An assessment of the consistency between satellite measurements of upper tropospheric water vapor

Eui-Seok Chung[1], Brian J. Soden[1], Xianglei Huang[2], Lei Shi[3], and Viju O. John[4]

[1]*Rosenstiel School of Marine and Atmospheric Science*
University of Miami, Miami, Florida

[2]*Department of Climate and Space Sciences and Engineering*
University of Michigan, Ann Arbor, Michigan

[3]*National Centers for Environmental Information*
NOAA, Asheville, North Carolina

[4]*Met Office Hadley Centre*
Exeter, United Kingdom

Submitted to *Journal of Geophysical Research-Atmospheres*
February 2016 (Revised Version)

Corresponding author address: Eui-Seok Chung, Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Causeway, Miami, FL 33149, USA. (echung@rsmas.miami.edu)

Key points

1. The consistency of the satellite-based observations of upper tropospheric water vapor is assessed

2. The consistency signifies the credibility of satellite-based monitoring on multi-decadal time scales

3. Both reanalyses and GCM simulations fail to capture satellite-observed long-term variability

Abstract

We assess the consistency of the satellite-based observations of upper tropospheric water vapor (UTWV) by comparing brightness temperature measurements from the channel 12 of High-Resolution Infrared Radiation Sounder (HIRS), the 183.31±1 GHz channel of Advanced Microwave Sounding Unit-B (AMSU-B)/Microwave Humidity Sounder (MHS), and spectral radiances from the Atmospheric Infrared Sounder (AIRS). All three products exhibit consistent spatial and temporal patterns of interannual variability. On decadal time-scales, the spatial patterns of trends are similar between all three products, however the amplitude of the regional trends is noticeably weaker in the HIRS measurements than in either the AMSU-B/MHS or AIRS data. This presumably reflects the greater clear-sky sampling limitations of HIRS relative to the other products. However, when averaged over tropical or near-global spatial scales, the trends between all three products are statistically indistinguishable from each other. The overall consistency between all three products provides important verification of their credibility for documenting long-term changes in UTWV. A similar analysis is performed for reanalysis-produced and model-simulated UTWV using the HIRS record as a benchmark. On decadal time-scales, both reanalysis data sets and the multi-model ensemble mean have difficulty in capturing the observed moistening of climatologically dry regions of the subtropics, although the model-simulated trends are more consistent with the HIRS measurements than are the reanalysis data.

1. Introduction

Water vapor is one of the most important components in the climate system. Not only is water vapor the ultimate source of clouds and precipitation, but also it plays a crucial role in energy transport through its phase changes. Moreover, the fact that water vapor is the strongest greenhouse gas in the atmosphere implies that changes in the amounts and distribution of water vapor may significantly amplify the extent of global warming induced by increases in anthropogenic greenhouse gases [e.g., *Held and Soden*, 2000; *Wentz and Schabel*, 2000; *Soden and Held*, 2006; *Mears* et al., 2007; *Dessler* et al., 2008; *Gettelman and Fu*, 2008; *Allan*, 2012].

Although water vapor in the upper troposphere accounts for a very small fraction of the column integrated water vapor, the absorption of outgoing terrestrial longwave radiation by water vapor increases logarithmically with the concentration of water vapor. This means that the strength of water vapor feedback is determined not by the absolute change, but by the fractional change in water vapor amount. Under a constant relative humidity scenario, the largest fractional changes in water vapor occur in the upper troposphere [e.g., *Held and Soden*, 2000; *Soden and Held*, 2006; *Sherwood* et al., 2010; *Allan*, 2012].

The distribution and variability of upper tropospheric water vapor (UTWV) are governed by the interplay among large-scale circulation, deep convection and lateral mixing by transient eddies [e.g., *Soden and Bretherton*, 1994; *Salathé and Hartmann*, 1997; *Pierrehumbert and Roca*, 1998; *Brogniez and Pierrehumbert*, 2006; *Sohn* et al., 2008; *Sherwood* et al., 2010; *Chung* et al., 2011]. Thus accurate and stable observations of UTWV are crucial for monitoring its changes and assessing the credibility of model projections. Although reanalysis assimilates both in situ

and satellite observations, substantial discrepancies exist between different reanalysis-produced UTWV data sets [e.g., *Dessler and Davis*, 2010].

Early studies of the global distribution and variability of UTWV used infrared satellite observations of the brightness temperatures (BTs) in the 6.3-μm water vapor absorption band [e.g., *Schmetz and Turpeinen*, 1988; *Soden and Bretherton*, 1993, 1994; *Bates and Jackson*, 2001; *Brogniez* et al., 2006]. In particular, BTs from the channel 12 of the High-Resolution Infrared Radiation Sounder (HIRS) onboard polar-orbiting satellites have been utilized to infer decadal-scale variability over the globe, to evaluate climate models, and to confirm the presence of strongly positive water vapor feedback in the climate system [e.g., *Bates and Jackson*, 2001; *Iacono* et al., 2003; *Soden* et al., 2005; *Shi and Bates*, 2011; *Chung* et al., 2014]. More recently, hyperspectral atmospheric sounders such as Atmospheric Infrared Sounder (AIRS), Infrared Atmospheric Sounding Interferometer (IASI), and the Cross-track Infrared Sounder (CrIS) have allowed retrieval of water vapor profiles with even higher vertical resolutions.

Microwave satellite observations of BTs in the 183 GHz water vapor absorption band have also been utilized to monitor changes in UTWV [e.g., *Engelen and Stephens*, 1999; *Sohn* et al., 2000; *Buehler* et al., 2008]. While a substantial amount of effort has been made to construct intersatellite-calibrated data sets suitable for long-term climate monitoring, a comprehensive assessment on their reliability has not been conducted. In this study, we assess the consistency and robustness of climatological variations in satellite-observed UTWV by comparing BT measurements from these three independent satellite data sets (i.e., HIRS channel 12, microwave 183.31±1 GHz channel, and AIRS). Because reanalysis data sets and climate model simulations

are frequently used to scrutinize and project the variability of key climate variables, we also analyze the patterns of UTWV variability from reanalysis data sets and climate model simulations using the HIRS record as a benchmark in order to examine whether reanalysis-produced and model-simulated UTWV data sets are suitable for analyzing UTWV changes on multi-decadal time scales.

2. Data Sets

The longest global archives of UTWV are provided by observations from the channel 12 of HIRS. The HIRS, a cross-track scanner with 20 spectral channels, has flown onboard the operational National Oceanic and Atmospheric Administration (NOAA) polar-orbiting satellites since 1978. The central wavelength of this channel is 6.7 μm for the HIRS version 2 (HIRS/2) onboard TIROS-N and NOAA-6 through NOAA-14 satellites. The noise equivalent differential radiance is 0.19 mW $(m^2$ sr $cm^{-1})^{-1}$ for this channel (https://eosweb.larc.nasa.gov/ACEDOCS/data/appen.c.1.html). Intersatellite-calibrated data sets have been constructed by correcting for intersatellite biases due to differences in instrument calibration and other error sources [*Bates* et al., 1996, 2001; *Bates and Jackson*, 2001; *Jackson and Soden*, 2007], in which decrease in BT with increasing viewing angle is adjusted by applying a multi-channel linear multivariate regression algorithm [*Jackson* et al., 2003]. Unfortunately, the central wavelength was shifted from 6.7 μm to 6.5 μm with the launch of HIRS/3 in 1998. This shift resulted in a BT difference of ~8 K relative to the HIRS/3 (NOAA-15, 16, and 17) and HIRS/4 (NOAA-18 and 19) instruments. To correct for this bias, Shi and Bates

[2011] quantified scene-temperature-dependent intersatellite biases by comparing the zonal averages between overlapping satellites. In this study, the intersatellite-calibrated HIRS channel 12 BTs from Shi and Bates [2011] are analyzed for the period 1979-2014.

Microwave observations in the 183.31±1 GHz channel are also sensitive to layer-averaged relative humidity in the upper troposphere similar to the infrared UTWV channels [e.g., *Sohn* et al., 2000, 2003]. The Advanced Microwave Sounding Unit-B (AMSU-B) and the Microwave Humidity Sounder (MHS) are carried on polar-orbiting satellites (NOAA and the European Meteorological Operational satellite programme (MetOp)) and have nearly identical 183.31±1 GHz channels. The nominal noise equivalent temperatures are 1.06 K for AMSU-B and 0.50 K for MHS [*John* et al., 2012]. Because microwave measurements are less affected by clouds compared to infrared measurements, they are less prone to clear-sky sampling biases and have improved space/time coverage [e.g., *Sohn and Schmetz*, 2004; *John* et al., 2011]. Recently, Chung et al. [2013b] produced an intersatellite-calibrated data set of 183.31±1 GHz measurements by adjusting for calibration and diurnal sampling biases between the AMSU-B or MHS instruments. They removed the viewing angle dependence of BT using a method based upon a simplified model of radiative transfer. This intersatellite-calibrated microwave data set covers the period 1999-2014.

Spectrally-resolved radiances observed from the AIRS onboard the National Aeronautics and Satellite Administration (NASA) Aqua satellite provide a third independent data set for UTWV. The AIRS consists of 2378 spectral channels in the thermal infrared band (3.7-15.4 μm) together with 4 channels in the visible/near-infrared band, which enables to retrieve the vertical

profiles of temperature and water vapor with significantly enhanced accuracy and vertical resolution compared to conventional atmospheric sounders. Based on the initial in-flight calibrations [*Pagano* et al., 2003; *Gaiser* et al., 2003], AIRS measurements have a radiometric accuracy ~0.3 K or higher for a target with 250 K BT and its spectral accuracy is better than 0.5% of the full width at half maximum of each channel. Aumann et al. [2006] further updated the calibration estimation with the calibration accuracy better than 0.2 K and the stability be better than 16 mK per year. To facilitate the comparison with HIRS observations, 16-day averages of AIRS nadir-view quality-controlled clear-sky radiances from 2003 to 2014 are used here [*Huang and Yung*, 2005; *Pan* et al., 2015]. 16 days is the repeating cycle of the Aqua satellite that hosts AIRS. Thus, 16-day averages can expect to achieve uniform sampling. The details of such AIRS data process can be found in Huang and Yung [2005] and Fig. 1 in Pan et al. [2015] illustrates the sampling pattern of the AIRS data used in such averaging process. The AIRS radiances (R_{AIRS}) are then convolved with the spectral response function (SRF) of HIRS channel 12:

$$R_{12} = \frac{\int_{\nu 1}^{\nu 2} R_{AIRS}(\nu) SRF(\nu) d\nu}{\int_{\nu 1}^{\nu 2} SRF(\nu) d\nu},$$

where ν denotes wavenumber or wavelength. Then, the AIRS-derived HIRS channel 12 radiances (R_{12}) are converted into BTs that cover the period of 2003 to 2014.

