

Evaluating CMIP5 Models using AIRS Tropospheric Air Temperature and Specific Humidity Climatology

Baijun Tian, Eric J. Fetzer, Brian H. Kahn, Joao Teixeira, and Evan Manning

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA

Thomas Hearty

Goddard Space Flight Center/Wyle Information Systems, Greenbelt, MD

Journal of Geophysical Research - Atmospheres

Submitted, 07/31/2012

Revised, 11/14/2012

Further revised, 12/11/2012

Corresponding author address: Baijun Tian, Jet Propulsion Laboratory, M/S 233-304, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109.

E-mail: baijun.tian@jpl.nasa.gov

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/jgrd.50117

Abstract

This paper documents the climatological mean features of the Atmospheric Infrared Sounder (AIRS) monthly mean tropospheric air temperature (t_a , K) and specific humidity (h_{us} , kg/kg) products as part of the Obs4MIPs project and compares them to those from NASA's Modern Era Retrospective analysis for Research and Applications (MERRA) for validation and sixteen Coupled Model Intercomparison Project (CMIP5) model simulations for CMIP5 model evaluation. MERRA is warmer than AIRS in the free troposphere but colder in the boundary layer with differences typically less than 1 K. MERRA is also drier ($\sim 10\%$) than AIRS in the tropical boundary layer but wetter ($\sim 30\%$) in the tropical free troposphere and the extratropical troposphere. In particular, the large MERRA-AIRS specific humidity differences are mainly located in the deep convective cloudy regions indicating that the low sampling of AIRS in the cloudy regions may be the main reason for these differences. In comparison to AIRS and MERRA, the sixteen CMIP5 models can generally reproduce the climatological features of tropospheric air temperature and specific humidity well, but several noticeable biases exist. The models have a tropospheric cold bias (around 2 K), especially in the extratropical upper troposphere, and a double-ITCZ problem in the troposphere from 1000 hPa to 300 hPa, especially in the tropical Pacific. The upper-tropospheric cold bias exists in the most (13 of 16) models, and the double-ITCZ bias is found in all 16 CMIP5 models. Both biases are independent of the reference dataset used (AIRS or MERRA).

1. Introduction

The fifth phase of the Coupled Model Intercomparison Project (CMIP5) involving around twenty climate modeling groups from around the world has produced a freely available state-of-the-art multi-model dataset designed to advance our knowledge of climate variability and climate change [Taylor *et al.*, 2012]. The next Intergovernmental Panel on Climate Change (IPCC) Assessment Report (AR5), scheduled for publication in 2014, will rely mainly on the peer-reviewed analyses of model outputs from the CMIP5 experiments. To increase the fidelity of the IPCC AR5, the CMIP5 model experiments need vigorous evaluation by comparing the model outputs to state-of-the-art observations and by quantifying the model errors that may lead to climate projection uncertainties [Randall *et al.*, 2007]. Thus, a wide variety of observationally based datasets are needed for CMIP5 model evaluation.

The National Aeronautics and Space Administration (NASA) Jet Propulsion Laboratory (JPL) and the Department of Energy (DOE) Program for Climate Model Diagnosis and Inter-comparison (PCMDI) have initiated the Obs4MIPs project for CMIP5 model evaluation [Teixeira *et al.*, 2011]. Obs4MIPs refers to a limited collection of well-established and well-documented datasets that have been organized according to the CMIP5 model output requirements and made available on the PCMDI Earth System Grid (ESG) - Center for Enabling Technologies (ESG-CET) (<http://pcmdi6.llnl.gov/esgcet/>) for CMIP5 model evaluation. Each Obs4MIPs dataset corresponds to a field that is output in one or more of the CMIP5 experiments. This technical alignment of observational products with climate model output can greatly

facilitate model-observation comparisons. This effort has been initiated with support from NASA and DOE with the intention of enabling additional data providers to contribute products. Please see the Obs4MIPs Wiki (<http://obs4mips.llnl.gov:8080/wiki/>) for more details. The NASA Atmospheric Infrared Sounder (AIRS) project [Aumann *et al.*, 2003; Chahine *et al.*, 2006] is a pioneer in the Obs4MIPs project. The AIRS project produced monthly averaged tropospheric air temperature (ta) and specific humidity (hus) products as part of the Obs4MIPs project.

Tropospheric water vapor is widely recognized to be a key climate variable. It is the dominant greenhouse gas and provides a key feedback for amplifying the climate sensitivity to external forcing [Bony *et al.*, 2006; Held and Soden, 2000; Soden and Held, 2006]. There has been a considerable effort to assess the credibility of model simulations of atmospheric water vapor using satellite measurements [e.g., Allan *et al.*, 2003; Bates and Jackson, 1997; Brogniez *et al.*, 2005; Soden and Bretherton, 1994]. Many of these studies have identified systematic biases in the climatological distribution of water vapor simulated by various models.

There are several advanced temperature and moisture sounding systems in space now, such as the AIRS/Advanced Microwave Sounding Unit (AMSU) on NASA Aqua satellite [Aumann *et al.*, 2003; Chahine *et al.*, 2006], the Infrared Atmospheric Sounding Interferometer (IASI) [Hilton *et al.*, 2012] and AMSU-A/Microwave Humidity Sounder (MHS) on the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Meteorological Operation (MetOp) satellites, as well as the Cross-track Infrared Sounder (CrIS)/Advanced Technology Microwave Sounder (ATMS) on Suomi National Polar-orbiting Partnership (NPP) satellite. Through multispectral coverage in

infrared and microwave channels, the AIRS/AMSU system obtains vertical profiles of atmospheric temperature and water vapor with vertical resolution of around 2 km, horizontal resolution of 45 km at nadir, twice daily temporal resolution, global coverage, and for cloud cover up to about 70%. Thus, AIRS is a valuable dataset to evaluate climate models in terms of the vertical structure of tropospheric temperature and water vapor.

The AIRS tropospheric temperature and water vapor data have been used to evaluate climate models. *Gettelman et al.* [2006] compared the AIRS relative humidity (RH) data to that simulated by the National Center for Atmospheric Research (NCAR) Community Atmosphere Model version 3 (CAM3). They found that the model does a good job of reproducing the mean RH distribution but is slightly moister than the observations in the middle and upper troposphere. Further, the CAM3 model has difficulties reproducing many scales of observed variability, particularly in the tropics. *Pierce et al.* [2006] and *John and Soden* [2007] compared the AIRS specific humidity measurements with those from coupled general circulation models (GCMs) of the CMIP third phase (CMIP3) archive. They found that most CMIP3 models have a large moist bias in the free troposphere (more than 100%) especially over the extratropics, but a dry bias in the boundary layer (up to 25%) over the tropics. *John and Soden* [2007] also found that the CMIP3 model simulated temperatures are systematically colder by 1–2 K throughout the troposphere. This cold bias generally increases with altitude in the free troposphere, with maxima located near 200 hPa in the extra-tropics. However, the above studies have mainly evaluated the CMIP3 models instead of CMIP5 models. Recently, *Jiang et al.* [2012] evaluated the cloud and water vapor simulations in CMIP5 models using the “A-Train” satellite observations including the AIRS specific humidity data.

They mainly studied CMIP5 model improvement from CMIP3 models as well as CMIP5 model performance relative to observations, including global, tropical, mid-latitude, and high-latitude profiles as well as global spatial correlations and variances expressed as Taylor diagrams. However, they did not discuss the detailed spatial maps of model specific humidity biases (i.e., model-AIRS) at various vertical levels that are the main topic of this paper. Furthermore, they did not discuss the CMIP5 model simulations of tropospheric air temperature at all.

The first main purpose of this paper is to document the characteristics of these two AIRS datasets in terms of data origin, data description, caveats for model-observation comparison. The second main purpose of this paper is to compare the tropospheric air temperature and specific humidity climatology between AIRS and NASA's Modern Era Retrospective analysis for Research and Applications (MERRA) for data validation and between AIRS and CMIP5 model simulations for CMIP5 model evaluation focusing on the spatial maps of model biases (i.e., model-AIRS). The rest of this paper is organized as follows. Section 2 describes these two AIRS datasets including its data origin, data description, considerations for model-observation comparison, and data validation. Section 3 describes the CMIP5 model output and model-observation comparison methodology. The main results of this paper, i.e., the comparison of AIRS, MERRA and CMIP5 tropospheric air temperature and specific humidity climatology, are presented in Section 4 followed by a summary and conclusions in Section 5.

2. AIRS Data

2.1. Data description

As part of the Obs4MIPs project, the AIRS project has generated two new AIRS

datasets for CMIP5 model evaluation including the monthly mean tropospheric t_a (K) and specific humidity (kg/kg) profile products from the standard AIRS Level-3 (L3) version 5 (V5) monthly mean tropospheric air temperature and specific humidity products [Susskind *et al.*, 2006]. These two AIRS datasets are provided for each calendar month from September 2002 to May 2011, on a global spatial grid at 1° -longitude by 1° -latitude resolution (the same spatial grid from the AIRS V5 L3), and on the 17 CMIP5 mandatory vertical pressure levels (1000, 925, 850, 700, 600, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, and 10 hPa). However, valid AIRS data are available only for the lowest 8 pressure levels from 1000 to 300 hPa (1000, 925, 850, 700, 600, 500, 400, and 300 hPa). A missing value (1.e20) is assigned for the 9 pressure levels above 300 hPa (250, 200, 150, 100, 70, 50, 30, 20, and 10 hPa) because AIRS water vapor measurements are not as reliable for pressure levels above 300 hPa as other instruments such as the Aura Microwave Limb Sounder (MLS) [Waters *et al.*, 2006], which is specially designed for the accurate measurements of the atmospheric air temperature and specific humidity profiles in the upper troposphere and lower stratosphere [Fetzer *et al.*, 2008; Read *et al.*, 2007]. The key strengths of these AIRS datasets include global coverage with high horizontal and vertical resolutions and frequent sampling, validated against radiosondes, radio occultation data and reanalyses, and co-located temperature and humidity measurements. Their key weaknesses are low sampling in cloudy regions, inconsistent sampling with latitude and incomplete diurnal cycle coverage.

2.2. Data origin

Launched on May 4, 2002, Aqua is part of NASA's "A-Train" satellite constellation, a series of high-inclination, Sun-synchronous satellites in low Earth orbit

designed to make long-term global observations of the land surface, biosphere, solid Earth, atmosphere, and oceans [Stephens *et al.*, 2002]. The AIRS and its partner microwave instruments AMSU and Humidity Sounder for Brazil (HSB) [Aumann *et al.*, 2003; Chahine *et al.*, 2006] share the Aqua satellite with the Moderate Resolution Imaging Spectroradiometer (MODIS), Clouds and the Earth's Radiant Energy System (CERES), and the Advanced Microwave Scanning Radiometer-EOS (AMSR-E) [Parkinson, 2003]. The Aqua spacecraft orbits the Earth every 98.8 minutes with an equatorial crossing time going north (ascending) at 1:30 P.M. local time (daytime) and going south (descending) at 1:30 A.M. local time (nighttime) in a Sun-synchronous, near polar orbit with an inclination of 98.2° and 705 km of operational altitude. HSB improves the water vapor retrieval capability of the AIRS system, but ceased operating in February 2003. The AIRS and AMSU instruments are each cross-track scanning nadir sounders that are co-aligned and have a swath roughly 1650 km wide. The AIRS instrument is a 2378-channel grating spectrometer measuring infrared radiance at wavelengths in the range 3.7–15.4 μm with a horizontal resolution of about 13.5 km at nadir [Aumann *et al.*, 2003]. These wavelengths are sensitive to temperature and moisture profiles, clouds, minor gases, and surface properties. The AMSU instrument is a fifteen-channel microwave radiometer with a horizontal resolution of about 45 km at nadir [Lambrigtsen, 2003]. Twelve AMSU channels are sensitive to temperature, with the other three channels are used for moisture measurement and precipitation detection.

The AIRS/AMSU geophysical retrieval method (no HSB data) uses an iterative, least-square physical inversion of cloud-cleared infrared radiances, obtained from a combination of nine AIRS fields of view contained within each AMSU field of view. The

algorithm is referred to as the AIRS/AMSU combined retrieval and was described by *Susskind et al.* [2003; 2006]. Following common practice, any discussion of AIRS in this work implicitly refers to the AIRS/AMSU system. The AIRS/AMSU sounding system produces about 324,000 air temperature and specific humidity profiles every day, though with varying information content. The horizontal resolution of air temperature and specific humidity profiles is 45 km at nadir and can be approximately 100 km at the high scan angles off-nadir, the same as the AMSU field of view. *Maddy and Barnett* [2008] found that, depending on the scene, AIRS Version 5 tropospheric air temperature (specific humidity) retrieval vertical resolution, which is as determined by the full-width at half-maximum of the averaging kernels, ranges between 2.5 km (2.7 km) near the surface and 7.1 km (4.3 km) near the tropopause. Those AIRS air temperature and specific humidity profiles are referred to as Level-2 (L2) products [*Susskind et al.*, 2006]. The AIRS L3 air temperature and specific humidity profile products are the gridded averages of the AIRS L2 swath air temperature and specific humidity profiles on horizontal 1° -latitude \times 1° -longitude grids, the 24 World Meteorological Organization (WMO) standard pressure levels from 1000 to 1 hPa for air temperature and 12 WMO standard layers from 1000 to 100 hPa for specific humidity. They are reported for the ascending (daytime) and descending (nighttime) orbits separately and three temporal resolutions: daily, 8-day (one half of the Aqua orbit repeat cycle) and monthly (calendar month). The multi-day products are simply the arithmetic mean weighted by the counts of the daily data combined in each grid box. As a general rule, AIRS L2 retrieved quantities

whose quality indicators are “best” (=0) or “good” (=1) are included in the sums that generate the L3 products. Please see the AIRS V5 L3 quick start guide¹ for more details.

The two new AIRS datasets of the monthly mean tropospheric air temperature and specific humidity profiles for CMIP5 model evaluation were derived from the standard AIRS L3 V5 monthly mean tropospheric air temperature and specific humidity profile products. The original 24 pressure levels for AIRS V5 L3 air temperature profile product are a superset of the 17 CMIP5 mandatory pressure levels and there is no need for reprocessing. However, the original AIRS L3 V5 specific humidity product is reported in terms of layer averages. In order to convert the layer amounts to level amounts, we treated the original layer averages as level amounts at the midpoint in logarithmic pressure of the layers and then logarithmically interpolated them in logarithmic pressure to the desired levels. For the 1000-hPa level, this interpolation was replaced by an extrapolation because there is no layer with a higher midpoint pressure than 1000 hPa. The extrapolation was done logarithmically in logarithmic pressure consistent with the interpolation.

In addition, simple arithmetic means of the daytime and nighttime monthly mean air temperature and specific humidity values were reported in these two AIRS datasets. A minimum of twenty observations per 1° grid box per month was required, except for latitudes beyond +/- 80°. Limits at high latitudes were relaxed to compensate for a much

¹ Available on line: http://disc.sci.gsfc.nasa.gov/AIRS/documentation/v5_docs/AIRS_V5_Release_User_Docs/V5_L3_QuickStart.pdf

lower number of observations because of oblique viewing angle (see Figure 1).

2.3. Caveats for model-observation comparison

There are several caveats that distinguish these AIRS tropospheric air temperature and specific humidity products from those climate model output counterparts.

