes:

- In 50-700 hPa, JoSFRA temperature biases are within ±1 K, with occasional situations reaching ±1.5 K near the tropopause.
- JoSFRA has a large warm bias during daytime January cases in the tropics.
- When the ECFs are small (0-0.1), JoSFRA has temperature RMSEs ~1 K in the middle troposphere, increasing to 1.5-2 K in the upper troposphere and boundary layer.
- Increasing ECFs enhances JoSFRA's RMSEs below 300 hPa to 2 K or more.
- JoSFRA tends to enhance its RMSEs from its first guesses.
- The patterns of RMSEs described are independent of season (January vs. July) or day and night.
- JoSFRA shows some single-footprint skill in the troposphere when there are moderately thick clouds (ECFs around 0.1-0.9).
- The means/medians of 3×3 footprint retrievals usually have smaller RMSEs than the single-footprint ones, implying that single footprint retrievals introduce additional noises.

c. Specific Humidity

When compared to ERA5:

- Wet bias less than 15% in the troposphere with an RMSE less than 70%
- No seasonal differences in the yield
- JoSFRA had skill against the first guess, particularly in the upper troposphere, but in the lower troposphere JoSFRA had skill < 0.
- JoSFRA had some skill in the tropics around 400 hPa and between 200 and 100 hPa
- Day/night differences were small.

Those conclusions also apply to the comparisons with AIRS.

When compared to IGRA radiosondes:

- JoSFRA has relative biases of specific humidity (20-40%).
- JoSFRA enhances the biases of its first guesses.
- JoSFRA tends to have a much moister tropical and mid-latitude troposphere in 500-800 hPa than radiosonde measurements.

- JoSFRA has RMSEs of 50% to over 100% in the troposphere.
- JoSFRA enhances its first guesses' RMSEs.
- JoSFRA has random sudden spikes in RMSEs in the vertical profiles that sometimes can boost the RMSEs > 80%. The origin of such discontinuous jumps is not clear.
- When ECFs are in the range of 0-0.1, JoSFRA has consistent single-footprint skill above 700 hPa. Such consistent single-footprint skills decrease in higher ECF ranges and are completely eliminated when the discontinuous jumps in RMSEs occur.

d. O_3 and CO_2:

- JoSFRA is able to retrieve global O_3 profiles, despite the relatively small information content used. JoSFRA O_3 profile retrieval is slightly better than AIRS, which could be caused by the fact that JoSFRA retrieval is more relaxed to its first guess. There is at most a 2-4% deviation from its first guess in the stratosphere, whereas in the troposphere the difference is less than 1%. Therefore, the retrieval score is almost uniformly zero in the troposphere and around the tropopause, suggesting that the mean square error of retrieval is comparable to the mean square error of the first guess.
- JoSFRA retrieval of CO_2 is currently experimental. We see that JoSFRA captures the northern hemispheric high CO_2 emission near the mid-latitude surface. This is because the retrieval is largely relaxed to its first guess in this region given the similarity of the two. Compared to CarbonTracker-CO_2, JoSFRA broadly overestimates CO_2 abundances in the Southern Hemisphere, which makes the comparisons bearing a large positive bias. This is also manifested in the first guess, providing another clue that the retrieval is relaxed to its first guess. JoSFRA shows elevated CO_2 near the tropical tropopause, which looks unreal given that there are no known chemical sources of CO_2 at that latitude. Given the very limited information content shown in the DoF (less than 1), the challenge is expected.

References

Copernicus Climate Change Service (C3S) (2017): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS), *date of access*. https://cds.climate.copernicus.eu/cdsapp#!/home

Hennermann, Karl. (2020). *What are the changes from ERA-Interim to ERA5?* - Copernicus Knowledge Base - ECMWF Confluence Wiki. https://confluence.ecmwf.int/pages/viewpage.action?pageId=74764925

Jacobson, A. R., K. N. Schuldt, J. B. Miller, T. Oda, P. Tans, A. Andrews, J. Mund, L. Ott, G. J. Collatz, T. Aalto, S. Afshar, K. Aikin, S. Aoki, F. Apadula, B. Baier, P. Bergamaschi, A. Beyersdorf, S. C. Biraud, A. Bollenbacher, D. Bowling, G. Brailsford, J. B. Abshire, G. Chen, H. Chen, L. Chmura, ..., and M. Zimnoch (2020). CarbonTracker CT2019B. Model published by NOAA Global Monitoring Laboratory. URL https://gml.noaa.gov/ccgg/carbontracker/CT2019B/.

McPeters, R. D., Labow, G. J., and Logan, J. A. (2007), Ozone climatological profiles for satellite retrieval algorithms, J. Geophys. Res., 112, D05308, doi:10.1029/2005JD006823.

Murphy, A. H. (1988), Skill scores based on the mean square error and their relationship to the correlation coefficient, *Mon. Weather Rev.*, **116**, 2417–2424.

Irion, F. W., Blaisdell, J. M., Fetzer, E. J., Fishbein, E., etc. (2023), Algorithm Theoretical Basis Document for the Joint Single Footprint Retrieval Algorithm (JoSFRA).

Thrastarson, H. Th. (Ed.), Manning, E., Irion, F.W., Monarrez, R. (2023). JoSFRA Level-2 Product User Guide. Jet Propulsion Laboratory, California Institute of Technology.

Wang, T., Payne, V., Manning, E., Fishbein, E., Yue, Q., Wong., S., Osterman, G., etc. (2022): Testing report for the CLIMCAPS-Aqua v02.39 and AIRS v7 Level-2 Ozone profiles. November 2022.

WMO (2012): Manual on the Global Data-Processing and Forecasting System, Volume I - Global Aspects. Geneva, Switzerland: World Meteorological Organization. II. 7-36 - II. 7-38. Online at: http://www.wmo.int/pages/prog/www/DPFS/documents/485_Vol_I_en_colour.pdf .

Wong, S., E. J. Fetzer, M. Schreier, G. Manipon, E. F. Fishbein, B. H, Kahn, Q. Yue, and F. W. Irion (2015), Cloud-induced uncertainties in AIRS and ECMWF temperature and specific humidity, *J. Geophys. Res.*, 120, doi:10.1002/2014JD022440.

Acknowledgement

The research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004).