In addition to these three homogenized long-term satellite products, the patterns of interannual and decadal variability of UTWV are analyzed for reanalysis data sets and climate model simulations. The reanalysis data sets analyzed in this study are the European Centre for Medium-Range Weather Forecasts (ECMWF) Interim Reanalysis [ERA-Interim, *Dee* et al., 2011]

for the period 1979-2014, the NASA Modern-Era Retrospective Analysis for Research and Applications [MERRA, *Rienecker* et al., 2011] for the period 1979-2014, and the Japanese 55-year Reanalysis [JRA-55, *Kobayashi* et al., 2015] for the period 1979-2013. For climate model simulations, we analyze output for the Atmospheric Model Intercomparison Project (AMIP) simulations for the period 1979-2008 from 18 climate models participating in the phase 5 of the Coupled Model Intercomparison Project (CMIP5) [*Taylor* et al., 2012]: ACCESS1-0, bcc-csm1-1, BNU-ESM, CanAM4, CCSM4, CNRM-CM5, CSIRO-Mk3-6-0, GFDL-CM3, GISS-E2-R, HadGEM2-ES, inmcm4, IPSL-CM5A-LR, IPSL-CM5A-MR, MIROC5, MPI-ESM-LR, MPI-ESM-MR, MRI-CGCM3, and NorESM1-M. To be consistent with the satellite observations, reanalysis-produced and model-simulated atmospheric profiles of temperature and specific humidity are inserted into a fast radiative transfer model (RTTOV version 10 [*Hocking* et al., 2011]) to simulate HIRS channel 12 BTs that would be observed by satellites for given atmospheric conditions. The ability of RTTOV to simulate radiances in the 6.3-μm water vapor absorption band is comparable to other fast radiative transfer models [e.g., *Saunders* et al., 2007].

Variability in BT for these water vapor channels is due to both changes in temperature and relative humidity in the upper troposphere. As demonstrated in Soden and Bretherton [1993], however, relative humidity exerts substantially greater influences on BT variability in comparison to temperature, indicating that a decrease (increase) in both satellite-observed and synthesized BT can be translated into an increase (decrease) in relative humidity in the upper troposphere.

3. Results

In this section, the consistency of the satellite-based observations of UTWV is assessed by analyzing the short-term and long-term variability of BTs among the three independent data sets. The simulated HIRS channel 12 BTs from reanalysis data sets and output from climate model simulations are also compared to the satellite observations in order to evaluate the interannual and decadal variability of reanalysis-produced and model-simulated UTWV. The BTs in these channels are proportional to the logarithm of the relative humidity averaged over a deep layer of the upper troposphere, which extends from roughly 200-500 hPa [*Soden and Bretherton*, 1993]. The altitude of this layer shifts depending upon the integrated water path, being higher for humid atmospheres and lower for dry atmospheres. In polar regions, where the atmosphere is extremely dry, the radiances can even receive significant contributions from the surface emission. For the microwave channel, this complication is further amplified because of substantial spatiotemporal variability in the surface emissivity. As a result, our analysis is confined to the region 60^{o}S-60^{o}N.

3.1 Short-term Variability

The red lines of Fig. 1 denote the time series of domain-averaged monthly-mean BT anomalies over two regions (60^{o}S-60^{o}N and 30^{o}S-30^{o}N) for HIRS channel 12. Anomalies are computed by removing the mean seasonal cycle over the period 2005-2014. The time series exhibit pronounced high-frequency variability over the period 1979-2014, but do not show distinct trends. The transition of HIRS/2 to HIRS/3, accompanied by pronounced shift in the central wavelength from 6.7 μm (HIRS/2) to 6.5 μm (HIRS/3), resulted in a nominal BT

difference of ~8 K between them, leading to the end of the HIRS/2 record with the decommission of NOAA-14 in the mid-2000s. However this discontinuity has been almost completely removed by a series of calibration processes conducted by Shi and Bates [2011], in which the influence of the filter change was adequately taken into account by a scene radiance dependent bias correction.

Figure 1 also presents the time series of BT anomalies for AMSU-B/MHS 183.31±1 GHz channel over the period 1999-2014 (blue) and AIRS-derived HIRS channel 12 over the period 2003-2014 (green), with the insets showing difference time series. Both of the anomaly time series show pronounced monthly to interannual variability but negligible trends. The correlation coefficient of the anomaly time series between HIRS channel 12 and AMSU-B/MHS 183.31±1 GHz channel is 0.89 (0.91) over the 60ºS-60ºN (30ºS-30ºN) domain for the period 1999-2014, underscoring the consistency of their high frequency variability. The HIRS channel 12 and AIRS are also highly correlated with the correlation coefficient of 0.83 over both of the domains for the period 2003-2014. These high correlations (statistically significant above 95% level) between independent observations demonstrate the credibility of the satellite-based observations of UTWV.

Figure 2 presents the time series of simulated HIRS channel 12 BT anomalies from ERA-Interim (purple), MERRA (blue), JRA-55 (green), and the multi-model ensemble mean of climate model simulations for CMIP5 AMIP experiment (orange) over the two domains. The BT anomalies are computed by removing the mean seasonal cycle over the period 1979-2008. For clarity, monthly time series are smoothed by applying 13-month running average. Despite the

fact that a number of in situ and satellite observations have been assimilated into these reanalysis systems, the dynamic ranges of the simulated BT anomalies are noticeably different from the HIRS record (black line). While the observed BT anomalies range from -0.3 K to 0.2 K over the 60ºS-60ºN domain, the range of simulated BT anomalies is from -0.2 K to 0.3 K for ERA-Interim; -0.2 K to 0.4 K for MERRA; -0.5 K to 0.3 K for JRA-55. The BT anomalies over the 30ºS-30ºN domain also show similar discrepancies in the dynamic range between the observation and reanalyses, which leads to a speculation that assimilated satellite and/or in situ observations might have biases, in particular, for JRA-55. The correlation coefficients with the HIRS record are 0.73 for ERA-Interim, 0.72 for MERRA, and 0.26 for JRA-55 over the 60ºS-60ºN domain. In the case of the 30ºS-30ºN domain, the correlation coefficients are 0.77 for ERA-Interim, 0.70 for MERRA, and 0.51 for JRA-55. The lower correlation coefficients for JRA-55 are due to the discontinuity and negative trend around the early 2000s, which is also evident in time series of domain-averaged relative humidity anomalies in the upper troposphere (Fig. S1). This moistening trend could be explained, in part, by the fact that a dry bias of the JRA-55 forecast model in the upper and middle troposphere has been reduced gradually with the increase of the number of observations from satellite water vapor channels [*Kobayashi* et al., 2015]. These results indicate the limited skill of the reanalysis data sets in replicating the observed UTWV trends. While these reanalysis data sets assimilate both in situ and satellite observations, only observed sea surface temperatures and sea ice concentration are prescribed in the AMIP simulations. As a result, the correlation coefficient between the HIRS measurements and the

multi-model ensemble mean is substantially low (0.39 over the 60°S-60°N domain and 0.38 over the 30°S-30°N domain).

In addition to the differences in the correlation coefficient, the reanalysis data sets show discrepancies in trends between them (Table 1 lists the linear least-squares trends and their statistical significance for reanalysis-simulated HIRS channel 12 BTs over the two domains). While JRA-55 exhibits a negative trend for the period 1979-2013, MERRA produces a positive trend for the period 1979-2014. The trend for ERA-Interim is closer to the HIRS measurements. Divergence among the reanalysis data sets in the simulated HIRS channel 12 BT trends is qualitatively consistent with Dessler and Davis [2010], in which trends in reanalysis-produced specific humidity were assessed using AIRS-retrieved specific humidity as a benchmark. In the case of the AMIP model simulations, the multi-model ensemble mean shows generally consistent decadal time-scale variability with the HIRS record.

Previous studies have shown that interannual variations of tropical UTWV are closely associated with El Niño-Southern Oscillation (ENSO) events [e.g., *Bates* et al., 2001; *Shi and Bates*, 2011]. By conducting an empirical orthogonal function (EOF) analysis over the tropics (30°S-30°N), we examine whether the ENSO-related variations in UTWV captured in the HIRS observations are consistent with the other two satellite observations. Figure 3a displays the spatial distribution of the leading EOF mode for HIRS channel 12 BTs over the period 1979-2014. In agreement with previous studies [*Bates* et al., 2001; *Shi and Bates*, 2011], the obtained spatial patterns highlight the dominant influence of ENSO on HIRS channel 12 BTs over the tropics: negative values (i.e., moist upper troposphere) along the equatorial central to eastern

Pacific Ocean are surrounded by positive values (i.e., dry upper troposphere) over the eastern Indian/western Pacific Ocean including the maritime continent, subtropical regions of the Pacific Ocean, and equatorial South America.

Performing EOF analysis over the period 2003-2014 (Fig. 3b), for which the AIRS and AMSU-B/MHS observations are available, produces similar spatial distribution of the leading EOF mode to that for the period 1979-2014. However, because of increasing occurrence frequency of non-canonical El Niño in recent decades [e.g., *Yeh* et al., 2007], noticeable discrepancies from the period 1979-2014 are found over the eastern Pacific Ocean (15ºS-Eq.); in addition, the center of the negative values located in the equatorial central Pacific shifted westward relative to the period 1979-2014. In spite of these differences, the contribution from the leading EOF mode is 9.4% for both of the periods. The influence of increasing occurrence frequency of non-canonical El Niño is also evident in the spatial distribution of the leading EOF mode for AMSU-B/MHS 183.31±1 GHz channel (Fig. 3c, explained variance: 10.5%) and AIRS (Fig. 3d, explained variance: 8%) for the period 2003-2014, which indicates that the different patterns from the period 1979-2014 are not artifacts. Due to differences in spectral characteristics and spatiotemporal sampling, the local values of the leading EOF mode show discernible discrepancies between the satellite data sets. Nonetheless, the position of the centers of maxima and minima is in good agreement between these independent satellite data sets.

The spatial patterns of the leading EOF mode for the reanalysis data sets are comparable to the HIRS observations for the period 1979-2014, implying the dominant influence of ENSO on reanalysis-produced UTWV (not shown). The percentages of the variance explained by the

leading EOF mode are 11.4% for ERA-Interim over the period 1979-2014, 11.2% for MERRA over the period 1979-2014, and 10.7% for JRA-55 over the period 1979-2013.

Figure 4a presents the corresponding principal component of the leading EOF mode for the three satellite products. The temporal evolution of the principal component for HIRS channel 12 over the period 2003-2014 is in phase with that for AMSU-B/MHS 183.31±1 GHz channel and AIRS, indicating that these three observations are consistent in the ENSO-related interannual variability with each other. These results demonstrate the high degree of consistency between the spatial and temporal variability of all three data sets.