2.3.1 Cloud effects on sampling

Because AIRS is an infrared instrument, its sensitivity to air temperature and specific humidity is reduced near and below clouds and AIRS coverage is limited by the presence of optically thick clouds. However, the combination of infrared (AIRS) and optically transparent microwave frequencies (AMSU) allows retrieval of high-resolution air temperature and specific humidity profiles for infrared cloud fraction (the product of emissivity and coverage) up to about 70% [Susskind *et al.*, 2006], with a rapid decrease of highest-quality retrievals at higher cloud fraction [Yue *et al.*, 2011]. This infrared sensitivity to clouds makes the AIRS observation scene dependent and in turn causes a spatially inhomogeneous sampling (or number of observations/retrievals) as illustrated in Figure 1, which shows the number of AIRS air temperature and specific humidity retrievals at each pressure level per 1° by 1° grid cell averaged for the whole data record from September 2002 to April 2011. The AIRS number of retrievals is low (~ 60 -100 retrievals per grid cell per month) in cloudy regions, such as the Intertropical Convergence Zone (ITCZ) and the South Pacific Convergence Zone (SPCZ) and the mid-latitude storm tracks around 60° north and south latitude (e.g., north Pacific, north Atlantic, Southern Ocean near Antarctica). The AIRS number of retrievals is high (~ 150 retrievals per grid cell per month) in clear regions, such as the subtropics and some mid-latitude land regions. If cloudiness is correlated with air temperature and/or specific

humidity as demonstrated by some studies [e.g., *John et al.*, 2011; *Kahn et al.*, 2009; *Soden*, 2000; *Tian et al.*, 2004; 2006; 2010; 2012], then AIRS air temperature and specific humidity observations in cloudy regions with low data sampling may not be representative of the mean air temperature and specific humidity over the full range of conditions, and may contain large ‘clear-sky biases’ similar to other infrared sounder data, such as Television Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) and High Resolution Infrared Radiation Sounder (HIRS) [*John et al.*, 2011; 2012; *Lanzante and Gahrs*, 2000; *Soden and Lanzante*, 1996; *Sohn et al.*, 2010]. AMSU mainly helps temperature retrievals but less so humidity retrievals since HSB has not been working since February 2003 and was not used for the AIRS/AMSU specific humidity retrievals. Thus, we expect that the cloud-induced sampling effect is more evident in AIRS specific humidity product than in AIRS air temperature product.

Because different quality flags were used for air temperature and specific humidity retrievals in AIRS V5 L3 data, the numbers of retrievals are different between air temperature and specific humidity in these two AIRS datasets (Figure 1). For air temperature, each retrieval from the top of the atmosphere down to the pressure level of PBest or PGood is averaged into the L3 grids. As a result, the number of air temperature retrievals is higher at higher altitudes and lower at lower altitudes. For specific humidity, Qual_H2O of “best” (=0) or “good” (=1) is applied. As a result, the number of retrievals is constant for specific humidity at all altitudes, except for the near-surface pressure levels over land that were excluded (see below). Because the quality control is much looser for specific humidity than air temperature, the number of retrievals is typically higher for specific humidity than for air temperature (Figure 1). However, in the next

version release (V6) of AIRS L3 data, early 2013, the same quality control will be used for both air temperature and specific humidity and they will have the same number of retrievals.

2.3.2 Temperature bias trend

There is a spurious cooling trend of about 0.05 K/year in the troposphere for AIRS air temperature retrievals. This is suspected to come from incorrect handling of rising atmospheric CO₂ levels [Divakarla *et al.*, 2006], and will be reduced in the next version release (V6) of AIRS data [Dang *et al.*, 2012].

2.3.3 Incomplete diurnal cycle sampling

Because the Aqua satellite is in a Sun-synchronous polar orbit, AIRS samples the atmosphere at two fixed local solar times (1:30 AM and 1:30 PM) at the Equator and cannot fully resolve the diurnal cycle [Parkinson, 2003]. In contrast, typical model monthly mean outputs contain the monthly averaged values of atmospheric variables at a higher fixed local solar time resolution (e.g. every six hours). For air temperature and specific humidity in the oceanic upper troposphere with a small diurnal cycle, this diurnal cycle sampling difference between AIRS and models is not likely a big problem. However, for air temperature and specific humidity in the boundary layer or over land regions strongly influenced by the diurnal cycle, this diurnal cycle sampling difference between AIRS and models should be considered.

2.3.4 Inhomogeneous sampling with latitude

Because the Aqua satellite is in a Sun-synchronous polar orbit, the diurnal sample increases from the Equator to the poles, with the sampling in the polar regions including

all times of day and night [Parkinson, 2003]. Thus, the data sampling is much higher (as frequent as eight times daily) at latitudes between 70° and 83° than at lower latitudes. However, as discussed above, regions poleward of 85° latitudes are viewed obliquely and have low sampling rates, as shown in Figure 1.

2.3.5 Limited time coverage

AIRS went into a safe mode at the end of October 2003 to avoid possible damage from a large solar flare, and did not resume operations until mid-November 2003. Our preparation of these two AIRS products for CMIP5 model evaluation added a requirement of a minimum number of observations (twenty) per month for each grid square from each of ascending and descending orbits. With only a half-month of data in November 2003, many grid cells do not meet this criterion of twenty samples per cell per month. Cells filled during November 2003 include only the first half of the month, leading to a potential bias. A similar safe mode event occurred from January 9–29, 2010. As a result, the January 2010 product has about one third the data of a full month; it too could have a slight bias.

2.3.6 Excluded near-surface retrievals over land

Because of known biases near land surfaces [e.g., Tobin *et al.*, 2006], AIRS data within 100 hPa of the land surface were excluded from these two datasets. As a result, AIRS data are missing for most grids over land for the 1000-hPa and 925-hPa pressure levels and some grids for 850-hPa and 700-hPa pressure levels (Figure 1).

2.3.7 Different vertical resolutions

The AIRS tropospheric air temperature (specific humidity) vertical resolution

ranges between 2.5 km (2.7 km) near the surface and 7.1 km (4.3 km) near the tropopause. On the other hand, CMIP5 vertical resolution ranges from 200 m near the surface to 2-3 km in the upper troposphere. Thus, AIRS cannot well identify the planetary boundary layer height and this may have some impacts on the AIRS and CMIP5 comparisons.

2.4. Data validation

The AIRS L2 tropospheric air temperature and specific humidity measurements have been validated against a variety of other tropospheric air temperature and specific humidity measurements from in situ instruments including aircraft, ship-launched balloons and operational radiosondes [e.g., *Divakarla et al.*, 2006; *Gettelman et al.*, 2004; *Hagan et al.*, 2004; *Nalli et al.*, 2006; *Pierce et al.*, 2006; *Tobin et al.*, 2006; *Wu*, 2009] and model-generated analysis and reanalysis data [e.g., *Divakarla et al.*, 2006]. The AIRS L2 and L3 total water vapor has been compared against several other ground-based and satellite remote sensing measurements as well as reanalysis data having implication for AIRS tropospheric specific humidity profiles [e.g., *Fetzer et al.*, 2006; 2008; *Thomas et al.*, 2011; *Wong et al.*, 2011]. Table 1 summarizes the major findings from these previous validation studies. The AIRS uncertainty estimates were calculated based on the difference between the AIRS retrievals and radiosonde observations. The AIRS sounding accuracy requirements are 1 K per 1 km layers for air temperature and 15% per 2 km layers for specific humidity in the troposphere [*Aumann et al.*, 2003]. Table 1 indicates that these requirements are generally met.

The AIRS L3 tropospheric air temperature and specific humidity measurements have also been compared with radiosondes, reanalyses and Global Positioning System

radio occultation (GPS RO) data in the context of the Madden-Julian Oscillation studies [e.g., *Tian et al.*, 2006; 2010; 2012]. These studies have shown a generally consistent MJO temperature and humidity structure among AIRS, reanalysis and GPS RO data. *Tian et al.* [2012] have also shown that the AIRS L3 upper tropospheric temperature anomalies associated with the MJO are clearly underestimated in the cloudy regions in comparison to GPS RO data. However, the impact of this cloud-induced low sampling on the AIRS L3 tropospheric air temperature and specific humidity climatology is still unclear.

In order to provide an estimate of uncertainty of the AIRS observations, especially those related to cloud-induced low sampling in the cloudy regions, we also analyze the air temperature and specific humidity fields from NASA's Modern Era Retrospective analysis for Research and Applications (MERRA) [*Rienecker et al.*, 2011]. MERRA was generated with version 5.2.0 of the Goddard Earth Observing System (GEOS) atmospheric model and data assimilation system. MERRA spans most of the satellite era and is intended to place observations from NASA's satellites, particularly those available since October 2002 from Aqua, into a climate context. MERRA proposed to improve upon the water cycle as a contribution to the science community and to reanalysis research, since several previous reanalyses have shown significant deficiencies in their representation of critical components of the hydrologic cycle such as precipitation, ocean surface evaporation, and water balance closure (i.e., precipitation minus evaporation) [*Trenberth et al.*, 2011]. Like other reanalysis efforts, MERRA ingests much of the conventional and satellite observational data stream and performs a series of extensive quality checks and bias corrections [*Rienecker et al.*, 2011]. MERRA does assimilate AIRS data, but uses only temperature and humidity sensitive clear-sky

radiances instead of AIRS physical temperature and humidity retrievals. Recent evaluation by *Bosilovich et al.* [2011] indicates that MERRA does demonstrate improved skill of spatial distribution of precipitation, especially in tropical oceanic regions. *Wong et al.* [2011] estimated the atmospheric total water vapor sink, which is equal to the surface water exchange (precipitation minus surface evaporation), using water vapor retrievals from AIRS with the wind fields from MERRA, as well as from the water vapor budget in the MERRA. The AIRS and MERRA data can reproduce the main large-scale patterns of the atmospheric water vapor sink, including the locations and variations of the ITCZ, summertime monsoons, and mid-latitude storm tracks in both hemispheres. The spectra of its regional temporal variations are generally consistent with those from other independent satellite observations including the annual and semiannual cycles, and intraseasonal variations. The study of *Brunke et al.* [2011] also shows that the MERRA is among the “best performing” reanalysis datasets for latent heat and sensible heat fluxes and inertial-dissipation wind stresses. However, there is no single study available to quantify the uncertainties of the MERRA tropospheric air temperature and specific humidity products, especially their vertical structure. Since MERRA may be in some circumstances largely model-based, significant uncertainties in the MERRA data are expected, especially in the water vapor field. Since we do not yet know whether AIRS or MERRA tropospheric air temperature and specific humidity products are more accurate, we assume that differences between AIRS and MERRA indicate their combined uncertainties.

For this study, we use the MERRA monthly mean values of specific humidity (q_v) in kg/kg and air temperature (t) in K from the MERRA history collection

“MAIMCPASM” files, a monthly version of the “MAI3CPASM” or “inst3_3d_asm_Cp”. These MERRA files contain instantaneous basic assimilated fields from the Incremental Analysis Update (IAU) corrector on 42 vertical pressure levels from 1000 to 0.1 hPa, at a reduced resolution spatial grid (i.e., 288-longitude \times 144-latitude with 1.25° by 1.25° resolution).

3. CMIP5 Model Outputs and Comparison Methodology

For AIRS-CMIP5 comparison, we analyze the monthly mean tropospheric air temperature and specific humidity outputs from ‘historical’ experiments, which are defined as simulations of recent past climate [Taylor *et al.*, 2012], of sixteen CMIP5 models. These sixteen CMIP5 models, listed in Table 2, were chosen because they were available at the PCMDI ESG-CET at the time of our analysis. Eight of them (CCSM4, CNRM-CM5, CSIRO-Mk3-6-0, GFDL-CM3, GISS-E2-H, GISS-E2-R, MIROC4h, and MRI-CGCM3) are coupled atmosphere-ocean general circulation models (AOGCMs), while eight others (BCC-CSM1.1, CanESM2, HadGEM2-CC, HadGEM2-ES, IPSL-CM5A-LR, MIROC-ESM, MPI-ESM-LR, and NorESM1-M) are Earth system models (ESMs). AOGCMs respond to specified, time-varying concentrations of various atmospheric constituents (e.g., greenhouse gases) and include an interactive representation of the atmosphere, ocean, land, and sea ice. ESMs are AOGCMs that are coupled to biogeochemical components accounting for the important fluxes of carbon between the ocean, atmosphere, and terrestrial biosphere carbon reservoirs, thereby “closing” the carbon cycle in the models. These ESMs have the capability of using time-evolving emissions of constituents from which concentrations can be computed interactively. They may in some cases also include interactive prognostic aerosol,

chemistry, and dynamical vegetation components.

In this paper, we focus on the climatology of monthly mean tropospheric air temperature and specific humidity from AIRS, MERRA and the multi-model ensemble mean (MMEM) of sixteen CMIP5 model outputs. The MMEM was chosen to highlight the common problems or biases for most models instead of those for an individual model. The AIRS and MERRA climatology was calculated as the nine-year mean from September 2002 to May 2011. The climatology for each CMIP5 model was calculated as the twenty-year mean from January 1986 to December 2005. The CMIP5 MMEM climatology is a simple average of the climatologies of these sixteen CMIP5 models. For a direct comparison, we re-gridded all AIRS, MERRA and CMIP5 data to a standard horizontal grid of 120×60 (longitude \times latitude) with 3° by 3° resolution. Although each CMIP5 model has a different number of vertical pressure levels (from 17 to 25), all 16 CMIP5 models have the 17 CMIP5 mandatory pressure levels from 1000 to 10 hPa. This is also true for MERRA. However, our analyses are restricted for the 8 pressure levels from 1000 to 300 hPa because of the availability of the AIRS data. For both zonal and temporal averages, the missing air temperature and specific humidity values for pressure levels below the surface are neglected in all data sets.

4. Comparison of AIRS, MERRA and CMIP5

4.1. Global zonal mean cross sections

Figure 2 shows the global zonal mean cross sections of climatological mean tropospheric air temperature (K, upper panels) (column 1: AIRS; column 2: MERRA; column 3: CMIP5 MMEM) and their differences (K, lower panels) (column 1: MERRA–AIRS; column 2: CMIP5–AIRS; column 3: CMIP5–MERRA). Similar

quantities for specific humidity are presented in Figure 3 in addition to the relative differences because the relative error provides a better measure of the water vapor's impact on the radiative transfer than does the absolute error [Soden *et al.*, 2005]. These figures can be compared to similar figures in *Pierce et al.* [2006] and *John and Soden* [2007].

As expected, the AIRS data show the well-known vertical and meridional structures of the climatological mean tropospheric air temperature and specific humidity (Figures 2 and 3). There exist strong meridional and vertical gradients in tropospheric air temperature and specific humidity that decrease with both latitude (warm/moist near the Equator and cold/dry near the poles) and altitude (warm/moist near the surface and cold/dry in the upper troposphere). For example, at 1000 hPa, tropospheric air temperature decreases from around 295 K near the Equator to around 255 K near the poles. Near the Equator, tropospheric air temperature decreases from around 295 K at 1000 hPa to around 240 K at 300 hPa. Similarly, at 1000 hPa, tropospheric specific humidity decreases from around 14 g/kg near the Equator to around 1 g/kg near the poles. Near the Equator, tropospheric specific humidity decreases from around 14 g/kg at 1000 hPa to around 0.5 g/kg at 300 hPa.

The general vertical and meridional structures of the climatological mean tropospheric air temperature and specific humidity are, as expected, very similar between AIRS and MERRA and are well captured by the CMIP5 MEM. However, several noticeable differences exist among the CMIP5 model simulations, the MERRA reanalysis and the AIRS observations.

MERRA is slightly warmer than AIRS in the free troposphere, especially in the

tropical upper troposphere (around 300 hPa) and the extratropical lower troposphere, but colder than AIRS in the boundary layer. However, their differences are typically less than 1 K. The largest difference is located in the Southern Ocean near Antarctica and around 925 hPa (see later Fig. 7). In comparison to AIRS, the CMIP5 models exhibit a cold bias of 1–2 K throughout the vast majority of the troposphere. The cold biases tend to be larger in the extratropics than the tropics. The cold biases also increase with height and reach maxima (around 3 K) in the upper troposphere at 300 hPa. The CMIP5 models also have a warm bias (around 2 K) in the boundary layer over the Southern Ocean near Antarctica. However, AIRS yields are typically low in this region due to cloud impacts (Figure 1) [Kahn *et al.*, 2011; Yue *et al.*, 2011] and the AIRS air temperature is subjected to large uncertainties due to cloud and ice influence. Thus, this warm bias in the models may not be real. In comparison to MERRA, the upper-tropospheric cold bias in CMIP5 models is even larger because MERRA is warmer than AIRS in the upper troposphere. However, the warm bias (around 2 K) in the boundary layer over the Southern Ocean near Antarctica is much smaller but still evident. The CMIP5-AIRS air temperature differences are systematically larger than the MERRA-AIRS air temperature differences (which can be seen as the first order estimate of AIRS data uncertainties) and should be considered a robust result and indicate a real bias for the CMIP5 models.