Figure 4b compares the temporal evolution of the principal component for simulated HIRS channel 12 BTs among the three reanalysis data sets. Despite the substantial discrepancies shown in Fig. 2, the reanalysis data sets exhibit highly correlated temporal evolution of the principal component. Furthermore, the temporal evolution of the principal component is consistent with the satellite observations with correlations of ~0.95. These results indicate that the reanalysis data sets reasonably capture the influence of ENSO on UTWV, which suggests the credibility of the sub-decadal variability of UTWV in the reanalysis data sets.

3.2 Long-term Variability

Figure 5a displays the spatial distribution of the trends of HIRS channel 12 BTs over the period 1979-2014 with superimposed contours indicating the annual mean BTs. Positive BT trends imply a decrease in upper tropospheric relative humidity and vice-versa. The equatorial central Pacific exhibits positive trends over the period, whereas the opposite sign is evident over

the surrounding regions. This spatial pattern reflects changes in ENSO activity over this period; i.e., El Niño conditions were more frequent prior to the end of the 1990s, and less common during the latter portion of the record. Shi and Bates [2011] also showed that HIRS channel 12 BTs are correlated with the Pacific Decadal Oscillation over this period. Compared to the tropical Pacific Ocean, the other ocean basins show a relatively uniform spatial distribution of the trends with the positive (negative) sign being prevalent in the North Pacific and the Southern Ocean (the North Atlantic). Most continental regions show positive trends except for Europe.

Figures 5b-d compare the trends in BT among the three satellite data sets over the period 2003-2014. All three products produce a similar spatial distribution of BT trends to the HIRS channel 12 over the period 1979-2014, but the HIRS trends are noticeably weaker than either the AMSU-B/MHS or AIRS trends. This may reflect coverage differences arising from the differing cloud clearance techniques and spatial resolution of the sensors, which are known to cause significant biases in the mean climatology of HIRS upper tropospheric humidity products [*John* et al., 2011]. The microwave products are only affected by heavy rain, whereas all cloud types affect the HIRS channel 12 BT. AIRS radiances are also affected by all cloud types, however its higher spatial resolution allows for improved sampling of partially cloudy scenes. The consistency between the AIRS and AMSU-B/MHS trends implies that the trends for the full HIRS record (1979-2014) might also be biased weak.

Decadal trends are also computed for simulated HIRS channel 12 BTs from the reanalysis data sets (Figs. 6b-d). The superimposed contours represent the annual mean BTs. Comparisons between the observed and simulated annual mean BTs indicate discrepancies in the annual mean

BTs that appear to result from the biases in the reanalysis-produced and model-simulated temperature and water vapor [e.g., *John and Soden*, 2007; *Chung* et al., 2011; *Tian* et al., 2013] in conjunction with the clear-sky sampling biases of infrared satellite observations [e.g., *Sohn and Schmetz*, 2004; *John* et al., 2011]. In spite of the discrepancies in the annual mean BTs, features of the general circulation depicted in the reanalysis data sets exhibit broad agreement with the HIRS observations for the period 1979-2014 (Fig. 6a): positive trends over the equatorial central Pacific are surrounded by negative trends over the maritime continent, the South Pacific convergence zone and the northern part of subtropical regions in the tropical Pacific. In addition, the northern Indian Ocean exhibits negative trends in agreement with the HIRS observations. However, it is questionable whether important climatological details are adequately captured by the reanalyses. For instance, distinct discrepancies relative to the HIRS observations are pronounced over the eastern Pacific Ocean (30ºS-Eq.) for all three reanalysis data sets and over Europe and the North Atlantic for ERA-Interim and MERRA. In particular, while there is a clear tendency for the climatologically dry regions of the subtropics to become more humid in the HIRS observations, the dry regions in the reanalysis data sets have become drier or exhibit spatially inhomogeneous decadal time-scale variability. Comparisons with Fig. 6a also indicate that the regions of negative (positive) trends have larger magnitude in JRA-55 (MERRA). These discrepancies demonstrate the limited skill of reanalyses in replicating observed trends.

Figures S2 and S3 display statistical confidence levels for the trends shown in Fig. 5 and Fig. 6, respectively, which are computed by considering reduced degrees of freedom to account

for auto-correlation in the time series. The estimated confidence levels are not much high for the satellite observations over the period 2003-2014. However, with increase in time span, a substantial portion of the analysis domain exhibits statistically significant (above 90% level) trends for the period 1979-2014.

Figure 7 compares the multi-model ensemble mean trends from the AMIP simulations with the observed trends for the period 1979-2008. The climate models integrated with prescribed sea surface temperatures and sea ice concentration produce spatial patterns of trends that reflect the changes in ENSO activity over that period. However, the trends over the tropical eastern Pacific are noticeably different from the HIRS observations, and the region of positive trends in the equatorial Pacific is displaced eastward compared to the HIRS observations. Similar to the reanalysis data sets, the ensemble-mean of the AMIP models appears to have difficulty in capturing the observed feature that the climatologically dry regions of the subtropics have become more humid. These discrepancies seem to imply that the temporal evolution of observed sea surface temperatures in the tropical eastern Pacific was not the primary factor governing the decadal-scale changes in UTWV over there during the satellite era.

Although zonal averaging masks detailed regional features, it may help to highlight any systematic difference between observed and simulated features by averaging out random errors. The zonal-mean trends of simulated BTs from the reanalysis data sets and climate model simulations for CMIP5 AMIP experiment are compared with the HIRS measurements (black) in Fig. 8. The vertical bars for observations (black) and reanalyses (purple, blue and green) indicate twice the standard error of the trends and are computed using the method of Weatherhead et al.

[1998]. The orange line and associated vertical bars represent the multi-model ensemble mean and intermodel spread (±2 standard deviation) of the zonal-mean trends, respectively. ERA-Interim (purple) shows generally similar, albeit not identical, latitudinal distributions to the HIRS measurements. In contrast, the other two reanalysis data sets exhibit noticeably different zonal-mean trends. In the case of MERRA (blue), the zonal-mean trends are roughly comparable to the HIRS measurements in the extra-tropics and near the equator, however positive biases are evident in the subtropical regions. The latitudinal distributions of the zonal-mean trends for JRA-55 (green) are similar to those for MERRA. However, unlike both MERRA and the HIRS measurements, the zonal-mean trends are negative for most of the latitudes, which implies that upper tropospheric relative humidity has spuriously increased in JRA-55. In the case of the multi-model ensemble mean of the AMIP simulations, the latitudinal variations in the zonal-mean trends apparent in both the observations and the reanalysis data sets virtually disappear. The noticeable intermodel spread represented by the error bars implies intermodel differences in the model physics and parameterization that determine the water budget in the upper troposphere.

To determine whether the biases in UTWV trends noted in Figures 5-7 are manifest at larger spatial scales, Table 1 lists the trends of observed and simulated BTs over the near-globe (60^oS-60^oN) and the tropics (30^oS-30^oN). For the period 2003-2014, the trends from HIRS are positive over both of the domains whereas the trends are negative for AMSU-B/MHS and AIRS. However the HIRS trends are within ±2 standard error interval of either the AMSU-B or AIRS trends; this indicates that at the larger spatial scales the trends between all three products are statistically indistinguishable from each other.

The trends of simulated BTs from ERA-Interim are greater than the HIRS measurements, but are statistically indistinguishable from the HIRS measurements. On the other hand, the other two reanalysis data sets produce trends that are significantly different from the HIRS measurements: while MERRA shows statistically significant positive trends, JRA-55 exhibits statistically significant negative trends. It is noted that the trends of simulated BTs are statistically distinguishable from each other among the three reanalysis data sets.

Table 1 also lists the multi-model ensemble mean and intermodel spread (±2 standard deviation) of the trends from climate model simulations for the AMIP experiment. In spite of distinctly low correlations with the observed HIRS time series (Fig. 2), the multi-model ensemble means of the trends are closer to the satellite-determined trends than to the reanalysis trends. To further examine the consistency with the satellite observations, we also simulate HIRS channel 12 BTs for the period 1979-2014 from coupled simulations of the CMIP5. These models were integrated with historical changes in anthropogenic greenhouse gases, aerosols, solar activity and volcanic eruptions through 2005 (historical scenario), followed by the representative concentration pathway (RCP) 4.5 scenario for 2006-2014. Sea surface temperatures and sea ice concentration are not prescribed in the coupled simulations. Taylor et al. [2012] provide detailed information on the historical and RCP4.5 experiments of CMIP5.

The rightmost column of Table 1 provides the multi-model mean and intermodel spread of the trends for the coupled ocean-atmosphere model simulations. Because of internal variability that is suppressed in the AMIP simulations through prescribed sea surface temperatures, the intermodel spread of the trends is greater compared to that for the AMIP simulations.

Nonetheless, the multi-model ensemble mean values are very close to both the HIRS measurements and the multi-model ensemble mean values for the AMIP simulations. This means that upper tropospheric relative humidity has remained nearly constant during the period 1979-2014. In contrast, the trends from the CMIP5 models lie outside the range of the trends from both MERRA and JRA-55, implying that the statistically significant negative (positive) trends from JRA-55 (MERRA) are artifacts.

4. Summary and Discussion

We have assessed the consistency of satellite-based observations of UTWV by comparing the variability and trends between three independent BT products: HIRS channel 12 [*Shi and Bates*, 2011], AMSU-B/MHS 183.31±1 GHz channel [*Chung* et al., 2013b] and AIRS radiances convolved with the SRF of HIRS channel 12.

Domain-averaged BT anomalies exhibit a highly-correlated temporal evolution among the three products over their common period (2003-2014) despite the shift in the central wavelength of HIRS channel 12 between the HIRS/2 and HIRS/3. This shift resulted in a BT difference of ~8 K between HIRS/2 and HIRS/3, however the good agreement between HIRS channel 12 with the other two observations confirms that a large portion of the bias has been accounted for by a series of intersatellite calibration processes [*Shi and Bates*, 2011]. The three products were also shown to exhibit similar patterns of interannual variability, which highlight the role of ENSO in modulating the distribution of water vapor on sub-decadal time scales.

On decadal time-scales, the spatial patterns of trends are similar between all three products, but the amplitude of the regional trends is noticeably weaker in the HIRS measurements than in either the AMSU-B/MHS or AIRS data. This presumably reflects the greater clear-sky sampling limitations of HIRS, relative to the other products. However, when averaged over tropical or near-global spatial scales, the trends between all three products are statistically indistinguishable from each other. The overall consistency between all three products provides important verification of their credibility for documenting long-term changes in UTWV.

Since reanalysis data sets and climate model simulations are frequently used to analyze the variability of key climate variables, we have compared the patterns of UTWV variability from reanalysis data sets and climate model simulations with the HIRS measurements. Reanalysis-produced UTWV data sets exhibit spatial and temporal patterns of interannual variability consistent with the satellite observations. However, the spatial patterns of reanalysis-produced trends are distinctly different from the HIRS observations. In particular, the reanalysis data sets fail to capture the observed feature that the climatologically dry regions of the subtropics have become more humid. The reanalysis data sets also show substantial discrepancies in trends between them. When averaged over tropical or near-global spatial scales, JRA-55 exhibits statistically significant negative trends, whereas MERRA produces statistically significant positive trends. In contrast, the trends from ERA-Interim are comparatively close to the satellite observations. Given that these reanalysis data sets assimilate a number of in situ and satellite observations, the satellite-reanalyses discrepancies imply that compared to the intersatellite-calibrated data sets assessed in this study assimilated satellite or in situ observations might have

spurious long-term trends as a result of degradation of the sensitivity of humidity sensors, satellite orbital drift, or changes in calibration systems [e.g., *Bates* et al., 1996, 2001; *Jackson and Soden*, 2007; *Shi and Bates*, 2011; *Chung* et al., 2013b]. In particular, the discontinuity between the HIRS/2 and HIRS/3 instruments might not be adequately corrected for in assimilation processes. The limited skill of reanalyses in replicating the observed UTWV trends might also be indicative of potential deficiency in model physics that affects the interplay among large-scale circulation, deep convection and lateral mixing by transient eddies [e.g., *Chung* et al., 2013a]. The spatial patterns of the trends computed from CMIP5 climate model simulations are also noticeably different from the HIRS observations. As with the reanalysis data sets, the ensemble-mean of the AMIP models fails to reproduce the observed feature that the climatologically dry regions of the subtropics have become more humid. However, the tropical or near-global trends are statistically indistinguishable from the HIRS observations, implying that upper tropospheric relative humidity has remained nearly constant during the satellite era.

Acknowledgments

We would like to thank three anonymous reviewers for their constructive and valuable comments which led to an improved version of the manuscript. We also thank Dr. David Parker and Dr. Shinya Kobayashi for their helpful comments on the manuscript. AIRS Level 1B infrared geolocated and calibrated radiances (version 005) were downloaded from Goddard Earth Science Data and Information Service Center (GES DISC, http://disc.sci.gsfc.nasa.gov/datacollection/AIRIBRAD_V005.html). ERA-Interim reanalysis

data were downloaded from the ECMWF data server (http://apps.ecmwf.int/datasets/data/interim-full-moda). MERRA reanalysis data were downloaded from the NASA Goddard Earth Sciences Data and information Services Center (http://disc.sci.gsfc.nasa.gov/daac-bin/DataHoldings.pl). JRA-55 reanalysis data were downloaded from the Research Data Archive at the National Center for Atmospheric Research (http://rda.ucar.edu/datasets/ds628.1/). We acknowledge the World Climate Research Programme's Working Group on Coupled Modeling, which is responsible for CMIP, and we thank the climate modelling groups for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. CMIP5 data are available from the PCMDI archive. This study was supported by NASA ROSES: Satellite Calibration Interconsistency Studies. Viju John was supported by the Joint UK DECC/Defra Met Office Hadley Centre Climate Programme (GA01101).

References

Allan, R. P. (2012), The role of water vapour in Earth's energy flows, *Surv. Geophys.*, **33**, 557-564, doi:10.1007/s10712-011-9157-8.

Aumann, H. H., S. Broberg, D. Elliott, S. Gaiser, and D. Gregorich (2006), Three years of Atmospheric Infrared Sounder radiometric calibration validation using sea surface temperatures, *J. Geophys. Res.*, **111**, D16S90, doi:10.1029/2005JD006822.

Bates, J. J., and D. L. Jackson (2001), Trends in upper-tropospheric humidity, *Geophys. Res. Lett.*, **28**, 1695-1698, doi:10.1029/2000GL012544.

Bates, J. J., X. Wu, and D. L. Jackson (1996), Interannual variability of upper-troposphere water vapor band brightness temperatures, *J. Clim.*, **9**, 427-438.

Bates, J. J., D. L. Jackson, F.-M. Bréon, and Z. D. Bergen (2001), Variability of tropical upper tropospheric humidity 1979-1998, *J. Geophys. Res.*, **106**, 32,271-32,281, doi: 10.1029/2001JD000347.

Brogniez, H., and R. T. Pierrehumbert (2006), Using microwave observations to assess large-scale control of free tropospheric water vapor in the mid-latitudes, *Geophys. Res. Lett.*, **33**, L14801, doi:10.1029/2006GL026240.

Brogniez, H., R. Roca, and L. Picon (2006), A clear-sky radiance archive from Meteosat ''water vapor'' observations, *J. Geophys. Res.*, **111**, D21109, doi:10.1029/2006JD007238.

Buehler, S. A., M. Kuvatov, V. O. John, M. Milz, B. J. Soden, D. L. Jackson, and J. Notholt (2008), An upper tropospheric humidity data set from operational satellite microwave data, *J. Geophys. Res.*, **113**, D14110, doi:10.1029/2007JD009314.

Chung, E.-S., B. J. Soden, B.-J. Sohn, and J. Schmetz (2011), Model-simulated humidity bias in the upper troposphere and its relation to the large-scale circulation, *J. Geophys. Res.*, **116**, D10110, doi:10.1029/2011JD015609.

Chung, E.-S., B. J. Soden, B. J. Sohn, and J. Schmetz (2013a), An assessment of the diurnal variation of upper tropospheric humidity in reanalysis data sets, *J. Geophys. Res. Atmos.*, **118**, 3425-3430, doi:10.1002/jgrd.50345.

Chung, E.-S., B. J. Soden, and V. O. John (2013b), Intercalibrating microwave satellite observations for monitoring long-term variations in upper- and midtropospheric water vapor, *J. Atmos. Oceanic Technol.*, **30**, 2303-2319, doi:10.1175/JTECH-D-13-00001.1.

Chung, E.-S., B. Soden, B. J. Sohn, and L. Shi (2014), Upper-tropospheric moistening in response to anthropogenic warming, *Proc. Natl. Acad. Sci. USA*, **111**, 11,636-11,641, doi: 10.1073/pnas.1409659111.

Dee, D. P., et al. (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, **137**, 553-597, doi:10.1002/qj.828.

Dessler, A. E., and S. M. Davis (2010), Trends in tropospheric humidity from reanalysis systems, *J. Geophys. Res.*, **115**, D19127, doi:10.1029/2010JD014192.

Dessler, A. E., Z. Zhang, and P. Yang (2008), Water-vapor climate feedback inferred from climate fluctuations, 2003–2008, *Geophys. Res. Lett.*, **35**, L20704, doi:10.1029/2008GL035333.

Engelen, R. J., and G. L. Stephens (1999), Characterization of water-vapor retrievals from TOVS/HIRS and SSM/T-2 measurements, *Q. J. R. Meteorol. Soc.*, **125**, 331-351.

Gaiser, S. L., H. H. Aumann, D. T. Gregorich, and T. J. Hearty (2003), In-flight refinement of the radiometric, spectral, and spatial calibration of the Atmospheric Infrared Sounder (AIRS), Proc. SPIE 5151, Earth Observing Systems VIII, 232-243, doi:10.1117/12.506380.

Gettelman, A., and Q. Fu (2008), Observed and simulated upper-tropospheric water vapor feedback, *J. Clim.*, **21**, 3282-3289, doi:10.1175/2007JCLI2142.1.

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

Hocking, J., P. Rayer, R. Saunders, M. Matricardi, A. Geer, and P. Brunel (2011), RTTOV v10 Users Guide. NWP SAF, EUMETSAT.

Huang, X., and Y. L. Yung (2005), Spatial and spectral variability of the outgoing thermal IR spectra from AIRS: A case study of July 2003, *J. Geophys. Res.*, **110**, D12102, doi: 10.1029/2004JD005530.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, and S. A. Clough (2003), Evaluation of upper tropospheric water vapor in the NCAR Community Climate Model (CCM3) using modeled and observed HIRS radiances, *J. Geophys. Res.*, **108**, 4037, doi:10.1029/2002JD002539.

Jackson, D. L., and B. J. Soden (2007), Detection and correction of diurnal sampling bias in HIRS/2 brightness temperatures, *J. Atmos. Oceanic Technol.*, **24**, 1425-1438, doi:10.1175/JTECH2062.1.

Jackson, D. L., D. P. Wylie, and J. J. Bates (2003), The HIRS pathfinder radiance data set (1979-2001), paper presented at 12th Conference on Satellite Meteorology and Oceanography, Am. Meteorol. Soc., Long Beach, Calif., 10-13 Feb.

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**, L18704, doi:10.1029/2007GL030429.

John, V. O., G. Holl, R. P. Allan, S. A. Buehler, D. E. Parker, and B. J. Soden (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, B. Candy, R. W. Saunders, and D. E. Parker (2012), Understanding intersatellite biases of microwave humidity sounders using global simultaneous nadir overpasses, *J. Geophys. Res.*, **117**, D02305, doi:10.1029/2011JD016349.

Kobayashi, S., et al. (2015), The JRA-55 Reanalysis: General specifications and basic characteristics, *J. Meteorol. Soc. Japan*, **93**, 5-48, doi:10.2151/jmsj.2015-001.

Mears, C. A., B. D. Santer, F. J. Wentz, K. E. Taylor, and M. F. Wehner (2007), Relationship between temperature and precipitable water changes over tropical oceans, *Geophys. Res. Lett.*, **34**, L24709, doi:10.1029/2007GL031936.

Pagano, T. S., H. H. Aumann, D. E. Hagan, and K. Overoye (2003), Prelaunch and in-flight radiometric calibration of the Atmospheric Infrared Sounder (AIRS), *IEEE Trans. Geosci. Remote Sens.*, **41**, 265-273, doi:10.1109/TGRS.2002.808324.

Pan, F., X. Huang, L. L. Strow, and H. Guo (2015), Linear trends and closures of 10-yr observations of AIRS stratospheric channels, *J. Clim.*, **28**, 8939-8950.

Pierrehumbert, R. T., and R. Roca (1998), Evidence for control of Atlantic subtropical humidity by large scale advection, *Geophys. Res. Lett.*, **25**, 4537-4540, doi:10.1029/1998GL900203.

Rienecker, M. M., et al. (2011), MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, *J. Clim.*, **24**, 3624-3648, doi:10.1175/JCLI-D-11-00015.1.

Salathé, E. P., and D. L. Hartmann (1997), A trajectory analysis of tropical upper-tropospheric moisture and convection, *J. Clim.*, **10**, 2533-2547.

Saunders, R., et al. (2007), A comparison of radiative transfer models for simulating Atmospheric Infrared Sounder (AIRS) radiances, *J. Geophys. Res.*, **112**, D01S90, doi:10.1029/2006JD007088.