To highlight the CMIP5-AIRS differences and the spread or consistency of CMIP5 model biases, Figure 4 shows the global zonal mean cross sections of climatological mean tropospheric air temperature biases relative to AIRS for sixteen individual CMIP5 models. Figure 4 indicates that the majority of the models (13 of 16) have a tropospheric cold bias similar to the MMEM shown in Figure 2, while only a few

models (3 of 16 - CanESM2, CCSM4, and MIROC4h) have a tropospheric warm bias opposite to MMEM. This indicates that a tropospheric cold bias is one common problem or bias for most CMIP5 models. The upper-tropospheric cold bias in CMIP5 models is also found in CMIP3 models as shown by *John and Soden* [2007] and AMIP-I models as shown by *Gates et al.* [1999]. The prevalence of the upper-tropospheric cold bias in most models suggests a common source of errors in coupled models.

Water vapor differences between AIRS and MERRA are large in many places (Figure 3). MERRA is drier (around 1 g/kg) than AIRS in the tropical (30°S-30°N) boundary layer (below 700 hPa) but wetter (around -1 g/kg) than AIRS in the tropical free troposphere (above 700 hPa) and the extratropical troposphere. The MERRA-AIRS relative differences in the tropical boundary layer is less than 10% but can be greater than 20% in the tropical upper troposphere and the extratropical troposphere. These large differences will be further examined in subsection 4.2.

In comparison to AIRS, the CMIP5 models are too dry (around -1 g/kg) in the tropical lower troposphere (below 600 hPa) but too moist in the tropical upper troposphere (above 500 hPa) and the extratropical troposphere, especially over the southern hemisphere. In terms of relative difference, the models' dry bias in the tropical lower troposphere is around 15% and the wet bias in the tropical upper troposphere and the extratropical troposphere approaches 30% for the MMEM. To highlight the CMIP5-AIRS differences and the spread or consistency of CMIP5 model biases, Figure 5 shows the global zonal mean cross sections of climatological mean tropospheric specific humidity relative biases against AIRS for sixteen individual CMIP5 models. Figure 5 indicates that the majority of the models have a similar cross section for tropospheric

specific humidity bias to the MMEM shown in Figure 3. The specific humidity bias can exceed 100% for some individual models (e.g., MRI-CGCM). This specific humidity bias pattern in CMIP5 models is similar to that in CMIP3 models presented by *Pierce et al.* [2006] and *John and Soden* [2007].

However, the CMIP5-AIRS specific humidity differences are of comparable magnitude to the MERRA-AIRS specific humidity differences, and have similar patterns. As a result, the CMIP5-MERRA specific humidity differences are much smaller than the CMIP5-AIRS specific humidity differences with a different pattern (Figure 3). In comparison to MERRA, CMIP5 models are drier throughout the depth of the tropical troposphere but wetter in the extratropical troposphere, especially over the southern hemisphere. Their relative differences are typically less than 20%. This indicates that the CMIP5-AIRS specific humidity differences are comparable to the uncertainties of the AIRS specific humidity data, as indicated by the MERRA-AIRS specific humidity differences here. So, the CMIP5-AIRS specific humidity differences may not reflect the true CMIP5 model specific humidity biases where they are as large as the MERRA-AIRS specific humidity differences.

4.2. Global spatial maps

Figures 6 and 7 show the global spatial maps of climatological mean tropospheric air temperature (K) at different pressure levels (column 1: AIRS; column 2: MERRA; column 3: CMIP5 MMEM) and its differences among AIRS, MERRA, and CMIP5 (column 1: MERRA-AIRS; column 2: CMIP5-AIRS; column 3: CMIP5-MERRA) . Note that different scales are used for different pressure levels. Figures 6 and 7 show not only the vertical and meridional structures but also the zonal variation of the tropospheric

air temperature. Figure 6 indicates that the zonal gradient of tropospheric air temperature is rather weak in the whole troposphere compared to its vertical and meridional gradients. This is particularly true in the free troposphere (≤ 700 hPa). For example, at 500 hPa, tropospheric air temperature is around 265 K over the whole equatorial belt.

The spatial maps of the climatological mean tropospheric air temperature are similar between AIRS and MERRA and are fairly well captured by the CMIP5 models. However, several noticeable differences exist among the CMIP5 model simulations, the MERRA reanalysis and the AIRS observations similar to the zonal mean cross sections. Furthermore, the maps also show some new information when compared with the zonal mean cross sections. The MERRA-AIRS air temperature differences are generally smaller (less than 1 K) and also tend to circle the globe longitudinally except for a few locations. Over the Saharan, Arabian and Gobi deserts, MERRA is warmer than AIRS by about 2 K, likely due to issues with AIRS retrievals over deserts. Over the Southern Ocean near Antarctica, the large MERRA-AIRS air temperature differences (around 2 K) are probably due to the difficulty of the AIRS retrievals in handling the cloud and ice in this region. Near the coasts of Peru, California, northeast Asia and northeast Canada, the low clouds may cause some biases in the AIRS air temperature retrievals in the boundary layer. Of course, MERRA is mostly model-driven in some regions and may be an issue as well. The CMIP5 MEM cold biases relative to AIRS and MERRA tend to circle around the globe longitudinally and there are no significant differences between ocean and land regions especially in the free troposphere.

Similar to Figure 6 and 7, Figures 8 and 9 show the global spatial maps of climatological mean tropospheric specific humidity at different pressure levels (column

1: AIRS; column 2: MERRA; column 3: CMIP5 MEMEM) and its differences among AIRS, MERRA, and CMIP5 (column 1: MERRA-AIRS; column 2: CMIP5-AIRS; column 3: CMIP5-MERRA). In addition, Figure 10 shows the global spatial maps of climatological mean tropospheric specific humidity relative differences among AIRS, MERRA, and CMIP5 at different pressure levels (column 1: (MERRA-AIRS)/AIRS; column 2: (CMIP5-AIRS)/AIRS; column 3: (CMIP5-MERRA)/MERRA). Note that different scales are used for different pressure levels. Figures 8, 9 and 10 show not only the vertical and meridional structures but also the zonal variation of the tropospheric specific humidity. Figure 8 indicates that the zonal gradient of tropospheric specific humidity is strong both within the boundary layer (≥ 850 hPa) and in the free troposphere (≤ 700 hPa) in contrast to that of the tropospheric air temperature. For example, at 500 hPa, tropospheric specific humidity is around 3 g/kg over the equatorial Indian Ocean and western Pacific but around 1.5 g/kg over the equatorial eastern Pacific. The spatial distribution of tropospheric specific humidity is a strong indicator of the atmospheric general circulation and deep convection patterns. It is moist over the regions of deep convection and the ascending branches of the Hadley and Walker circulations, such as the ITCZ and the SPCZ (including equatorial Africa, equatorial South America and the Maritime Continents) and the northwestern Pacific and Atlantic, but it is dry over the clear-sky regions and descending branches of the Hadley and Walker circulations, such as the northeastern subtropical Pacific and Atlantic.

The spatial maps of the climatological mean tropospheric specific humidity differences between MERRA and AIRS reveal similar information to the zonal mean cross sections. For example, MERRA is drier in the tropical boundary layer but wetter in

the tropical free troposphere and the extratropical troposphere when compared to AIRS. These differences can be greater than 40% in the upper troposphere. Furthermore, the maps show some new information that zonal mean cross sections are not able to do. For example, the MERRA-AIRS specific humidity differences are mainly located in the deep convective cloudy regions, such as the ITCZ, the SPCZ and the mid-latitude storm tracks. This indicates that the low sampling of AIRS in the cloudy regions due to cloud impact may be the key reason for the large differences between AIRS and MERRA specific humidity. Also the AIRS and MERRA differences are bigger in the specific humidity field than those in the air temperature field indicating that the low sampling bias may be the main reason for these differences because the spatial gradient between the cloudy and clear regions is big for specific humidity but small for air temperature. This indicates that the AIRS tropospheric air temperature and specific humidity climatology in cloudy regions may contain a significant ‘clear-sky bias’ similar to other infrared sounder data, such as TOVS and HIRS [John *et al.*, 2011; 2012; Lanzante and Gahrs, 2000; Soden and Lanzante, 1996; Sohn *et al.*, 2010].

To further support this speculation, Figure 11 shows global maps of one-year (1 September 2008 – 31 August 2009) mean MERRA-AIRS specific humidity (left panels) and air temperature (right panels) differences (MERRA-AIRS) at 500 hPa when MERRA is subsampled for different cloud amount thresholds (top panels: MERRA cloud fraction (CF)<0.5; middle panels: CF<0.8; bottom panels: CF≤1.0). The one-year MERRA-AIRS specific humidity and air temperature differences when MERRA is sampled at the all CF are very similar to the nine-year MERRA-AIRS specific humidity and air temperature differences shown in Figures 7 and 10 except the magnitudes are much smaller for the

nine-year average. This demonstrates that it is reasonable to use one-year MERRA and AIRS data to test the cloud-related sampling impact on the nine-year MERRA-AIRS differences. Figure 11 indicates that when MERRA is subsampled only for CF less than certain thresholds (0.8 or 0.5) similar to the AIRS sampling, the MERRA-AIRS specific humidity and air temperature differences are significantly reduced or even change signs, such as over the tropical and mid-latitude convective regions with high amount of clouds. This indicates that the MERRA-AIRS specific humidity and air temperature differences for MERRA with full CF are largely due to the low sampling of AIRS in the cloudy regions. More detailed analyses regarding the cloud-induced biases in AIRS specific humidity and air temperature are needed and will be reported in a future paper.

In comparison to AIRS and MERRA, the general spatial pattern of the climatological mean tropospheric specific humidity is also captured by the CMIP5 MEMM (Figure 8). However, there are big differences between the CMIP5 MEMM and AIRS/MERRA and the spatial patterns of the differences depend strongly on the chosen reference dataset. In comparison to AIRS, CMIP5 models are generally wetter in the upper troposphere and drier in the lower troposphere. However, CMIP5 models are generally drier in the free troposphere and no moisture bias in the lower troposphere in comparison to MERRA. This is consistent with the zonal mean cross sections shown earlier. This again indicates that the CMIP5-AIRS specific humidity differences are within the large uncertainties of the AIRS specific humidity data as indicated by the MERRA-AIRS specific humidity differences here and may not indicate the real CMIP5 model specific humidity biases. Of course, MERRA is mostly model-driven in some regions and may be an issue as well.

However, the spatial maps of the climatological mean tropospheric specific humidity differences between CMIP5 and AIRS/MERRA (in Figure 10) reveal some new information that zonal mean cross sections are not able to do. For example, regardless of the reference dataset chosen, AIRS or MERRA, the CMIP5 models have a double-ITCZ problem that has persisted in coupled GCMs for a long time [e.g. *Lin, 2007; Liu et al., 2012; Zhang et al., 2007*]. The double-ITCZ problem typically refers to the pattern of excessive precipitation off the equator but insufficient precipitation on the equator, which is often associated with an excessive and overly narrow sea surface temperature (SST) cold tongue that extends too far west into the western Pacific in coupled models [*Lin, 2007*]. Figure 10 indicates that the double-ITCZ problem shows up in the water vapor field too. Over the tropical Pacific, the model-simulated moist SPCZ extends far too east while the model-simulated equatorial dry tongue extends far too west. As a result, the models are too dry over the equatorial convective regions while too moist over both sides off the Equator. The equatorial dry bias is typically less than 20% because of the large mean specific humidity. However, the off-equatorial moist bias can be extremely large (reaching 200%) in terms of the relative difference because of the small mean specific humidity. The equatorial dry bias and the off-equatorial moist bias extend throughout the whole troposphere from 1000 hPa to 300 hPa.

To highlight the robustness of the double-ITCZ problem in the CMIP5 models, Figure 12 shows the global spatial maps of climatological mean tropospheric specific humidity relative biases relative to AIRS for sixteen individual CMIP5 models at 500 hPa (other levels show similar results). There is a considerable spread among models in terms of the overall (or zonal mean) moist or dry biases. Some models (CNRM-CM5, GISS-

E2-H, GISS-E2-R, HadGEM2-CC, and HadGEM2-ES) are drier than AIRS while others (BCC-CSm1, CanESM2, CCSM4, MRI-CGCM3, and NorESM1-M) may be wetter than AIRS. Nevertheless, all sixteen models have the double-ITCZ problem in some detail. All models are too dry in the equatorial convective regions, such as the western Pacific warm pool, the ITCZ and equatorial South America, but they are all too moist over both off-equatorial sides of these convective regions, especially over the tropical Pacific. The moist bias is extremely high in the off-equatorial eastern Pacific and can reach around 200%. Given the dry biases of the AIRS specific humidity climatology in the cloudy regions due to its low sampling caused by clouds, the model dry bias in the equatorial convective regions may be underestimated as well as the double-ITCZ problem in the models. In summary, the double-ITCZ is a universal problem among all sixteen CMIP5 models examined here.

5. Summary and Conclusions

The peer-reviewed analyses of multi-model outputs from the CMIP5 experiments [Taylor *et al.*, 2012] will be the most important basis for the IPCC AR5, scheduled for publication in 2014. To increase the fidelity of the IPCC AR5, NASA/JPL and DOE/PCMDI have initiated an Obs4MIPs project for CMIP5 model evaluation [Gleckler *et al.*, 2011; Teixeira *et al.*, 2011]. The Obs4MIPs refers to a limited collection of well-established and well-documented datasets that have been organized according to the CMIP5 model output requirements and made available on the PCMDI ESG-CET for CMIP5 model evaluation. As part of the Obs4MIPs project, the AIRS project has generated two new AIRS datasets for CMIP5 model evaluation including the monthly mean tropospheric air temperature and specific humidity profile products from the

standard AIRS L3 V5 monthly mean tropospheric air temperature and specific humidity profile products. These two AIRS datasets are provided for each calendar month from September 2002 to May 2011, on a global spatial grid at 1°-longitude by 1°-latitude resolution, and on the 17 CMIP5 mandatory vertical pressure levels with valid data only for the 8 pressure levels from 1000 to 300 hPa (1000, 925, 850, 700, 600, 500, 400, and 300 hPa).

By examining the global zonal mean cross sections and global spatial maps of climatological mean tropospheric air temperature and specific humidity from AIRS, we find that the nine-year AIRS data show the well-known climatological features of the tropospheric air temperature and specific humidity. For example, tropospheric air temperature and specific humidity have strong meridional and vertical gradients. They both decrease with both latitude (warm/moist near the Equator and cold/dry near the poles) and altitude (warm/moist near the surface and cold/dry in the upper troposphere). The zonal air temperature gradient is rather weak especially in the free troposphere (≤ 700 hPa) but the zonal specific humidity gradient is strong both within the boundary layer (≥ 850 hPa) and in the free troposphere (≤ 700 hPa). The spatial distribution of tropospheric specific humidity is a strong indicator of the atmospheric general circulation and deep convection pattern. It is moist over the regions of deep convection and ascending branches of Hadley and Walker circulation, while it is dry over the clear regions and descending branches of Hadley and Walker circulation.

In comparison to MERRA, AIRS is colder in the free troposphere but warmer in the boundary layer with differences typically less than 1 K. However, there are a few locations that the AIRS air temperature retrievals may have large uncertainties, such as

the Saharan, Arabian and Gobi deserts, the Southern Ocean near Antarctica, and the coasts of Peru, California, northeast Asia and northeast Canada. In comparison to MERRA, AIRS is wetter ($\sim 10\%$) in the tropical boundary layer but drier (around 30%) in the tropical free troposphere and the extratropical troposphere. In particular, the large specific humidity differences between AIRS and MERRA are mainly located in the cloudy regions, such as the ITCZ, the SPCZ and the mid-latitude storm tracks. This indicates that the low sampling of AIRS in the cloudy regions may be the main reason for these differences.