Schmetz, J., and O. M. Turpeinen (1988), Estimation of the upper tropospheric relative humidity field from METEOSAT water vapor image data, *J. Appl. Meteor.*, **27**, 889-899.

Sherwood, S. C., R. Roca, T. M. Weckwerth, and N. G. Andronova (2010), Tropospheric water vapor, convection, and climate, *Rev. Geophys.*, **48**, RG2001, doi:10.1029/2009RG000301.

Shi, L., and J. J. Bates (2011), Three decades of intersatellite-calibrated High-Resolution Infrared Radiation Sounder upper tropospheric water vapor, *J. Geophys. Res.*, **116**, D04108, doi:10.1029/2010JD014847.

Soden, B. J., and F. P. Bretherton (1993), Upper tropospheric relative humidity from the GOES 6.7 μm channel: Method and climatology for July 1987, *J. Geophys. Res.*, **98**, 16,669-16,688, doi:10.1029/93JD01283.

Soden, B. J., and F. P. Bretherton (1994), Evaluation of water vapor distribution in general circulation models using satellite observations, *J. Geophys. Res.*, **99**, 1187-1210, doi:10.1029/93JD02912.

Soden, B. J., and I. M. Held (2006), An assessment of climate feedbacks in coupled ocean-atmosphere models, *J. Clim.*, **19**, 3354-3360, doi:10.1175/JCLI3799.1.

Soden, B. J., D. L. Jackson, V. Ramaswamy, M. D. Schwarzkopf, and X. Huang (2005), The radiative signature of upper tropospheric moistening, *Science*, **310**, 841-844, doi: 10.1126/science.1115602.

Sohn, B.-J., and J. Schmetz (2004), Water vapor-induced OLR variations associated with high cloud changes over the Tropics: A study from Meteosat-5 observations, *J. Clim.*, **17**, 1987-1996.

Sohn, B. J., et al. (2000), Intercalibration of the Meteosat-7 water vapor channel with SSM/T-2, *J. Geophys. Res.*, **105**, 15,673-15,680, doi:10.1029/2000JD900188.

Sohn, B.-J., E.-S. Chung, J. Schmetz, and E. A. Smith (2003), Estimating upper-tropospheric water vapor from SSM/T-2 satellite measurements, *J. Appl. Meteror.*, **42**, 488-504.

Sohn, B.-J., J. Schmetz, and E.-S. Chung (2008), Moistening processes in the tropical upper troposphere observed from Meteosat measurements, *J. Geophys. Res.*, **113**, D13109, doi:10.1029/2007JD009527.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, *Bull. Am. Meteorol. Soc.*, **93**, 485–498, doi:10.1175/ BAMS-D-11-00094.1.

Tian, B., E. J. Fetzer, B. H. Kahn, J. Teixeira, E. Manning, and T. Hearty (2013), Evaluating CMIP5 models using AIRS tropospheric air temperature and specific humidity climatology, *J. Geophys. Res. Atmos.*, **118**, 114-134, doi:10.1029/2012JD018607.

Weatherhead, E. C., et al. (1998), Factors affecting the detection of trends: Statistical considerations and applications to environmental data, *J. Geophys. Res.*, **103**(D14), 17,149-17,161, doi:10.1029/98JD00995.

Wentz, F. J., and M. Schabel (2000), Precise climate monitoring using complementary satellite data sets, *Nature*, **403**, 414-416.

Yeh, S.-W., et al. (2007), El Niño in a changing climate, *Nature*, **461**, 511-514.

List of Tables

Table 1: The linear least-squares trend and ±2 standard errors of the linear trend for satellite-observed BTs and simulated HIRS channel 12 BTs from ERA-Interim, MERRA and JRA-55. The rightmost two columns provide the multi-model mean and intermodel spread (±2 standard deviation) of the linear least-squares trend for simulated HIRS channel 12 BTs from CMIP5 climate models for the AMIP experiment for the period 1979-2008 and for the combined historical and RCP4.5 experiments for the period 1979-2014. The units are in K decade^{-1}. A trend can be considered to meet the 95% confidence level if the magnitude of the trend is larger than twice the standard error. Numerals in bold indicate the trend is statistically indistinguishable from the trend from HIRS channel 12 for the period 1979-2014.

	HIRS channel 12		AMSU-B/MHS 183.31±1 GHz channel	AIRS-derived HIRS channel 12	ERA-Interim	MERRA	JRA-55	AMIP	Historical+ RCP4.5
Period	1979-2014	2003-2014	2003-2014	2003-2014	1979-2014	1979-2014	1979-2013	1979-2008	1979-2014
60°S-60°N	0.063 ± 0.022	0.038 ± 0.074	-0.055 ± 0.090	-0.020 ± 0.081	**0.086 ± 0.025**	0.150 ± 0.025	-0.155 ± 0.049	**0.055 ±0.021**	**0.047 ± 0.039**
30°S-30°N	0.039 ± 0.029	0.006 ± 0.108	-0.089 ± 0.125	-0.076 ± 0.138	**0.079 ± 0.035**	0.179 ± 0.034	-0.200 ± 0.050	**0.065 ±0.033**	**0.050 ± 0.048**

List of Figures

Figure 1: Time series of domain-averaged monthly mean BT anomalies for HIRS channel 12, AMSU-B/MHS 183.31±1 GHz channel and AIRS-derived HIRS channel 12: (a) 60°S-60°N and (b) 30°S-30°N. The insets display difference time series (blue: AMSU-B/MHS minus HIRS, green: AIRS minus HIRS).

Figure 2: Time series of domain-averaged monthly mean HIRS channel 12 BT anomalies simulated from ERA-Interim, MERRA, JRA-55 and climate models for CMIP5 AMIP simulations for the period 1979-2008: (a) 60°S-60°N and (b) 30°S-30°N. The black lines denote the HIRS observations, and the orange lines represent the multi-model ensemble mean. Monthly time series are smoothed by applying 13-month running average for clarity.

Figure 3: Spatial distribution of the leading EOF mode for BTs over the tropics: (a) HIRS channel 12 for the period 1979-2014, (b) HIRS channel 12 for the period 2003-2014, (c) AMSU-B/MHS 183.31±1 GHz channel for the period 2003-2014, and (d) AIRS-derived HIRS channel 12 for the period 2003-2014.

Figure 4: Principal component of the leading EOF mode over the tropics: (a) satellite-observed BTs for HIRS channel 12 (1979-2014, black), HIRS channel 12 (2003-2014, red), AMSU-B/MHS 183.31±1 GHz channel (2003-2014, blue) and AIRS-derived HIRS channel 12 (2003-

2014, green) and (b) simulated HIRS channel 12 BTs from ERA-Interim (1979-2014, purple), MERRA (1979-2014, blue) and JRA-55 (1979-2013, green).

Figure 5: Trends of BTs: (a) HIRS channel 12 for the period 1979-2014, (b) HIRS channel 12 for the period 2003-2014, (c) AMSU-B/MHS 183.31±1 GHz channel for the period 2003-2014, and (d) AIRS-derived HIRS channel 12 for the period 2003-2014. The unit is in K decade^{-1}. Note that the color scale is different between the period 1979-2014 and the period 2003-2014. Contours represent the annual mean BTs (unit: K).

Figure 6: Trends of satellite-observed and reanalysis-simulated HIRS channel 12 BTs: (a) observations for the period 1979-2014, (b) ERA-Interim for the period 1979-2014, (c) MERRA for the period 1979-2014, and (d) JRA-55 for the period 1979-2013. The unit is in K decade^{-1}. Contours represent the annual mean BTs (unit: K).

Figure 7: Trends of satellite-observed and model-simulated HIRS channel 12 BTs for the period 1979-2008: (a) observations and (b) multi-model ensemble mean for AMIP simulations. The unit is in K decade^{-1}. Contours represent the annual mean BTs (unit: K).

Figure 8: Zonal-mean trends of simulated HIRS channel 12 BTs from ERA-Interim (1979-2014), MERRA (1979-2014), JRA-55 (1979-2013) and climate models for CMIP5 AMIP simulations (1979-2008). The black line denotes the HIRS observations for the period 1979-2014, and the

orange line represents the multi-model ensemble mean. In the case of the observations and the reanalysis data sets, the error bars denote ±2 standard errors of the linear trends. For the AMIP simulations, the error bars indicate ±2 intermodel standard deviation.

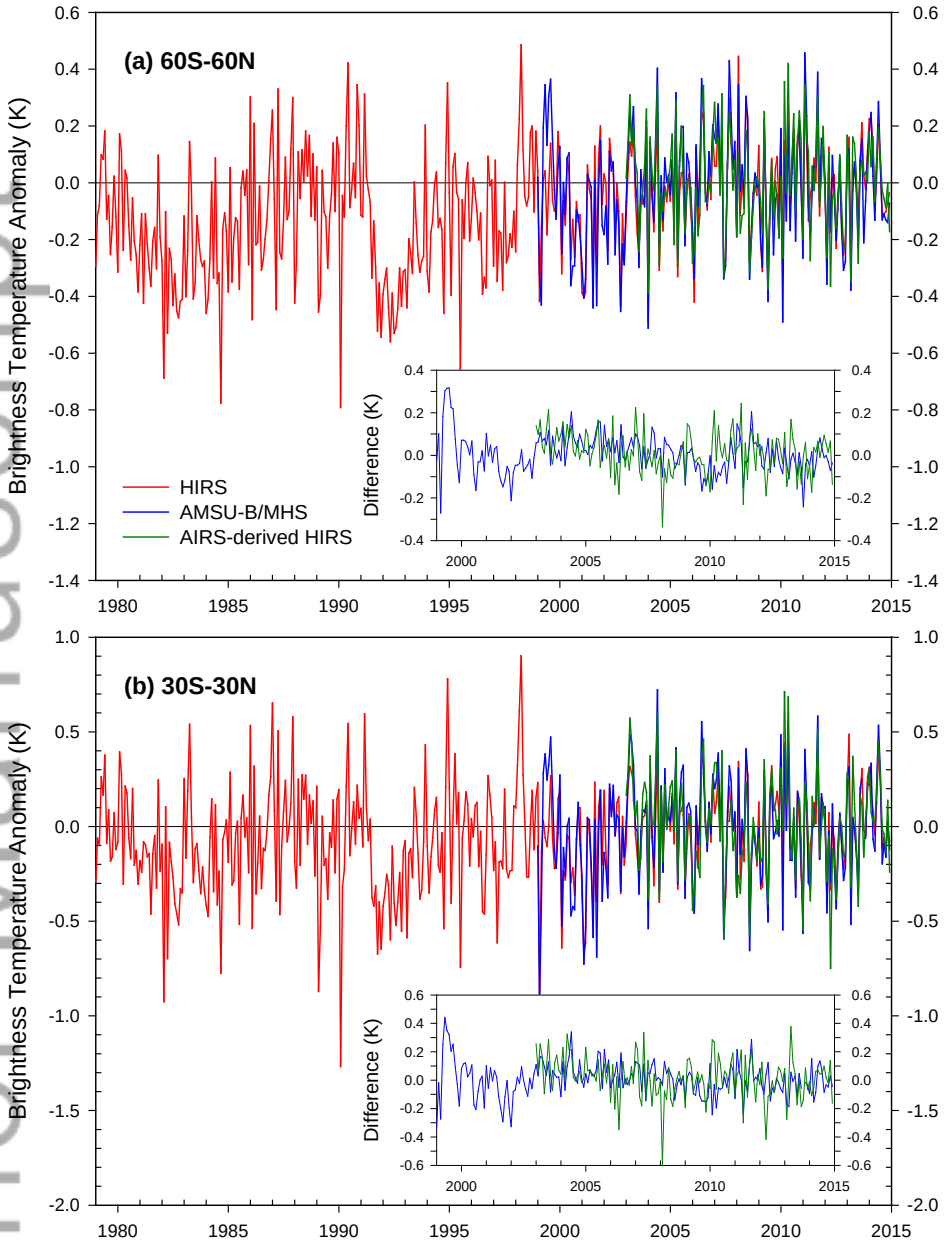


Figure 1: Time series of domain-averaged monthly mean BT anomalies for HIRS channel 12, AMSU-B/MHS 183.31±1 GHz channel and AIRS-derived HIRS channel 12: (a) 60°S-60°N and

(b) 30°S-30°N. The insets display difference time series (blue: AMSU-B/MHS minus HIRS, green: AIRS minus HIRS).