In comparison to AIRS and MERRA, the sixteen CMIP5 models examined here can generally simulate the climatological features of tropospheric air temperature and specific humidity well. However, several noticeable differences exist between the CMIP5 model simulations, the MERRA reanalysis and the AIRS observations. In comparison to AIRS, the CMIP5 models have a cold bias (around 2 K) in the troposphere, especially in the extratropical upper troposphere (around 3 K), and a warm bias in the boundary layer over the Southern Ocean near Antarctica. In comparison to MERRA, the upper-tropospheric cold bias in the CMIP5 models is even larger but the warm bias in the boundary layer over the Southern Ocean near Antarctica is much smaller although still evident. Thus, the upper-tropospheric cold bias should be considered to be a robust result and a real bias for the CMIP5 models. Examining individual models indicates that this upper-tropospheric bias is found in the majority (thirteen) of the sixteen CMIP5 models and is one common problem or bias for most CMIP5 models.

For the global zonal mean cross section, in comparison to AIRS, the CMIP5 models are too dry in the tropical lower troposphere but too moist in the tropical upper

troposphere and extratropical troposphere, especially over the southern hemisphere. However, in comparison to MERRA, CMIP5 models are too dry in the tropical whole troposphere but too moist in the extratropical troposphere, especially over the southern hemisphere. The CMIP5-AIRS specific humidity differences are within or at least comparable to the large uncertainties of the AIRS specific humidity data as indicated by the MERRA-AIRS specific humidity differences here and may not indicate the true CMIP5 model specific humidity bias.

Examining the spatial maps of tropospheric specific humidity between AIRS/MERRA and CMIP5, we do find salient water vapor biases in the CMIP5 models. No matter which reference dataset is used, AIRS or MERRA, the CMIP5 models have the double-ITCZ problem that persists in the coupled GCMs for a long time. Over the tropical Pacific, the model-simulated moist SPCZ extends too far east while the model-simulated equatorial dry tongue extends too far west. As a result, the models are too dry over the equatorial convective regions while too moist over both sides off the Equator. The equatorial dry bias is typically less than 20% because of the large mean specific humidity. However, the off-equatorial relative moist bias can be extremely large (reaching 200%) because of the small mean specific humidity. The equatorial dry bias and the off-equatorial moist bias all extend through the whole troposphere from 1000 hPa to 300 hPa. Examining individual model indicates that the double-ITCZ exists in all sixteen CMIP5 models and is a common problem for all CMIP5 models.

Acknowledgments

This research was performed at Jet Propulsion Laboratory (JPL), California Institute of Technology (Caltech), under a contract with National Aeronautics and Space

Administration (NASA). It was supported by the Atmospheric Infrared Sounder (AIRS) project at JPL. We acknowledge the World Climate Research Program's (WCRP) Working Group on Coupled Modeling (WGCM), which is responsible for Coupled Model Intercomparison Project (CMIP), and the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison (PCMDI), which provides coordinating support and leads development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We thank the climate modeling groups around the world including BCC, CCCMA, CNRM-CERFACS, CSIRO-QCCCE, GFDL, GISS, IPSL, MIROC, MOHC, MPI-M, MRI, NCAR, and NCC (listed in Table 2 of this paper) for producing and making available their model output. We also want to thank Duane Waliser, Evan Fishbein, Sun Wong, Hui Su, Jonathan Jiang, Steve Ghan (editor) and three anonymous reviewers for valuable suggestions. © 2012. All rights reserved.

References

- Allan, R. P., M. A. Ringer, and A. Slingo (2003), Evaluation of moisture in the Hadley Centre climate model using simulations of HIRS water-vapour channel radiances, *Q. J. Roy. Meteor. Soc.*, *129*(595), 3371-3389, 10.1256/qj.02.217.
- Arora, V. K., J. F. Scinocca, G. J. Boer, et al. (2011), Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases, *Geophys. Res. Lett.*, *38*, L05805, doi:10.1029/2010gl046270.
- Aumann, H. H., M. T. Chahine, C. Gautier, et al. (2003), AIRS/AMSU/HSB on the aqua mission: Design, science objectives, data products, and processing systems, *IEEE Trans. Geosci. Remote Sens.*, *41*(2), 253-264, 10.1109/tgrs.2002.808356.
- Bates, J. J., and D. L. Jackson (1997), A comparison of water vapor observations with

- AMIP I simulations, *J. Geophys. Res.*, *102*(D18), 21837-21852, 10.1029/97jd01769.
- Bony, S., R. Colman, V. M. Kattsov, et al. (2006), How well do we understand and evaluate climate change feedback processes?, *J. Climate*, *19*(15), 3445-3482, 10.1175/jcli3819.1.
- Bosilovich, M. G., F. R. Robertson, and J. Y. Chen (2011), Global energy and water budgets in MERRA, *J. Climate*, *24*(22), 5721-5739, 10.1175/2011jcli4175.1.
- Brogniez, H., R. Roca, and L. Picon (2005), Evaluation of the distribution of subtropical free tropospheric humidity in AMIP-2 simulations using METEOSAT water vapor channel data, *Geophys. Res. Lett.*, *32*(19), L19708, doi:10.1029/2005gl024341.
- Brunke, M. A., Z. Wang, X. B. Zeng, et al. (2011), An assessment of the uncertainties in ocean surface turbulent fluxes in 11 reanalysis, satellite-derived, and combined global datasets, *J. Climate*, *24*(21), 5469-5493, 10.1175/2011jcli4223.1.
- Chahine, M. T., T. S. Pagano, H. H. Aumann, et al. (2006), AIRS: Improving weather forecasting and providing new data on greenhouse gases, *Bull. Amer. Meteor. Soc.*, *87*(7), 911-926, 10.1175/bams-87-7-911.
- Dang, H. V. T., B. Lambrigtsen, E. Manning, et al. (2012), AIRS/AMSU/HSB version 6 level 2 performance and test report.
- Divakarla, M. G., C. D. Barnett, M. D. Goldberg, et al. (2006), Validation of Atmospheric Infrared Sounder temperature and water vapor retrievals with matched radiosonde measurements and forecasts, *J. Geophys. Res.*, *111*(D9), D09s15, doi:10.1029/2005jd006116.
- Fetzer, E. J., B. H. Lambrigtsen, A. Eldering, et al. (2006), Biases in total precipitable water vapor climatologies from Atmospheric Infrared Sounder and Advanced Microwave Scanning Radiometer, *J. Geophys. Res.*, *111*(D9), D09s16, doi:10.1029/2005jd006598.

- Fetzer, E. J., W. G. Read, D. Waliser, et al. (2008), Comparison of upper tropospheric water vapor observations from the Microwave Limb Sounder and Atmospheric Infrared Sounder, *J. Geophys. Res.*, *113*(D22), D22110, doi:10.1029/2008jd010000.
- Gates, W. L., J. S. Boyle, C. Covey, et al. (1999), An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I), *Bull. Amer. Meteor. Soc.*, *80*(1), 29-55, 10.1175/1520-0477(1999)080<0029:aootro>2.0.co;2.
- Gent, P. R., G. Danabasoglu, L. J. Donner, et al. (2011), The Community Climate System Model version 4, *J. Climate*, *24*(19), 4973-4991, 10.1175/2011jcli4083.1.
- Gettelman, A., E. M. Weinstock, E. J. Fetzer, et al. (2004), Validation of Aqua satellite data in the upper troposphere and lower stratosphere with in situ aircraft instruments, *Geophys. Res. Lett.*, *31*(22), L22107, doi:10.1029/2004gl020730.
- Gettelman, A., W. D. Collins, E. J. Fetzer, et al. (2006), Climatology of upper-tropospheric relative humidity from the Atmospheric Infrared Sounder and implications for climate, *J. Climate*, *19*(23), 6104-6121, 10.1175/jcli3956.1.
- Gleckler, P. J., R. Ferraro, and D. E. Waliser (2011), Improving use of satellite data in evaluating climate models, *EOS*, *92*(20), 172, 10.1029/2011EO200005.
- Griffies, S. M., M. Winton, L. J. Donner, et al. (2011), The GFDL CM3 coupled climate model: Characteristics of the ocean and sea ice simulations, *J. Climate*, *24*(13), 3520-3544, 10.1175/2011jcli3964.1.
- Hagan, D. E., C. R. Webster, C. B. Farmer, et al. (2004), Validating AIRS upper atmosphere water vapor retrievals using aircraft and balloon in situ measurements, *Geophys. Res. Lett.*, *31*(21), L21103, doi:10.1029/2004gl020302.
- Held, I. M., and B. J. Soden (2000), Water vapor feedback and global warming, *Annu. Rev. Energy Environ.*, *25*, 441-475, 10.1146/annurev.energy.25.1.441.
- Hilton, F., R. Armante, T. August, et al. (2012), Hyperspectral earth observation from IASI five years of accomplishments, *Bull. Amer. Meteor. Soc.*, *93*(3), 347-370,

10.1175/bams-d-11-00027.1.

Jiang, J. H., H. Su, C. Zhai, et al. (2012), Evaluation of cloud and water vapor simulations in CMIP5 climate models using NASA 'A-Train' satellite observations, *J. Geophys. Res.*, *117*(D14), D14105, 10.1029/2011jd017237.

John, V. O., and B. J. Soden (2007), Temperature and humidity biases in global climate models and their impact on climate feedbacks, *Geophys. Res. Lett.*, *34*(18), L18704, doi:10.1029/2007gl030429.

John, V. O., G. Holl, R. P. Allan, et al. (2011), Clear-sky biases in satellite infrared estimates of upper tropospheric humidity and its trends, *J. Geophys. Res.*, *116*, D14108, doi:10.1029/2010jd015355.

John, V. O., G. Holl, S. A. Buehler, et al. (2012), Understanding intersatellite biases of microwave humidity sounders using global simultaneous nadir overpasses, *J. Geophys. Res.*, *117*, D02305, doi:10.1029/2011jd016349.

Jones, C. D., J. K. Hughes, N. Bellouin, et al. (2011), The HadGEM2-ES implementation of CMIP5 centennial simulations, *Geosci. Model Dev.*, *4*(3), 543-570, 10.5194/gmd-4-543-2011.

Kahn, B. H., A. Gettelman, E. J. Fetzer, et al. (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., S. L. Nasiri, M. M. Schreier, et al. (2011), Impacts of subpixel cloud heterogeneity on infrared thermodynamic phase assessment, *J. Geophys. Res.*, *116*, D20201, doi:10.1029/2011jd015774.

Lambrigtsen, B. H. (2003), Calibration of the AIRS microwave instruments, *IEEE Trans. Geosci. Remote Sens.*, *41*(2), 369-378, 10.1109/tgrs.2002.808247.

Lanzante, J. R., and G. E. Gahrs (2000), The "clear-sky bias" of TOVS upper-tropospheric humidity, *J. Climate*, *13*(22), 4034-4041, 10.1175/1520-0442(2000)013<4034:tcsbot>2.0.co;2.

- Lin, J. L. (2007), The double-ITCZ problem in IPCC AR4 coupled GCMs: Ocean-atmosphere feedback analysis, *J. Climate*, 20(18), 4497-4525, 10.1175/jcli4272.1.
- Liu, H. L., M. H. Zhang, and W. Y. Lin (2012), An investigation of the initial development of the double-ITCZ warm SST biases in the CCSM, *J. Climate*, 25(1), 140-155, 10.1175/2011jcli4001.1.
- Maddy, E. S., and C. D. Barnett (2008), Vertical resolution estimates in version 5 of AIRS operational retrievals, *IEEE Trans. Geosci. Remote Sens.*, 46(8), 2375-2384, 10.1109/tgrs.2008.917498.
- Nalli, N. R., P. Clemente-Colon, P. J. Minnett, et al. (2006), Ship-based measurements for infrared sensor validation during aerosol and ocean science expedition 2004, *J. Geophys. Res.*, 111(D9), D09s04, doi:10.1029/2005jd006385.
- Parkinson, C. L. (2003), Aqua: An earth-observing satellite mission to examine water and other climate variables, *IEEE Trans. Geosci. Remote Sens.*, 41(2), 173-183, 10.1109/tgrs.2002.808319.
- Pierce, D. W., T. P. Barnett, E. J. Fetzer, et al. (2006), Three-dimensional tropospheric water vapor in coupled climate models compared with observations from the AIRS satellite system, *Geophys. Res. Lett.*, 33(21), L21701, doi:10.1029/2006gl027060.
- Raddatz, T. J., C. H. Reick, W. Knorr, et al. (2007), Will the tropical land biosphere dominate the climate-carbon cycle feedback during the twenty-first century?, *Climate Dyn.*, 29(6), 565-574, 10.1007/s00382-007-0247-8.
- Randall, D. A., R. A. Wood, S. Bony, et al. (2007), Climate models and their evaluation, Ch. 8 in *Climate change 2007: The physical science basis. Contribution of working group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon, et al., pp. 589-662, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Read, W. G., A. Lambert, J. Bacmeister, et al. (2007), Aura Microwave Limb Sounder

- upper tropospheric and lower stratospheric H₂O and relative humidity with respect to ice validation, *J. Geophys. Res.*, *112*(D24), D24s35, doi:10.1029/2007jd008752.
- Rienecker, M. M., M. J. Suarez, R. Gelaro, et al. (2011), MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, *J. Climate*, *24*(14), 3624-3648, 10.1175/jcli-d-11-00015.1.
- Rotstayn, L. D., M. A. Collier, M. R. Dix, et al. (2010), Improved simulation of Australian climate and ENSO-related rainfall variability in a global climate model with an interactive aerosol treatment, *Int. J. Climatol.*, *30*(7), 1067-1088, 10.1002/joc.1952.
- Sakamoto, T. T., Y. Komuro, M. Ishii, et al. (2012), MIROC4h—a new high-resolution atmosphere-ocean coupled general circulation model, *J. Meteorol. Soc. Jpn*, in press, available at <http://web.gps.caltech.edu/~btian/tmp/>.
- Schmidt, G. A., R. Ruedy, J. E. Hansen, et al. (2006), Present-day atmospheric simulations using GISS ModelE: Comparison to in situ, satellite, and reanalysis data, *J. Climate*, *19*(2), 153-192, 10.1175/jcli3612.1.
- Soden, B. J., and F. P. Bretherton (1994), Evaluation of water-vapor distribution in general-circulation models using satellite-observations, *J. Geophys. Res.*, *99*(D1), 1187-1210, 10.1029/93jd02912.
- Soden, B. J., and J. R. Lanzante (1996), An assessment of satellite and radiosonde climatologies of upper-tropospheric water vapor, *J. Climate*, *9*(6), 1235-1250, 10.1175/1520-0442(1996)009<1235:aaosar>2.0.co;2.
- Soden, B. J. (2000), The diurnal cycle of convection, clouds, and water vapor in the tropical upper troposphere, *Geophys. Res. Lett.*, *27*(15), 2173-2176, 10.1029/2000gl011436.
- Soden, B. J., D. L. Jackson, V. Ramaswamy, et al. (2005), The radiative signature of upper tropospheric moistening, *Science*, *310*(5749), 841-844,

10.1126/science.1115602.

Soden, B. J., and I. M. Held (2006), An assessment of climate feedbacks in coupled ocean-atmosphere models, *J. Climate*, *19*(14), 3354-3360, 10.1175/jcli3799.1.

Sohn, B. J., T. Nakajima, M. Satoh, et al. (2010), Impact of different definitions of clear-sky flux on the determination of longwave cloud radiative forcing: NICAM simulation results, *Atmos. Chem. Phys.*, *10*(23), 11641-11646, 10.5194/acp-10-11641-2010.