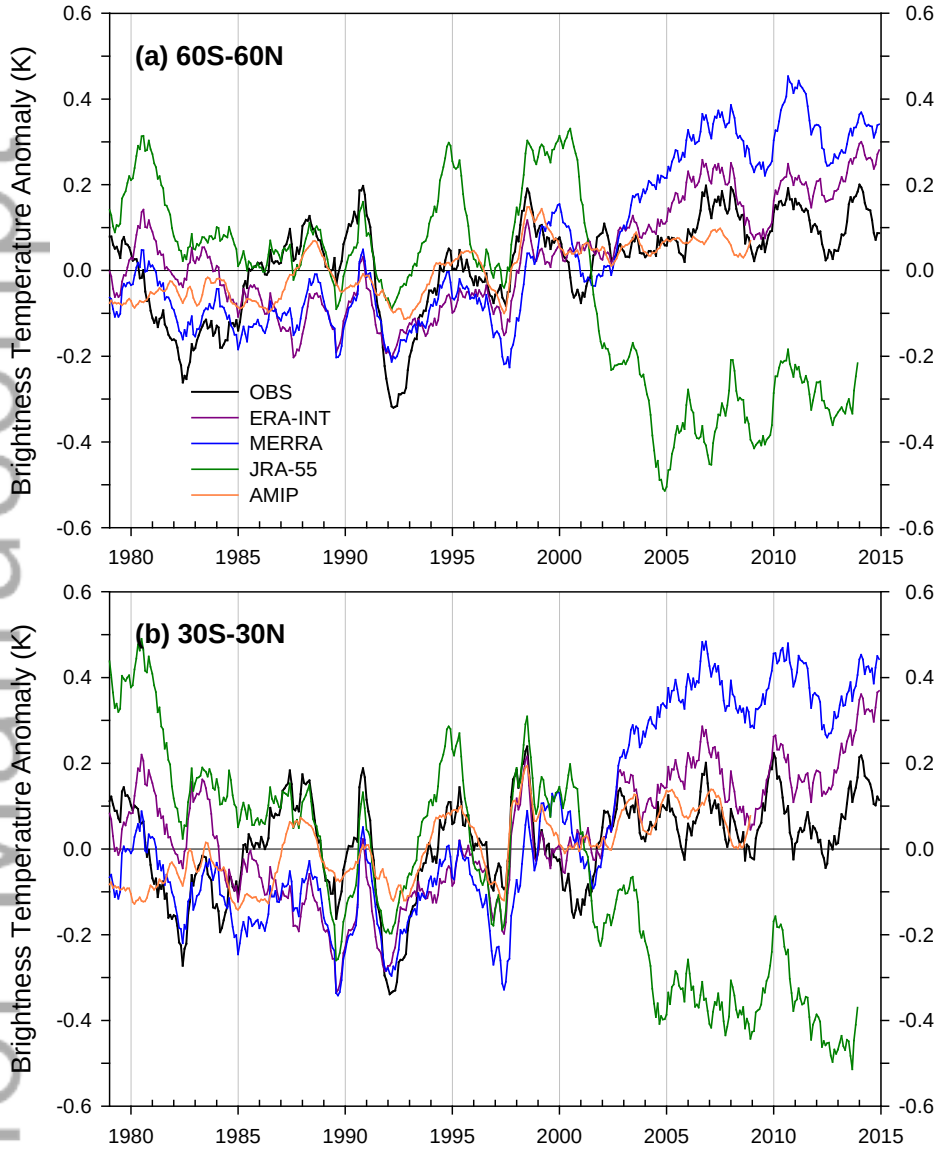


Figure 2: Time series of domain-averaged monthly mean HIRS channel 12 BT anomalies simulated from ERA-Interim, MERRA, JRA-55 and climate models for CMIP5 AMIP simulations for the period 1979-2008: (a) 60°S-60°N and (b) 30°S-30°N. The black lines denote

the HIRS observations, and the orange lines represent the multi-model ensemble mean. Monthly time series are smoothed by applying 13-month running average for clarity.

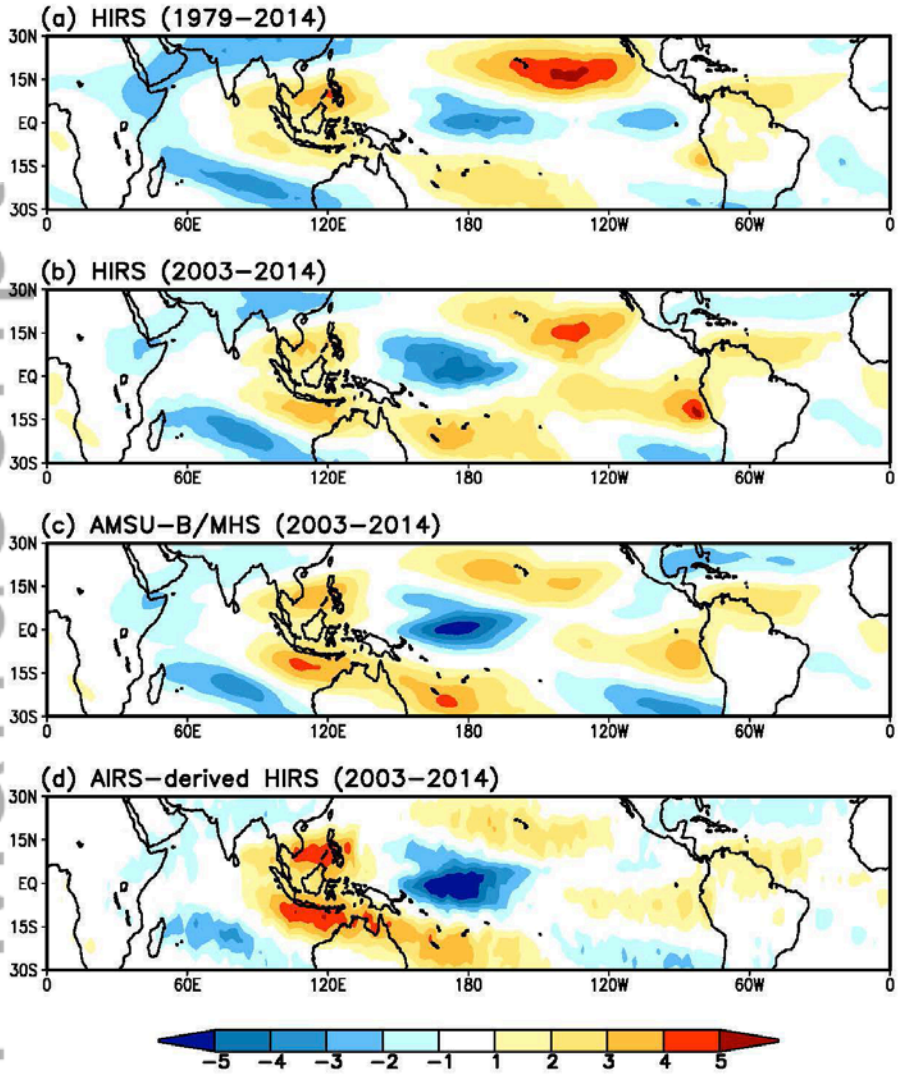


Figure 3: Spatial distribution of the leading EOF mode for BTs over the tropics: (a) HIRS channel 12 for the period 1979-2014, (b) HIRS channel 12 for the period 2003-2014, (c) AMSU-B/MHS 183.31±1 GHz channel for the period 2003-2014, and (d) AIRS-derived HIRS channel 12 for the period 2003-2014.