Stephens, G. L., D. G. Vane, R. J. Boain, et al. (2002), The CloudSat mission and the A-Train - a new dimension of space-based observations of clouds and precipitation, *Bull. Amer. Meteor. Soc.*, *83*(12), 1771-1790, 10.1175/bams-83-12-1771.

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, *IEEE Trans. Geosci. Remote Sens.*, *41*(2), 390-409, 10.1109/tgrs.2002.808236.

Susskind, J., C. Barnet, J. Blaisdell, et al. (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*(D9), D09s17, doi:10.1029/2005jd006272.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, *Bull. Amer. Meteor. Soc.*, *93*(4), 485-498, 10.1175/bams-d-11-00094.1.

Teixeira, J., D. E. Waliser, R. Ferraro, et al. (2011), Satellite observations for CMIP5 simulations, *CLIVAR Exchanges*, *16*(2).

Thomas, I. D., M. A. King, P. J. Clarke, et al. (2011), Precipitable water vapor estimates from homogeneously reprocessed GPS data: An intertechnique comparison in Antarctica, *J. Geophys. Res.*, *116*, D04107, doi:10.1029/2010jd013889.

Tian, B., B. J. Soden, and X. Q. Wu (2004), Diurnal cycle of convection, clouds, and water vapor in the tropical upper troposphere: Satellites versus a general

- circulation model, *J. Geophys. Res.*, *109*(D10), D10101, doi:10.1029/2003jd004117.
- Tian, B., D. E. Waliser, E. J. Fetzer, et al. (2006), Vertical moist thermodynamic structure and spatial-temporal evolution of the MJO in AIRS observations, *J. Atmos. Sci.*, *63*(10), 2462-2485, 10.1175/jas3782.1.
- Tian, B., D. E. Waliser, E. J. Fetzer, et al. (2010), Vertical moist thermodynamic structure of the Madden-Julian Oscillation in Atmospheric Infrared Sounder retrievals: An update and a comparison to ECMWF interim re-analysis, *Mon. Wea. Rev.*, *138*(12), 4576-4582, 10.1175/2010mwr3486.1.
- Tian, B., C. O. Ao, D. E. Waliser, et al. (2012), Intraseasonal temperature variability in the upper troposphere and lower stratosphere from the GPS radio occultation measurements, *J. Geophys. Res.*, *117*(D15), D15110, doi:10.1029/2012jd017715.
- Tobin, D. C., H. E. Revercomb, R. O. Knuteson, et al. (2006), Atmospheric Radiation Measurement site atmospheric state best estimates for Atmospheric Infrared Sounder temperature and water vapor retrieval validation, *J. Geophys. Res.*, *111*(D9), D09s14, doi:10.1029/2005jd006103.
- Trenberth, K. E., J. T. Fasullo, and J. Mackaro (2011), Atmospheric moisture transports from ocean to land and global energy flows in reanalyses, *J. Climate*, *24*(18), 4907-4924, 10.1175/2011jcli4171.1.
- Voltaire, A., E. Sanchez-Gomez, D. Salas y Mélia, et al. (2012), The CNRM-CM5.1 global climate model: Description and basic evaluation, *Climate Dyn.*, submitted, available at <http://web.gps.caltech.edu/~btian/tmp/>.
- Watanabe, S., T. Hajima, K. Sudo, et al. (2011), MIROC-ESM 2010: Model description and basic results of CMIP5-20c3m experiments, *Geosci. Model Dev.*, *4*(4), 845-872, 10.5194/gmd-4-845-2011.
- Waters, J. W., L. Froidevaux, R. S. Harwood, et al. (2006), The Earth Observing System Microwave Limb Sounder (EOS MLS) on the Aura satellite, *IEEE Trans. Geosci.*

- Remote Sens.*, 44(5), 1075-1092, 10.1109/tgrs.2006.873771.
- Wong, S., E. J. Fetzer, B. H. Kahn, et al. (2011), Closing the global water vapor budget with AIRS water vapor, MERRA reanalysis, TRMM and GPCP precipitation, and gsstf surface evaporation, *J. Climate*, 24(24), 6307-6321, 10.1175/2011jcli4154.1.
- Wu, L. G. (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.
- Wu, T. W., and coauthors (2012), The 20th century global carbon cycle from the Beijing Climate Center Climate System Model (BCC-CSM). *Climate Dyn.*, submitted, available at <http://web.gps.caltech.edu/~btian/tmp/>.
- Yue, Q., B. H. Kahn, E. J. Fetzer, et al. (2011), Relationship between marine boundary layer clouds and lower tropospheric stability observed by AIRS, CloudSat, and CALIOP, *J. Geophys. Res.*, 116, D18212, doi:10.1029/2011jd016136.
- Yukimoto, S., Y. Adachi, M. Hosaka, et al. (2012), A new global climate model of the Meteorological Research Institute: MRI-CGCM3-model description and basic performance, *J. Meteorol. Soc. Jpn*, 90A, 23-64, 10.2151/jmsj.2012-A02.
- Zhang, X. H., W. Y. Lin, and M. H. Zhang (2007), Toward understanding the double Intertropical Convergence Zone pathology in coupled ocean-atmosphere general circulation models, *J. Geophys. Res.*, 112(D12), D12102, doi:10.1029/2006jd007878.
- Zhang, Z. S., K. Nisancioglu, M. Bentsen, et al. (2012), Pre-industrial and mid-pliocene simulations with NorESM-L, *Geosci. Model Dev.*, 5(2), 523-533, 10.5194/gmd-5-523-2012.

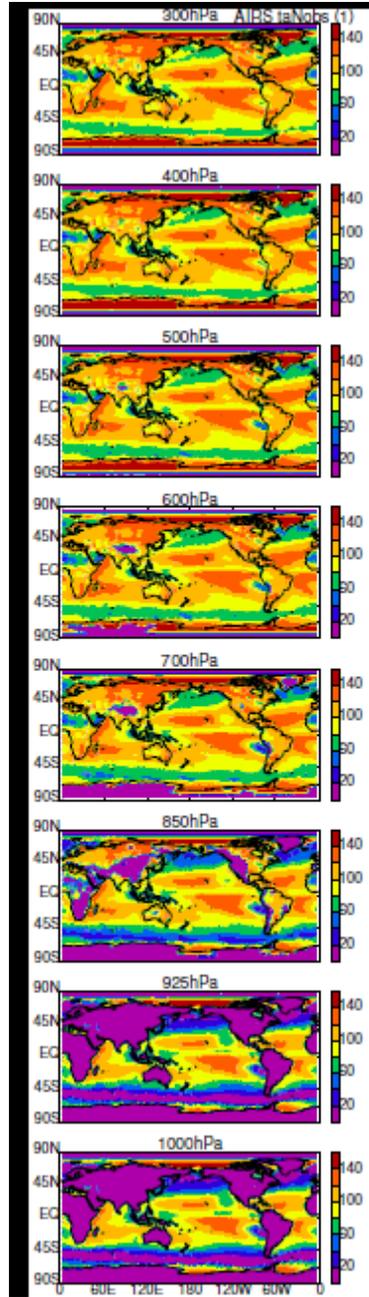
Table 1: AIRS L2 tropospheric air temperature (t_a , K) and specific humidity (h_{us} , kg/kg) uncertainty estimates for different geophysical conditions

Geophysical conditions	t_a	h_{us}
Non-polar ocean, surface to 300 hPa	1K	15–25%
Non-polar land, 2 km to 300 hPa	1K	15–25%
Non-polar land, surface to 2 km	1–2 K	30–40%
Polar	1–2 K	30–40%

Table 2: List of CMIP5 models used in this study

#	Modeling Center (or Group)	Institution ID	Model Name	Type	Horizontal Resolution	Vertical Resolution	Reference
1	Beijing Climate Center, China Meteorological Administration	BCC	BCC-CSM1.1	ESM	2.8°x2.8° (T42L26)	17 standard pressure levels	Wu et al. [2012]
2	Canadian Center for Climate Modeling and Analysis, Canada	CCCMA	CanESM2	ESM	2.8°x2.8° (T63L35)	22 pressure levels (17 std + 5 ext)	Arora et al. [2011]
3	National Center for Atmospheric Research	NCAR	CCSM4	AOGCM	1.25°x0.9°	17 standard pressure levels (L26)	Gent et al. [2011]
4	Centre National de Recherches Meteorologiques / Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique	CNRM-CERFACS	CNRM-CM5	AOGCM	1.4°x1.4° (T127L31)	17 standard pressure levels	Voltaire et al. [2012]
5	Australian Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence	CSIRO-QCCCE	CSIRO-Mk3-6-0	AOGCM	1.875°x1.875° (T63)	17 standard pressure levels (L18)	Rotstayn et al. [2010]
6	NOAA Geophysical Fluid Dynamics Laboratory	NOAA GFDL	GFDL-CM3	AOGCM	2.5°x2.0°	23 pressure levels (17 std + 6 ext)	Griffles et al. [2011]
7	NASA Goddard Institute for Space Studies	NASA GISS	GISS-E2-H	AOGCM	2.5°x2.0°	17 standard pressure levels (L40)	Schmidt et al. [2006]
8	NASA Goddard Institute for Space Studies	NASA GISS	GISS-E2-R	AOGCM	2.5°x2.0°	17 standard pressure levels (L40)	Schmidt et al. [2006]
9	Met Office Hadley Centre	MOHC	HadGEM2-CC	ESM	1.875°x1.25°	23 pressure levels (17 std + 6 ext)	Jones et al. [2011]
10	Met Office Hadley Centre	MOHC	HadGEM2-ES	ESM	1.875°x1.25°	17 standard pressure levels	Jones et al. [2011]
11	Institut Pierre-Simon Laplace	IPSL	IPSL-CM5A-LR	ESM	3.75°x1.9°	17 standard pressure levels (L39)	
12	Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Atmosphere and Ocean Research Institute	MIROC	MIROC4h	AOGCM	0.56°x0.56° (T213L56)	22 pressure levels (17 std + 5 ext)	Sakamoto et al. [2012]

	(AORI)(The University of Tokyo), and National Institute for Environmental Studies (NIES)						
13	AORI, NIES, and JAMSTEC	MIROC	MIROC-ESM	ESM	2.8°x2.8° (T42L80)	23 pressure levels (17 std + 6 ext)	Watanabe et al. [2011]
14	Max Planck Institute for Meteorology	MPI-M	MPI-ESM-LR	ESM	1.8°x1.8° (T63L47)	25 pressure levels (17 std + 8 ext)	Raddatz et al. [2007]
15	Meteorological Research Institute	MRI	MRI-CGCM3	AOGCM	1.1°x1.1° (T159L48)	23 pressure levels (17 std + 6 ext)	Yukimoto et al. [2012]
16	Norwegian Climate Centre	NCC	NorESM1-M	ESM	2.5°x1.9° (f19L26)	17 standard pressure levels	Zhang et al. [2012]



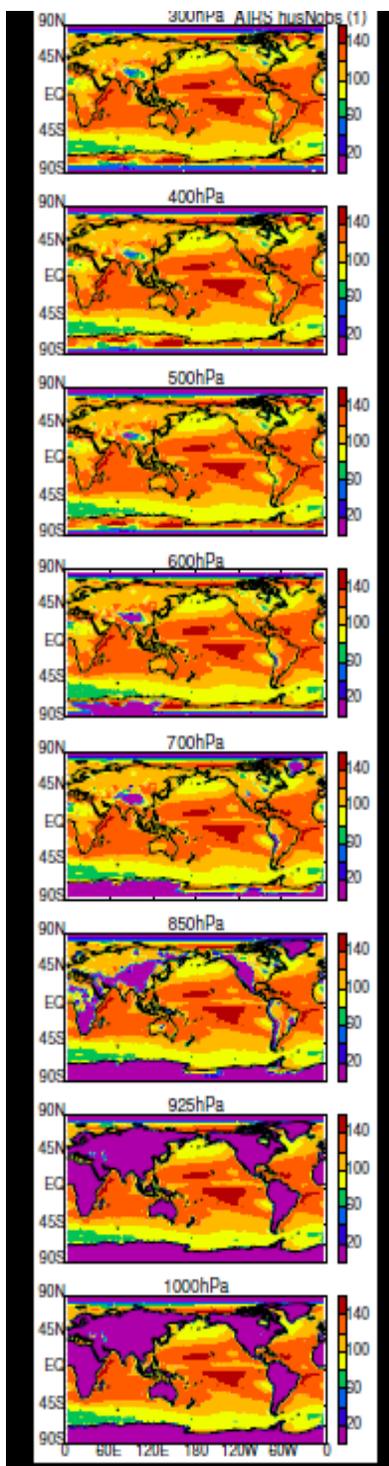


Figure 1: Climatological mean AIRS tropospheric air temperature (t_a , left panel) and specific humidity (h_{us} , right panel) number of observations (Nobs) per 1° by 1° grid cell at each pressure level.

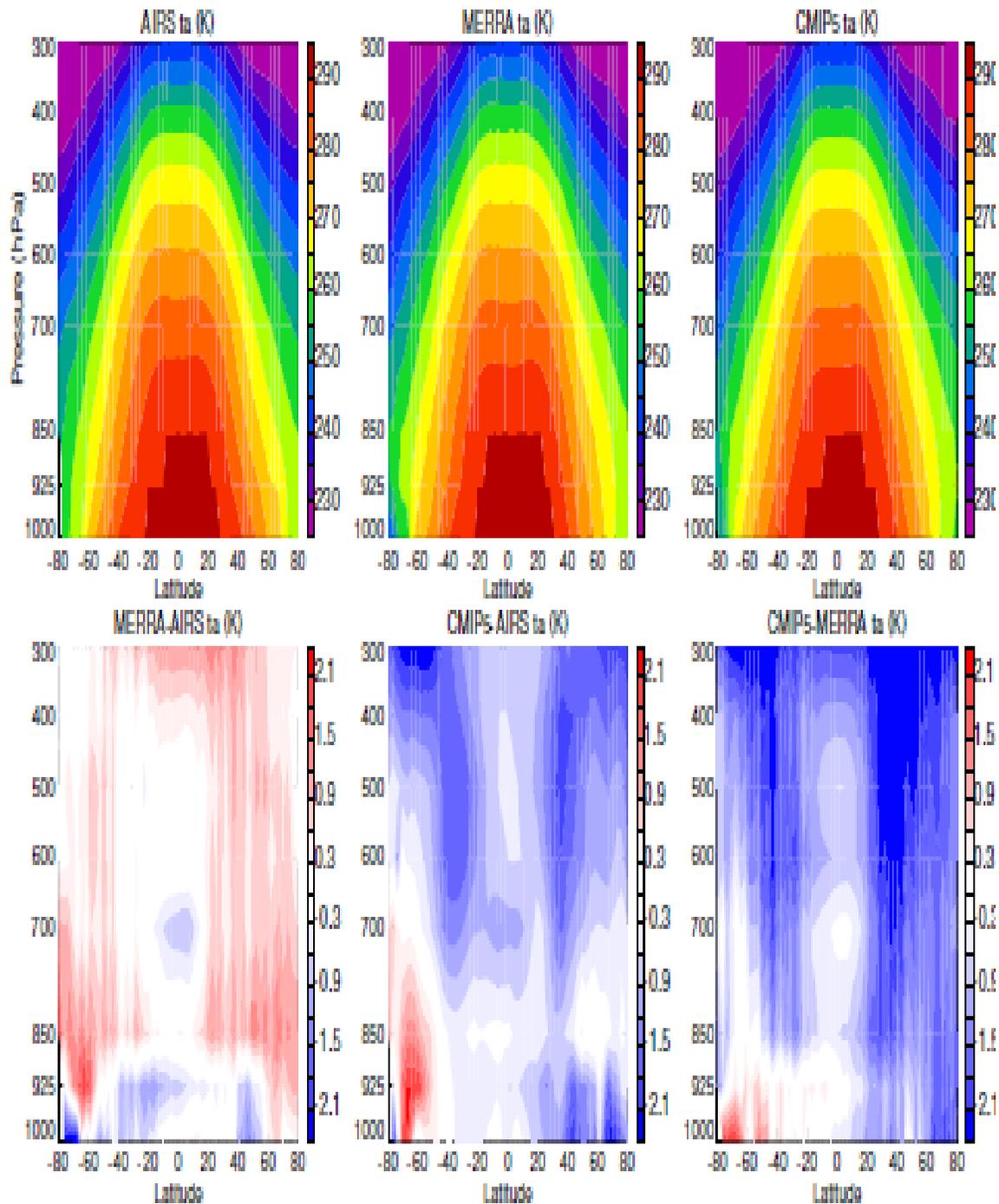


Figure 2: Global zonal mean cross sections of climatological mean tropospheric air temperature (K, upper panels) (column 1: AIRS; column 2: MERRA; column 3: CMIP5 MMEM) and their differences (K, lower panels) (column 1: MERRA-AIRS; column 2: CMIP5-AIRS; column 3: CMIP5-MERRA).