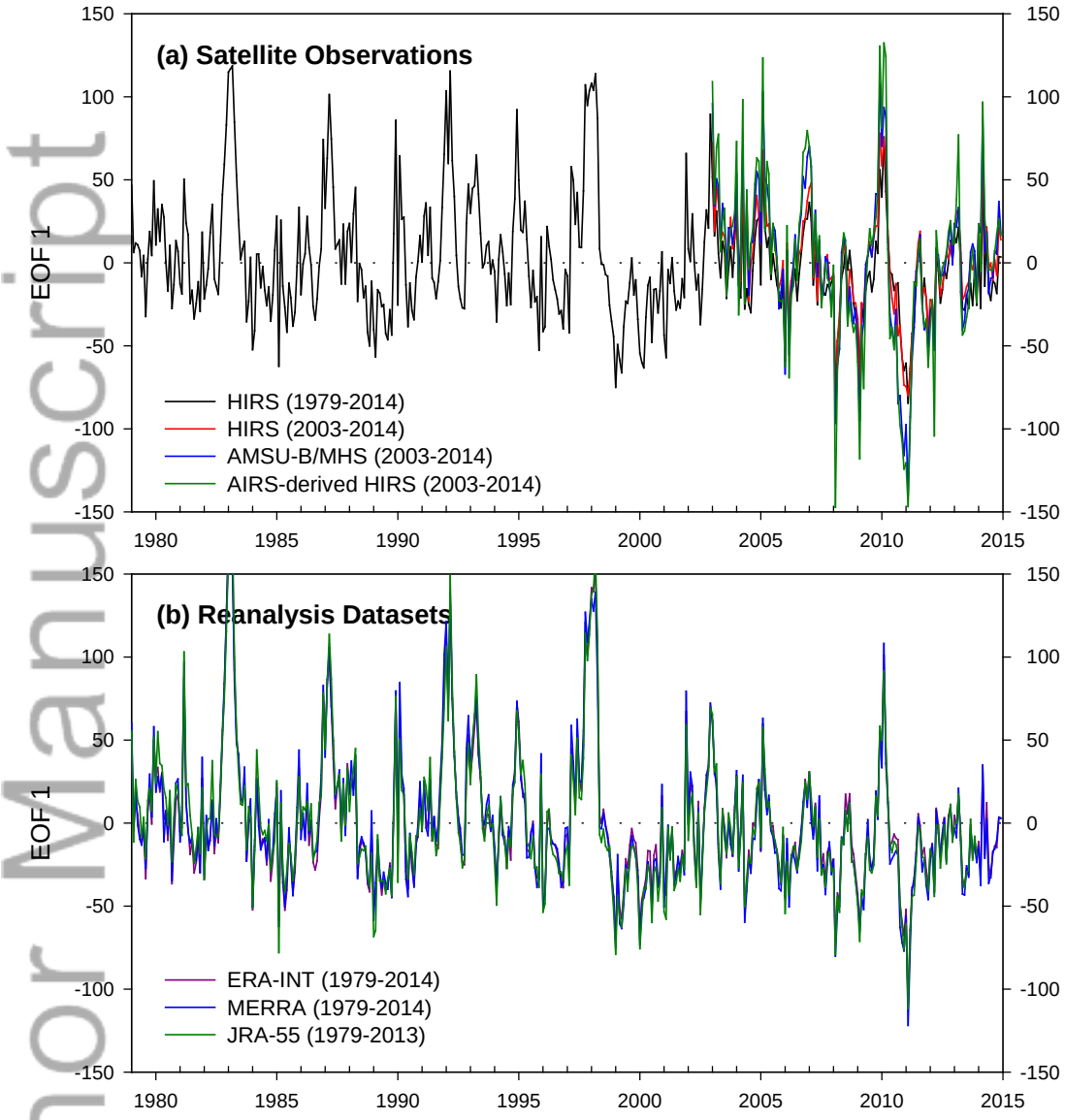


Figure 4: Principal component of the leading EOF mode over the tropics: (a) satellite-observed BTs for HIRS channel 12 (1979-2014, black), HIRS channel 12 (2003-2014, red), AMSU-B/MHS 183.31±1 GHz channel (2003-2014, blue) and AIRS-derived HIRS channel 12 (2003-

2014, green) and (b) simulated HIRS channel 12 BTs from ERA-Interim (1979-2014, purple), MERRA (1979-2014, blue) and JRA-55 (1979-2013, green).

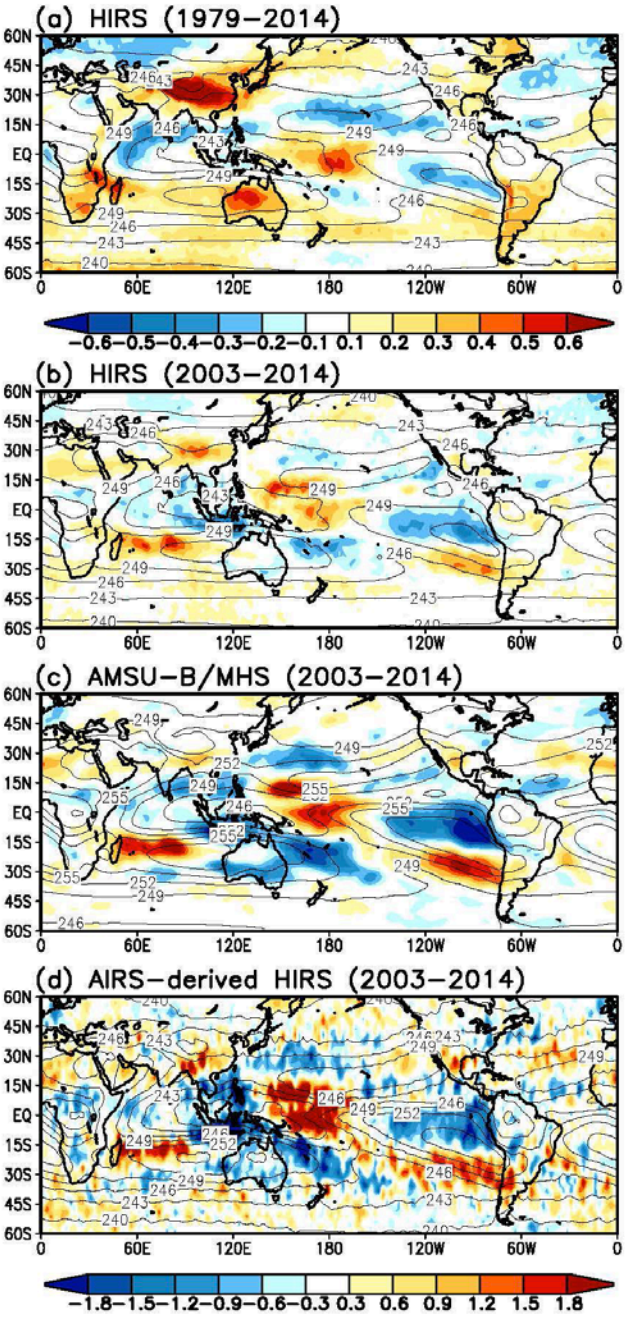


Figure 5: Trends of BTs: (a) HIRS channel 12 for the period 1979-2014, (b) HIRS channel 12 for the period 2003-2014, (c) AMSU-B/MHS 183.31±1 GHz channel for the period 2003-2014,

and (d) AIRS-derived HIRS channel 12 for the period 2003-2014. The unit is in K decade^{-1}. Note that the color scale is different between the period 1979-2014 and the period 2003-2014. Contours represent the annual mean BTs (unit: K).

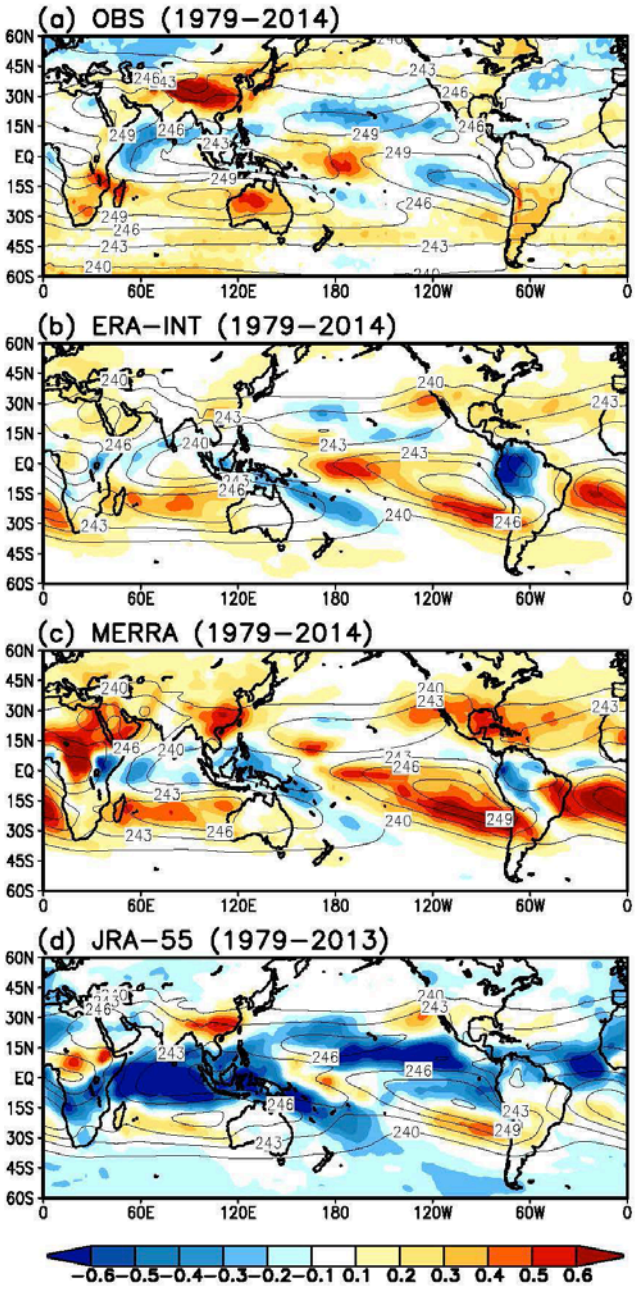


Figure 6: Trends of satellite-observed and reanalysis-simulated HIRS channel 12 BTs: (a) observations for the period 1979-2014, (b) ERA-Interim for the period 1979-2014, (c) MERRA

for the period 1979-2014, and (d) JRA-55 for the period 1979-2013. The unit is in K decade^{-1}. Contours represent the annual mean BTs (unit: K).

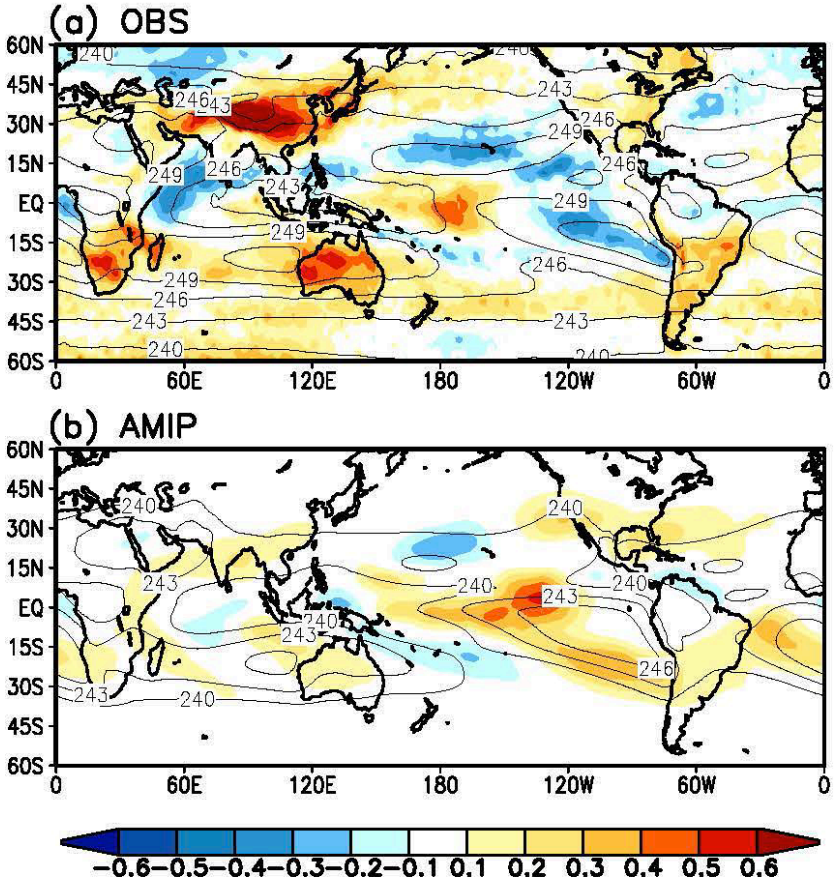


Figure 7: Trends of satellite-observed and model-simulated HIRS channel 12 BTs for the period 1979-2008: (a) observations and (b) multi-model ensemble mean for AMIP simulations. The unit is in K decade^{-1}. Contours represent the annual mean BTs (unit: K).