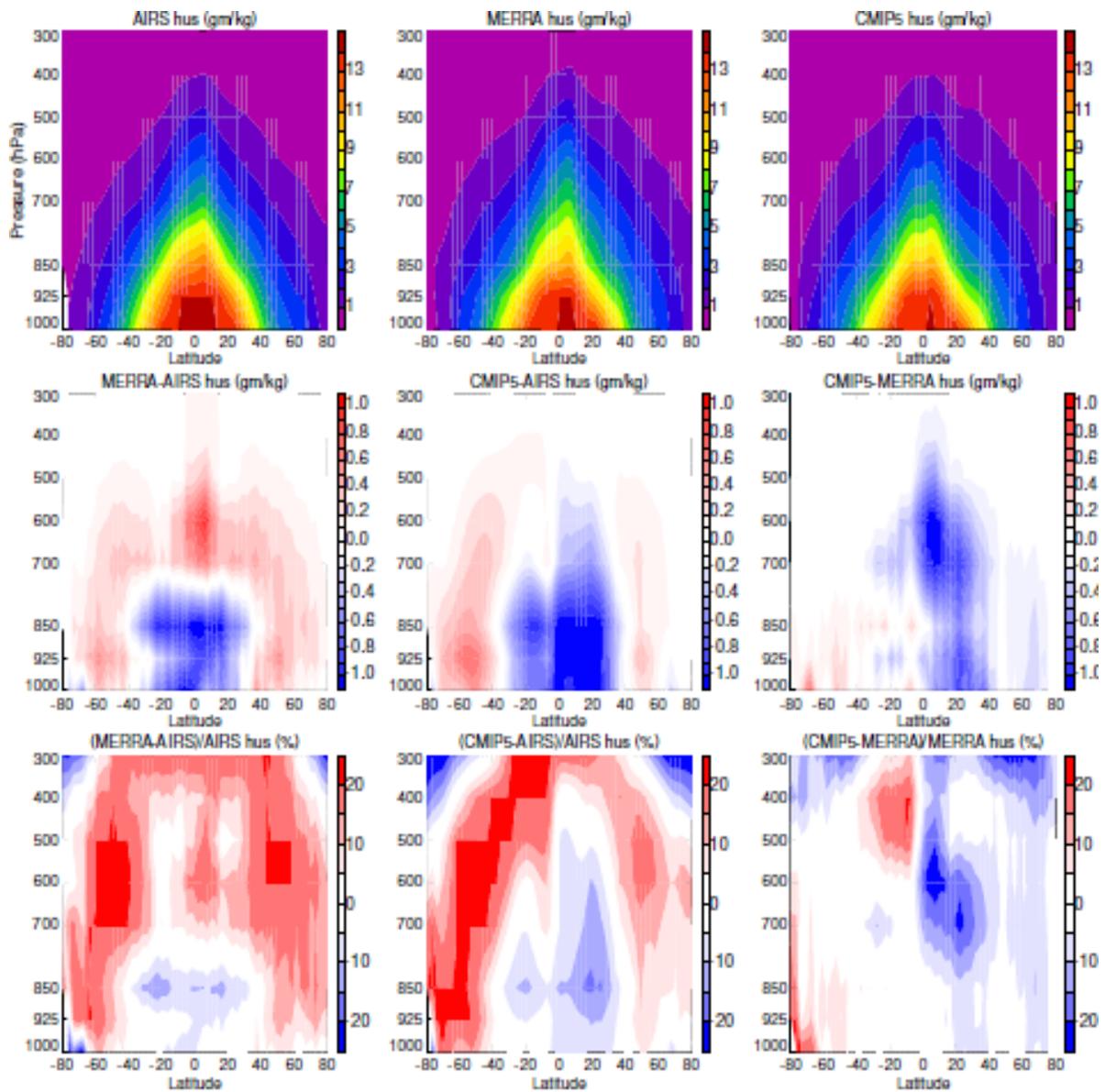


Figure 3: Global zonal mean cross sections of climatological mean tropospheric specific humidity (g/kg, upper panels) (column 1: AIRS; column 2: MERRA; column 3: CMIP5 MEM), their differences (g/kg, middle panels) (column 1: MERRA–AIRS; column 2: CMIP5–AIRS; column 3: CMIP5–MERRA), and their relative differences (% , lower panels) (column 1: (MERRA–AIRS)/AIRS; column 2: (CMIP5–AIRS)/AIRS; column 3: (CMIP5–MERRA)/MERRA).

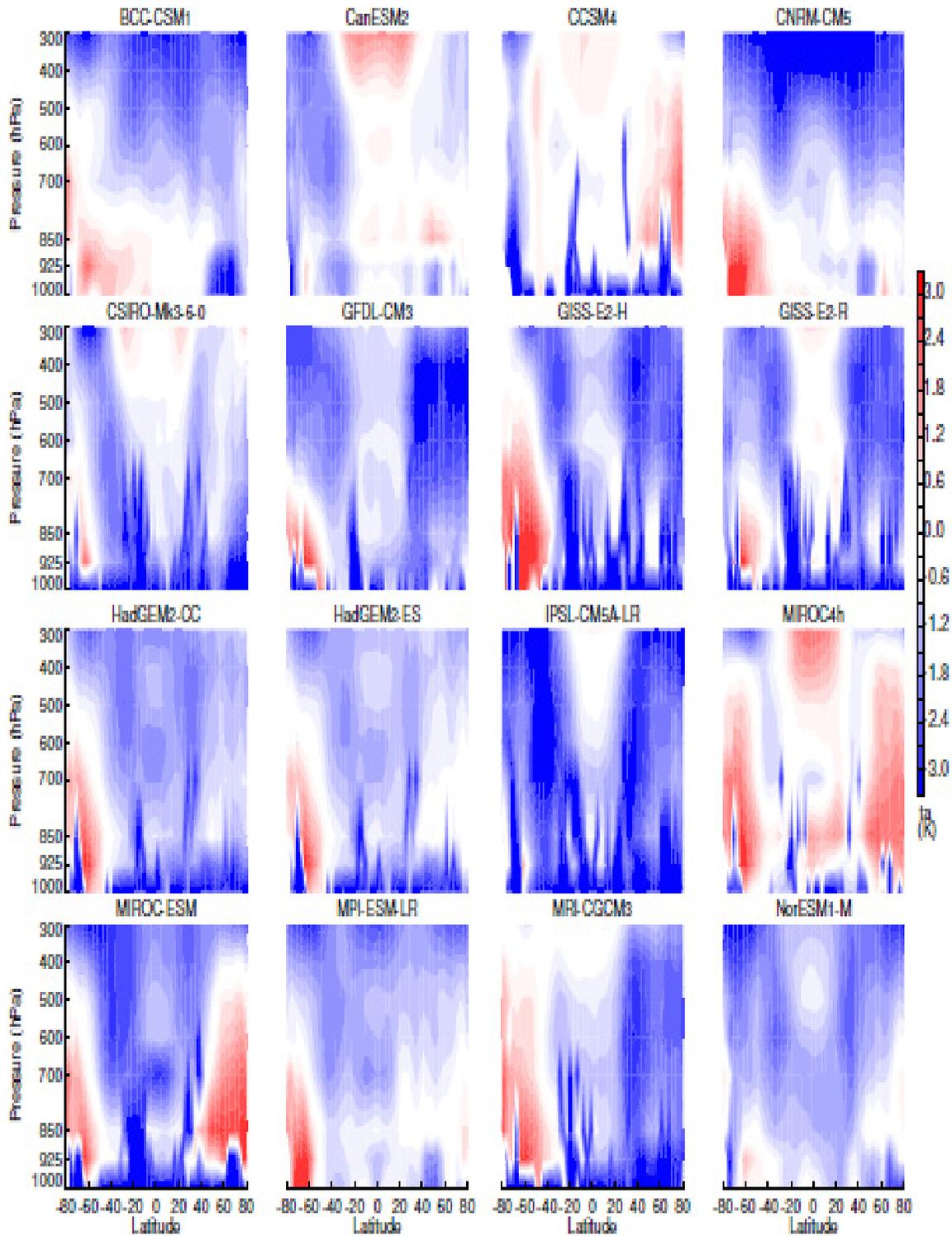


Figure 4: Global zonal mean cross sections of climatological mean tropospheric air temperature biases relative to AIRS for sixteen individual CMIP5 models.

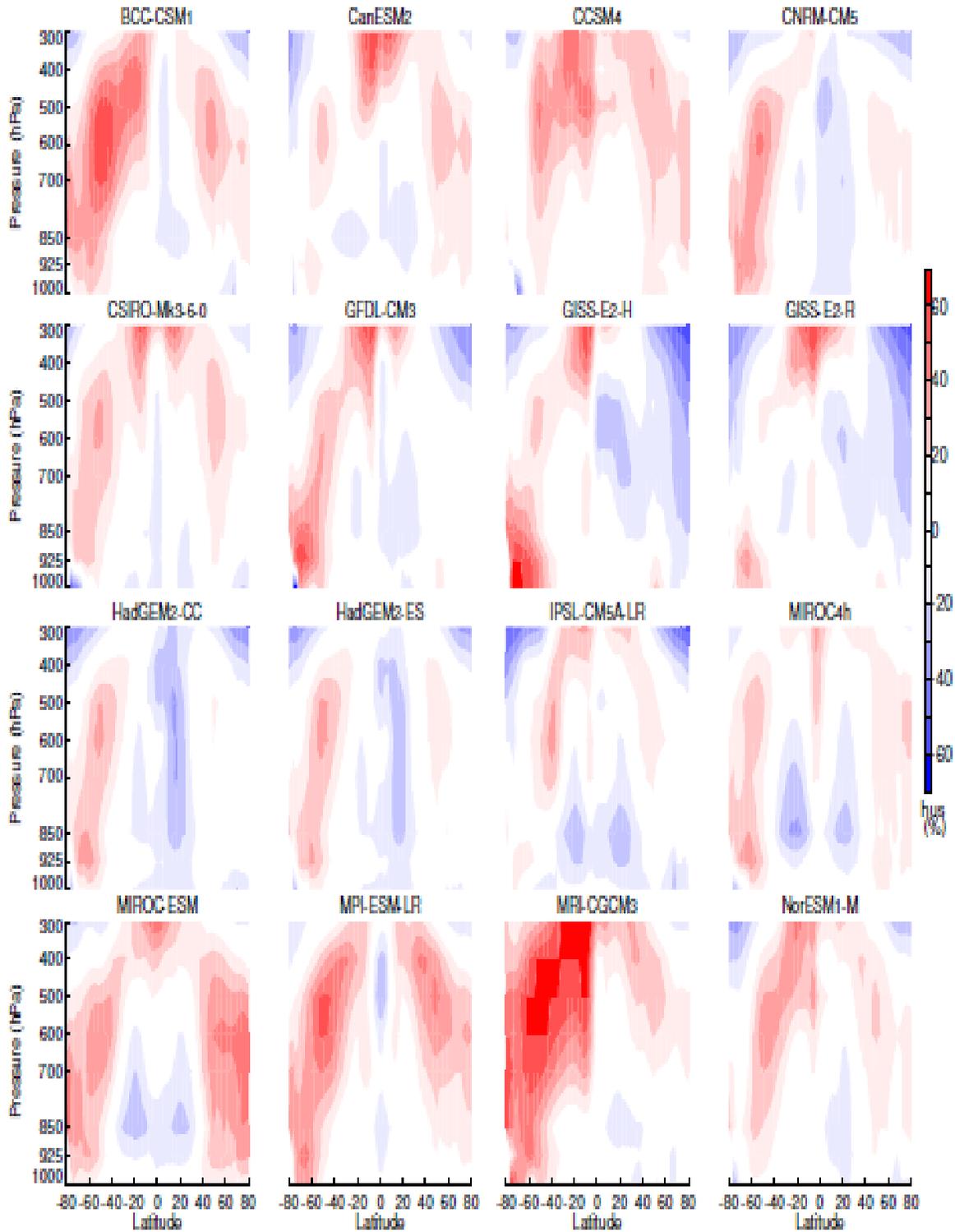
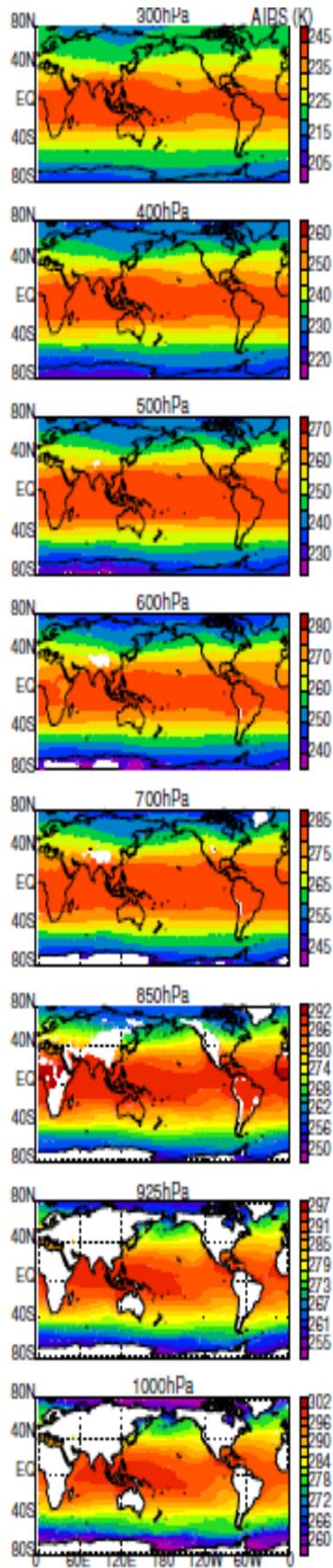
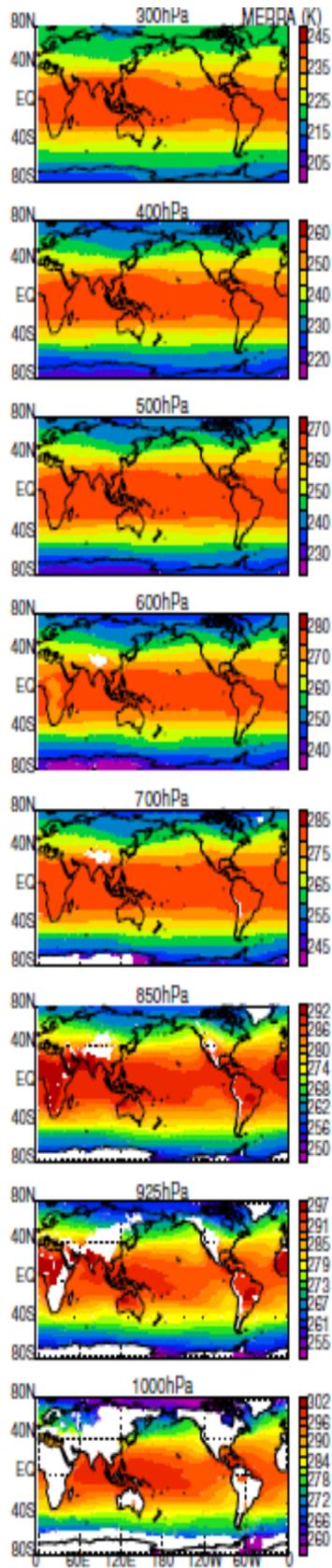


Figure 5: Global zonal mean cross sections of climatological mean tropospheric specific humidity biases relative to AIRS for sixteen individual CMIP5 models.





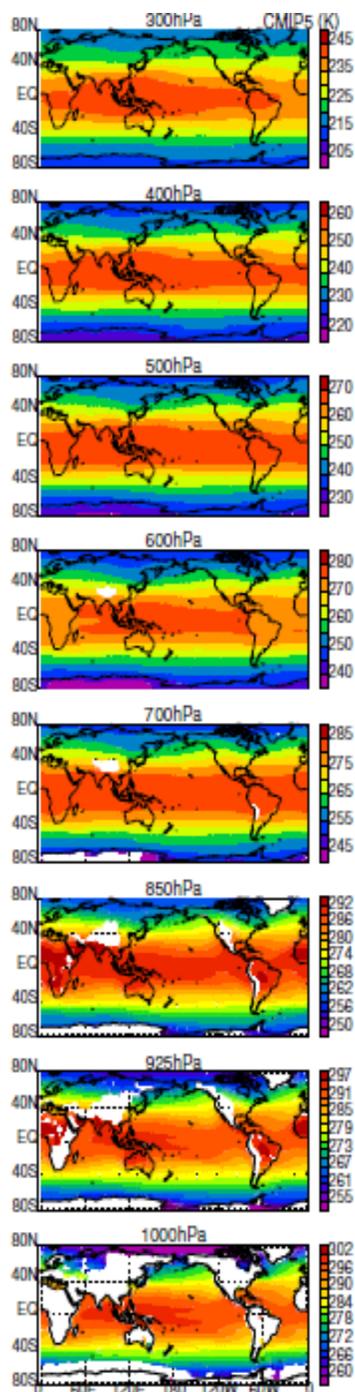
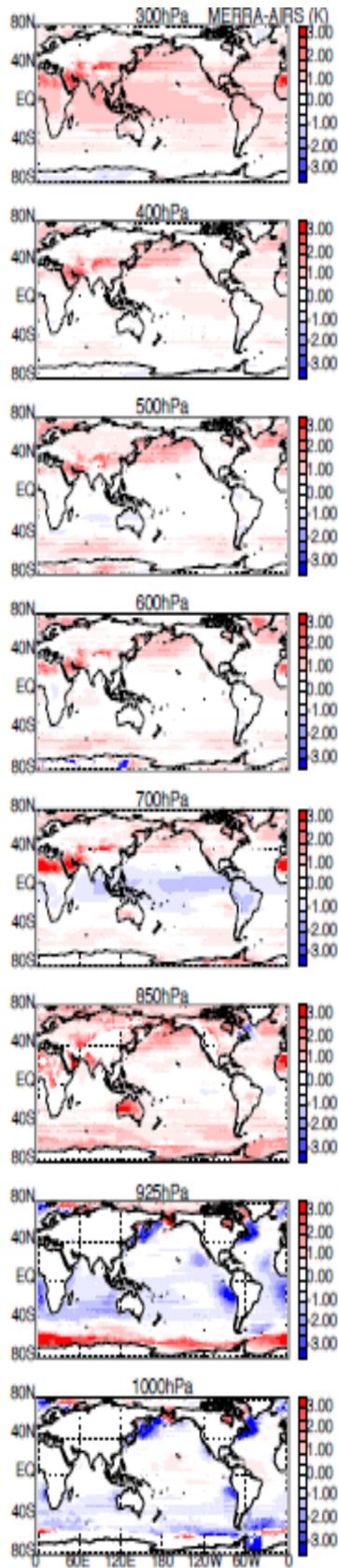
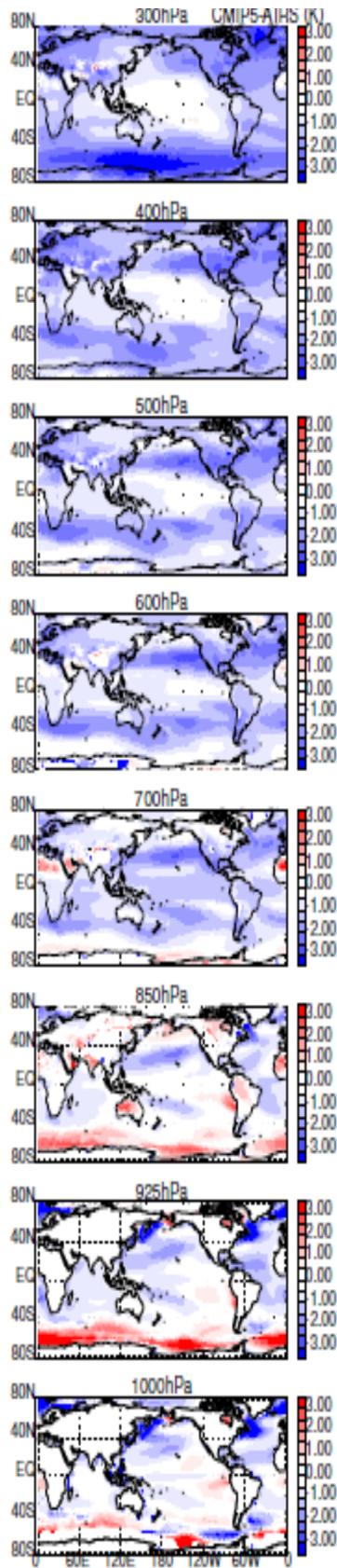


Figure 6: Global spatial maps of climatological mean tropospheric air temperature (K) at different pressure levels (column 1: AIRS; column 2: MERRA; column 3: CMIP5 MMEM). Note that different scales are used for different pressure levels.





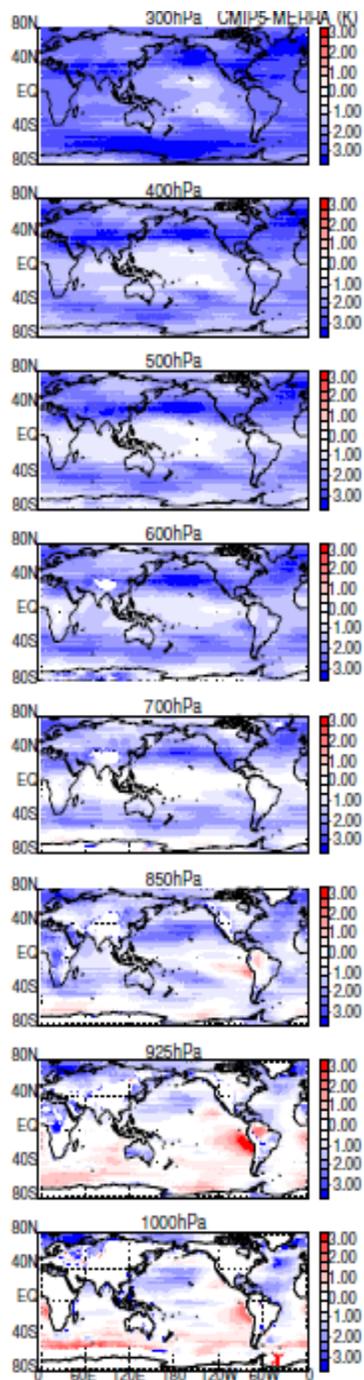
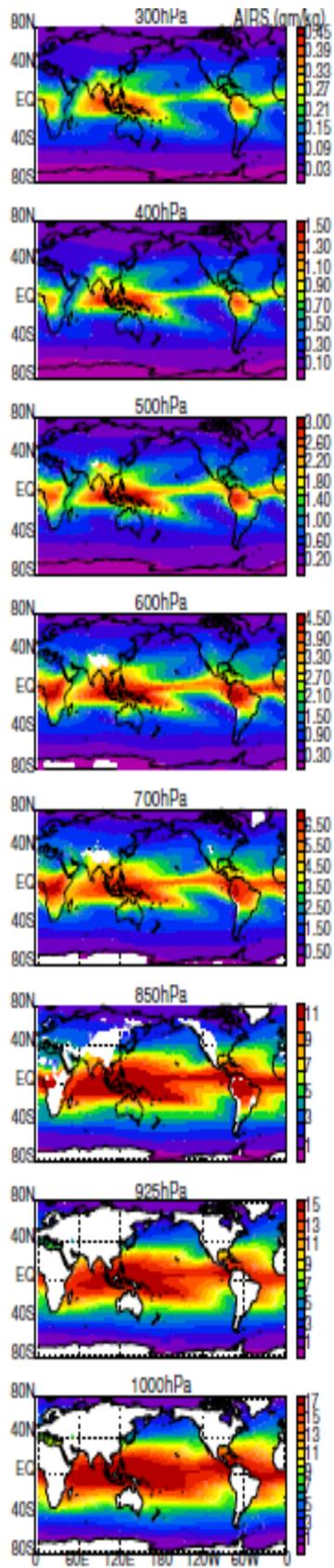
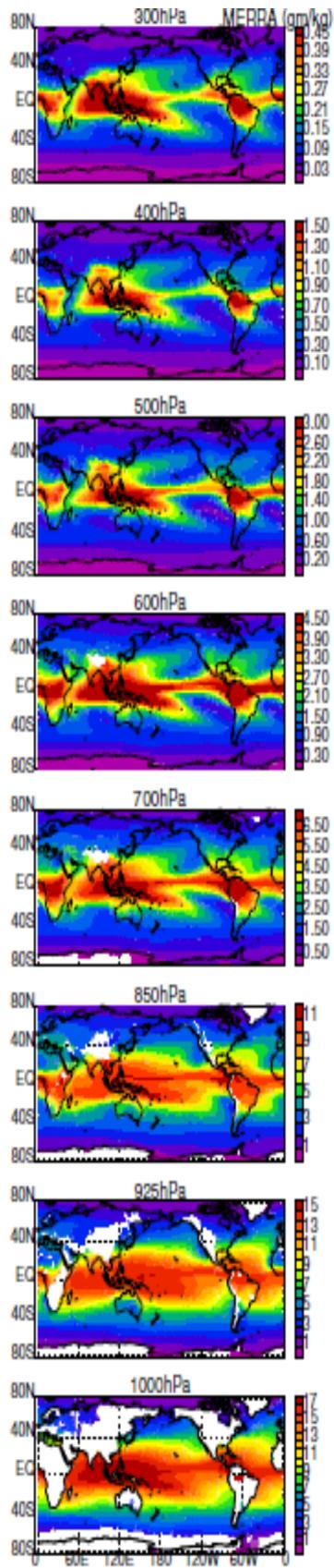


Figure 7: Global spatial maps of climatological mean tropospheric air temperature differences (K) among AIRS, MERRA, and CMIP5 at different pressure levels (column 1: MERRA–AIRS; column 2: CMIP5–AIRS; column 3: CMIP5–MERRA).





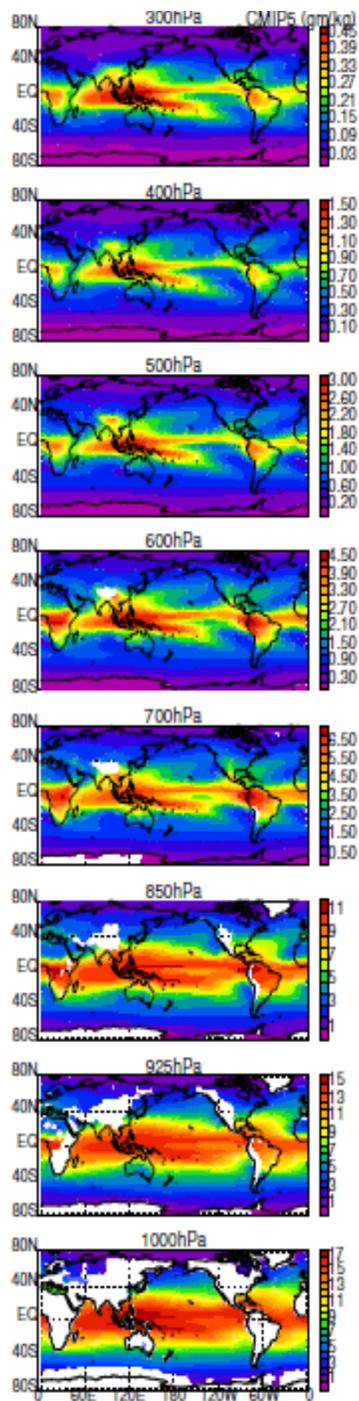
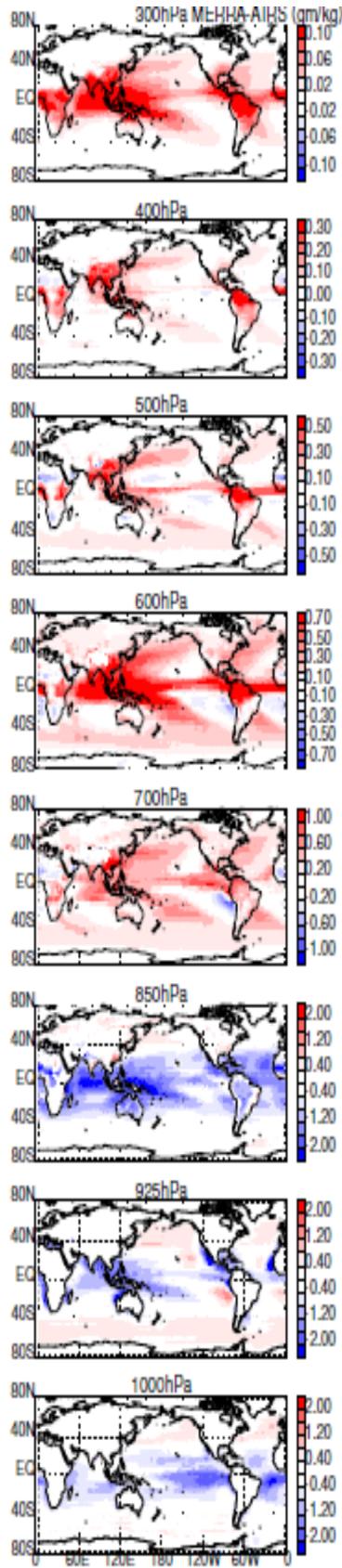
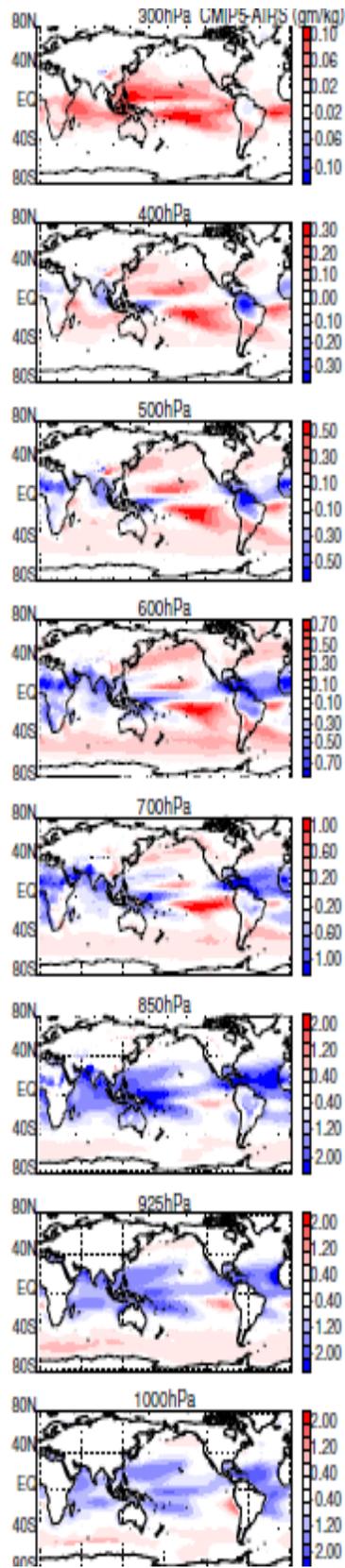


Figure 8: Global spatial maps of climatological mean tropospheric specific humidity (g/kg) at different pressure levels (column 1: AIRS; column 2: MERRA; column 3: CMIP5 MMEM). Note that different scales are used for different pressure levels.





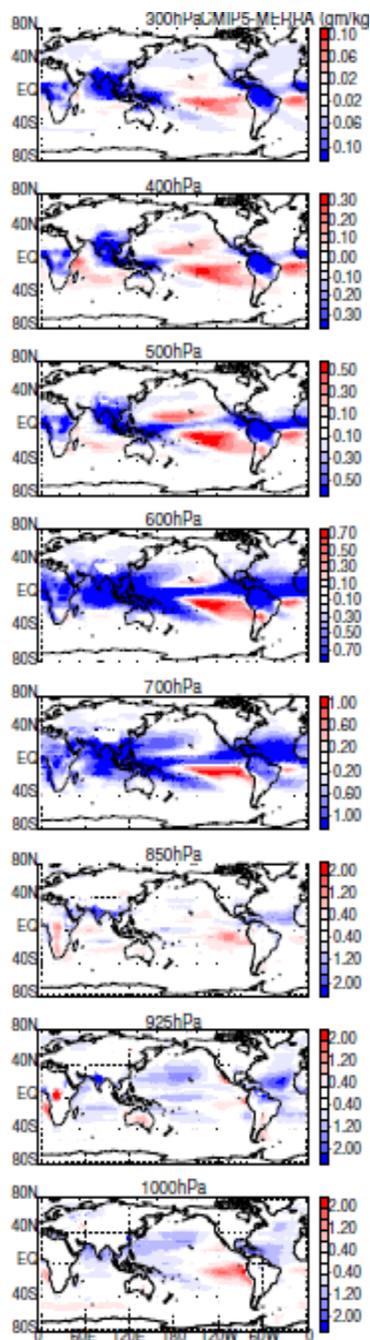
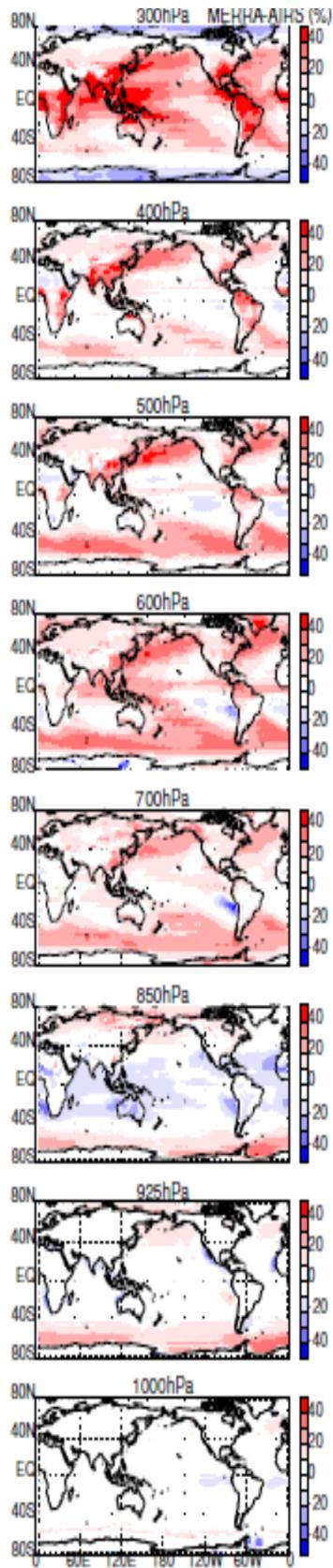
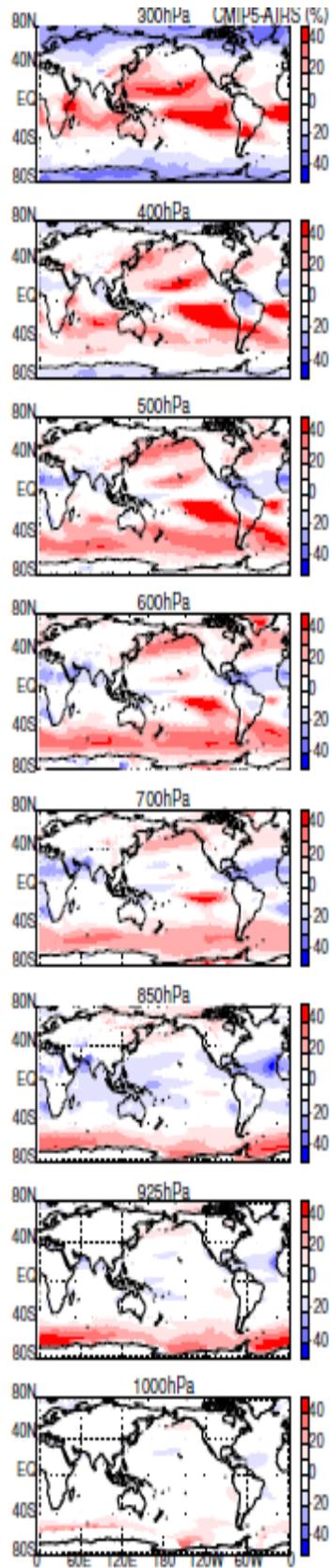


Figure 9: Global spatial maps of climatological mean tropospheric specific humidity differences (g/kg) among AIRS, MERRA, and CMIP5 at different pressure levels (column 1: MERRA–AIRS; column 2: CMIP5–AIRS; column 3: CMIP5–MERRA). Note that different scales are used for different pressure levels.





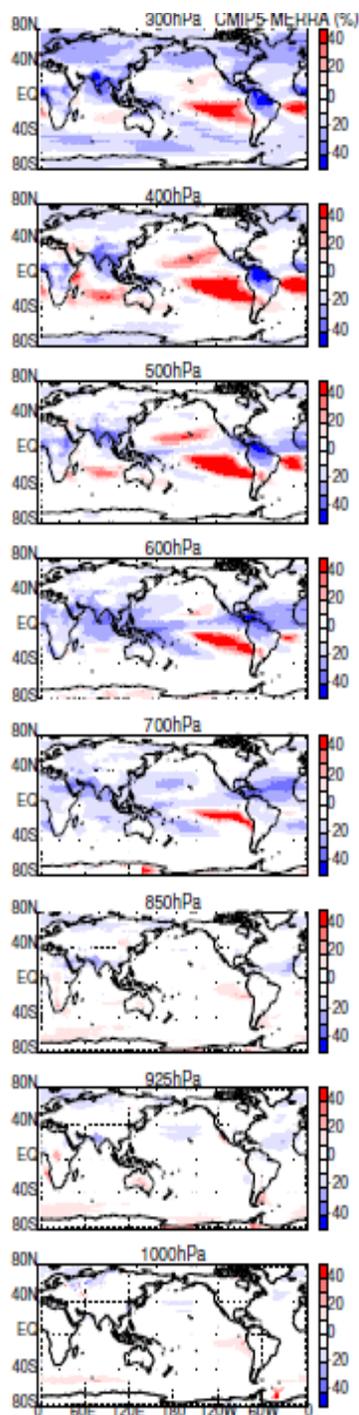


Figure 10: Global spatial maps of climatological mean tropospheric specific humidity relative differences (%) among AIRS, MERRA, and CMIP5 at different pressure levels (column 1: $-(\text{MERRA}-\text{AIRS})/\text{AIRS}$; column 2: $(\text{CMIP5}-\text{AIRS})/\text{AIRS}$; column 3: $(\text{CMIP5}-\text{MERRA})/\text{MERRA}$).

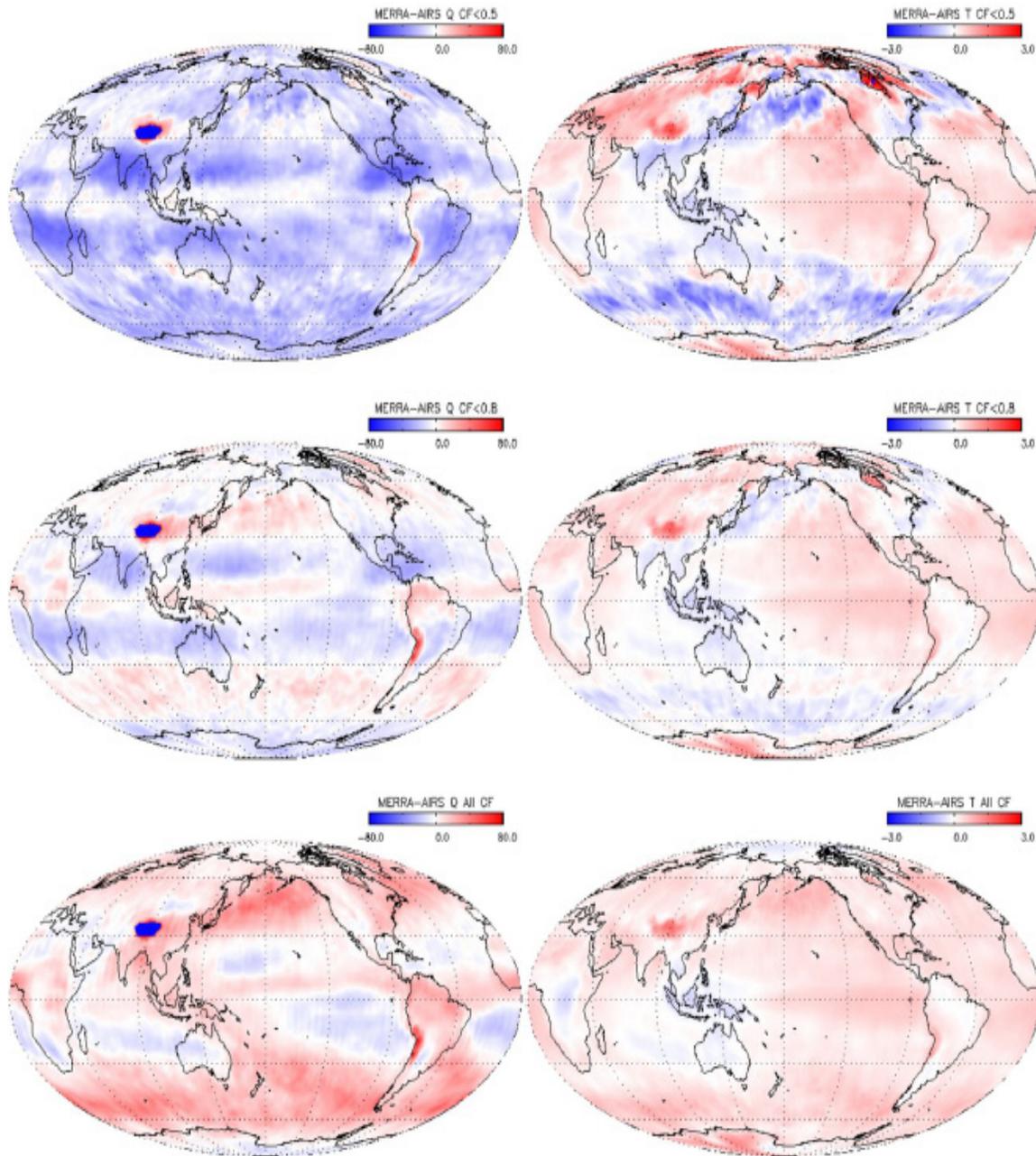


Figure 11: Global spatial maps of one-year (1 September 2008 – 31 August 2009) mean MERRA-AIRS specific humidity (left panels) and air temperature (right panels) differences (MERRA-AIRS) at 500 hPa for MERRA subsampled for different cloud amounts (top panels: MERRA cloud fraction (CF)<0.5; middle panels: CF<0.8; bottom panels: CF \leq 1.0).

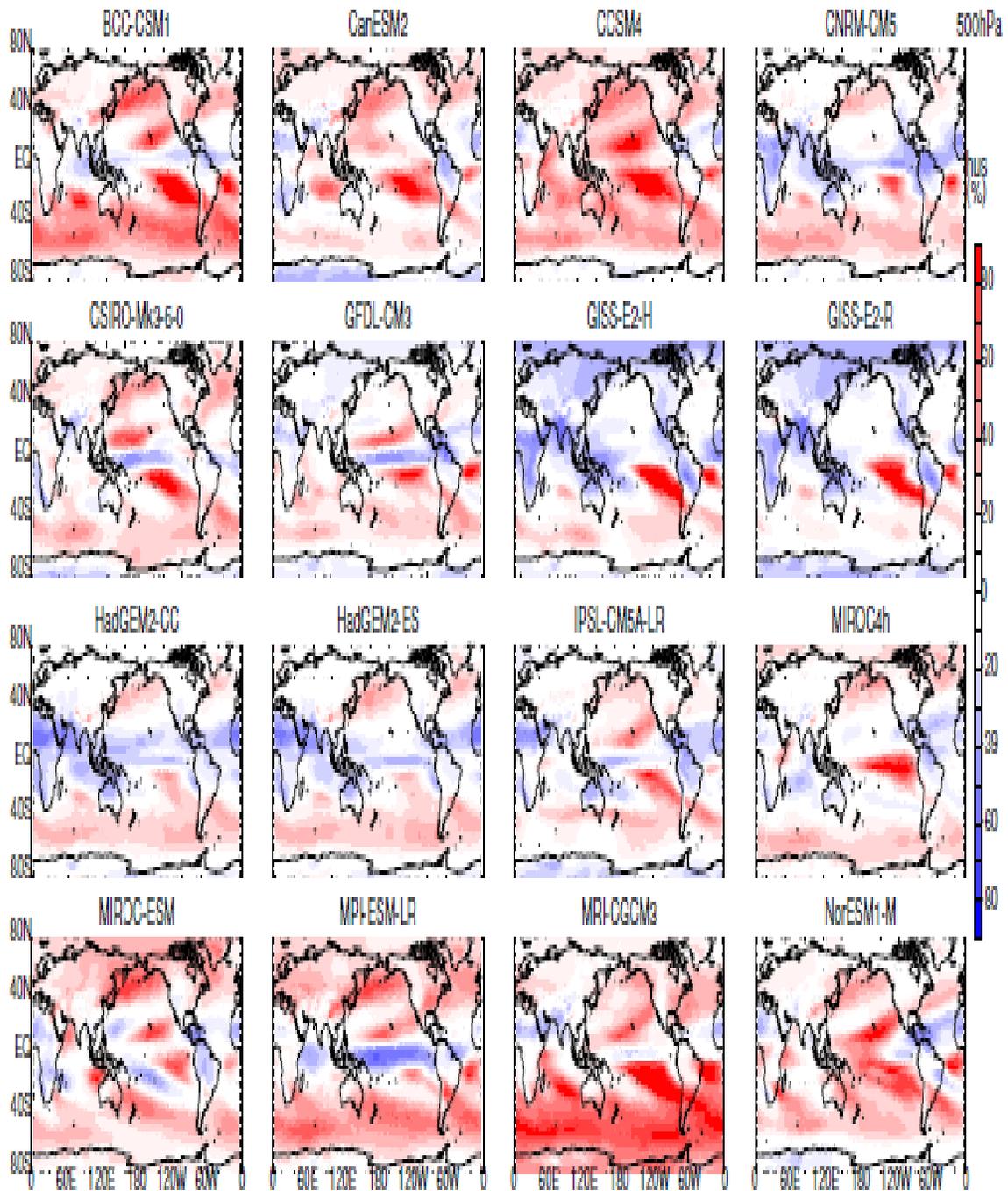


Figure 12: Global spatial maps of climatological mean tropospheric specific humidity relative biases (%) relative to AIRS for sixteen individual CMIP5 models at 500 hPa.