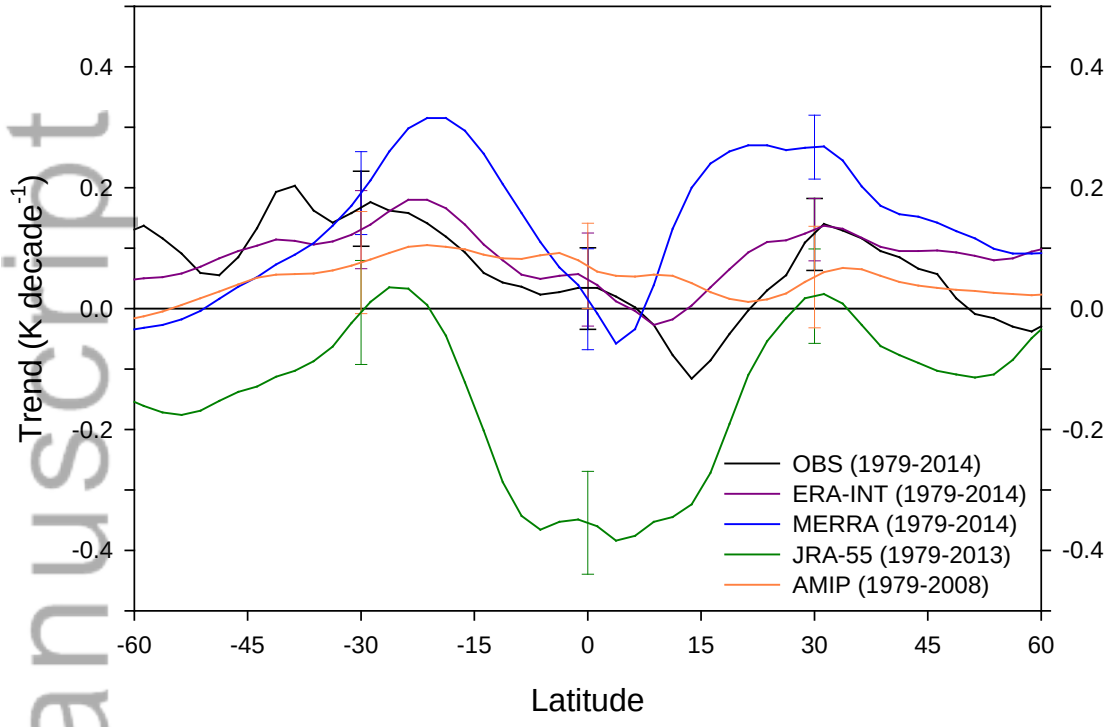
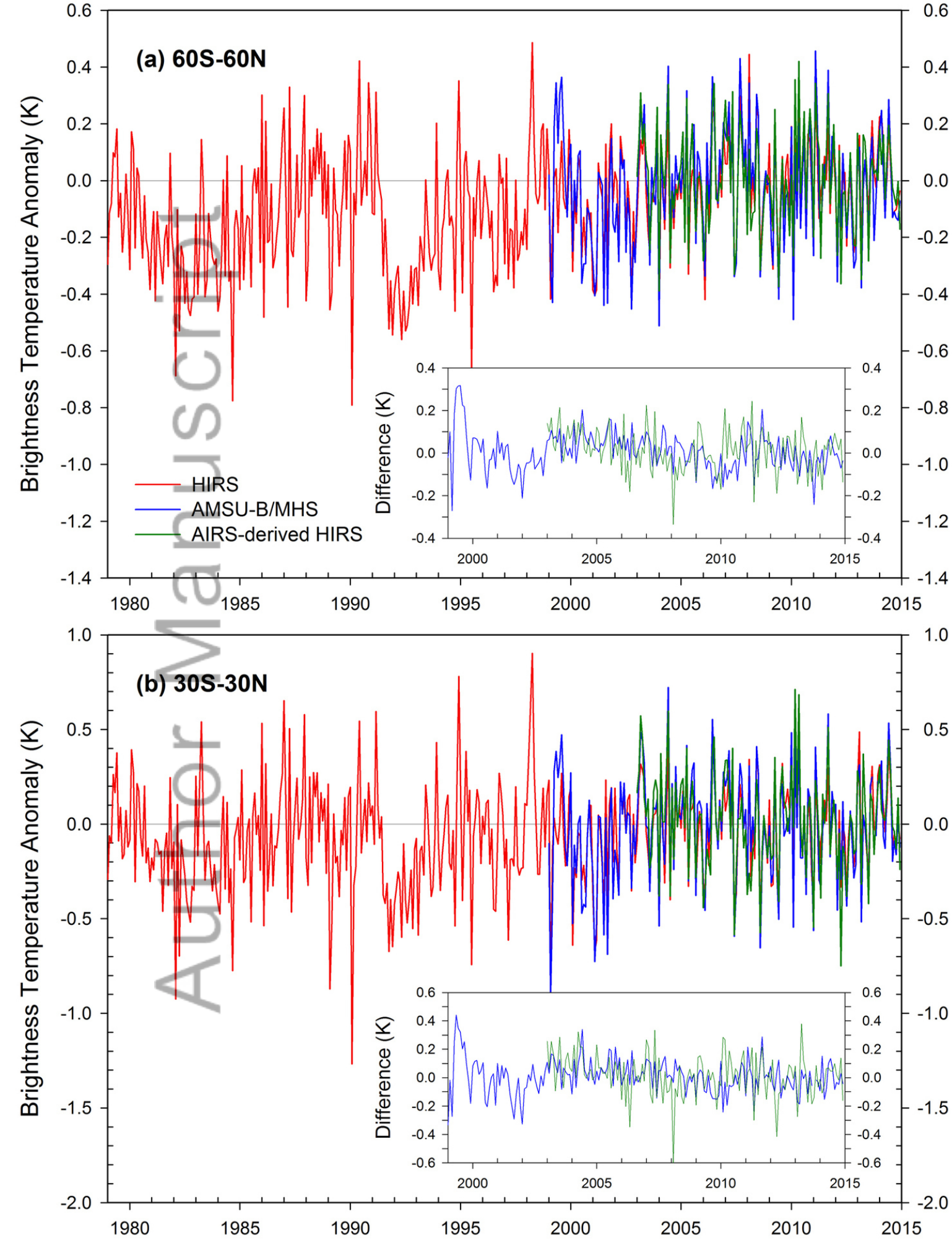
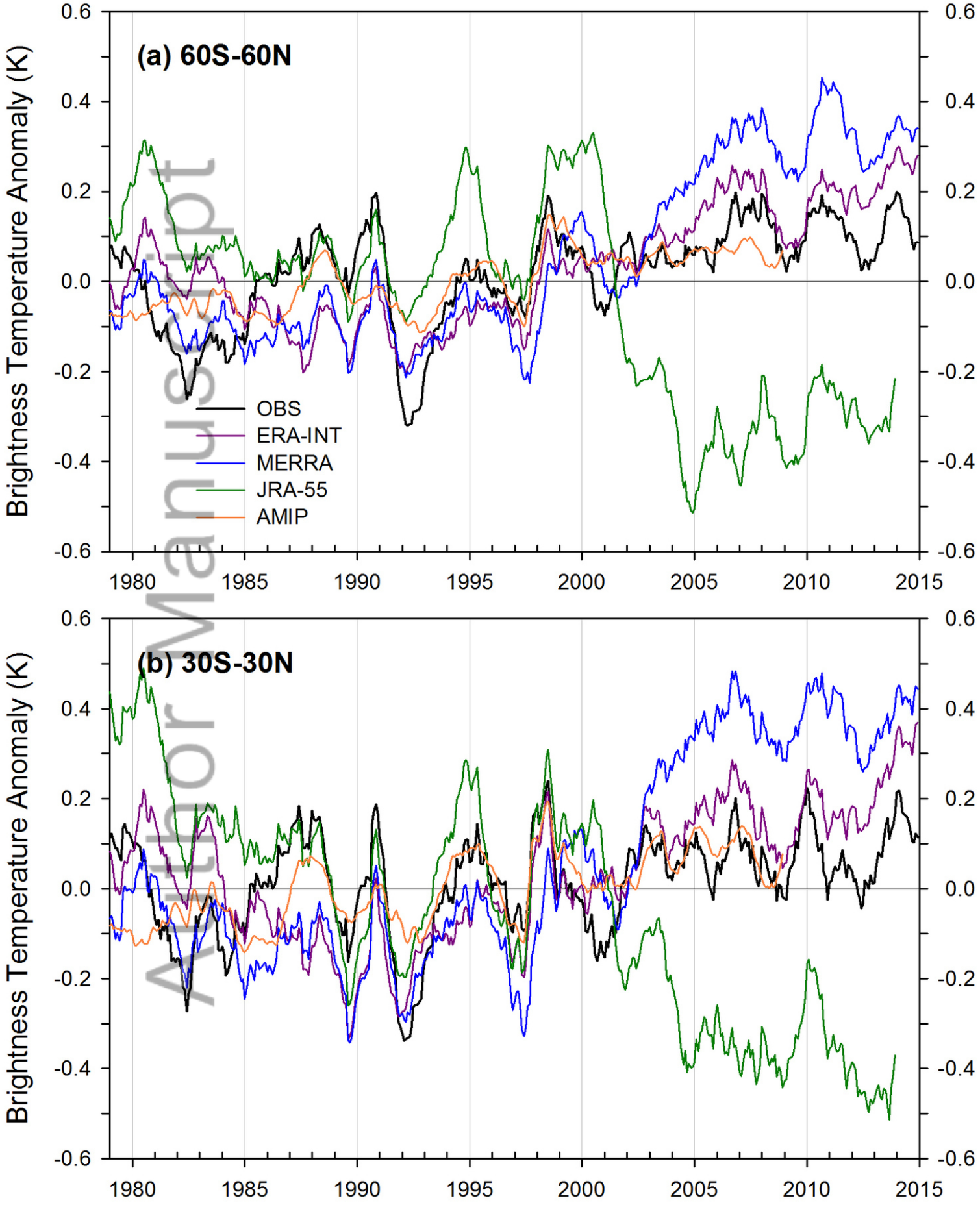
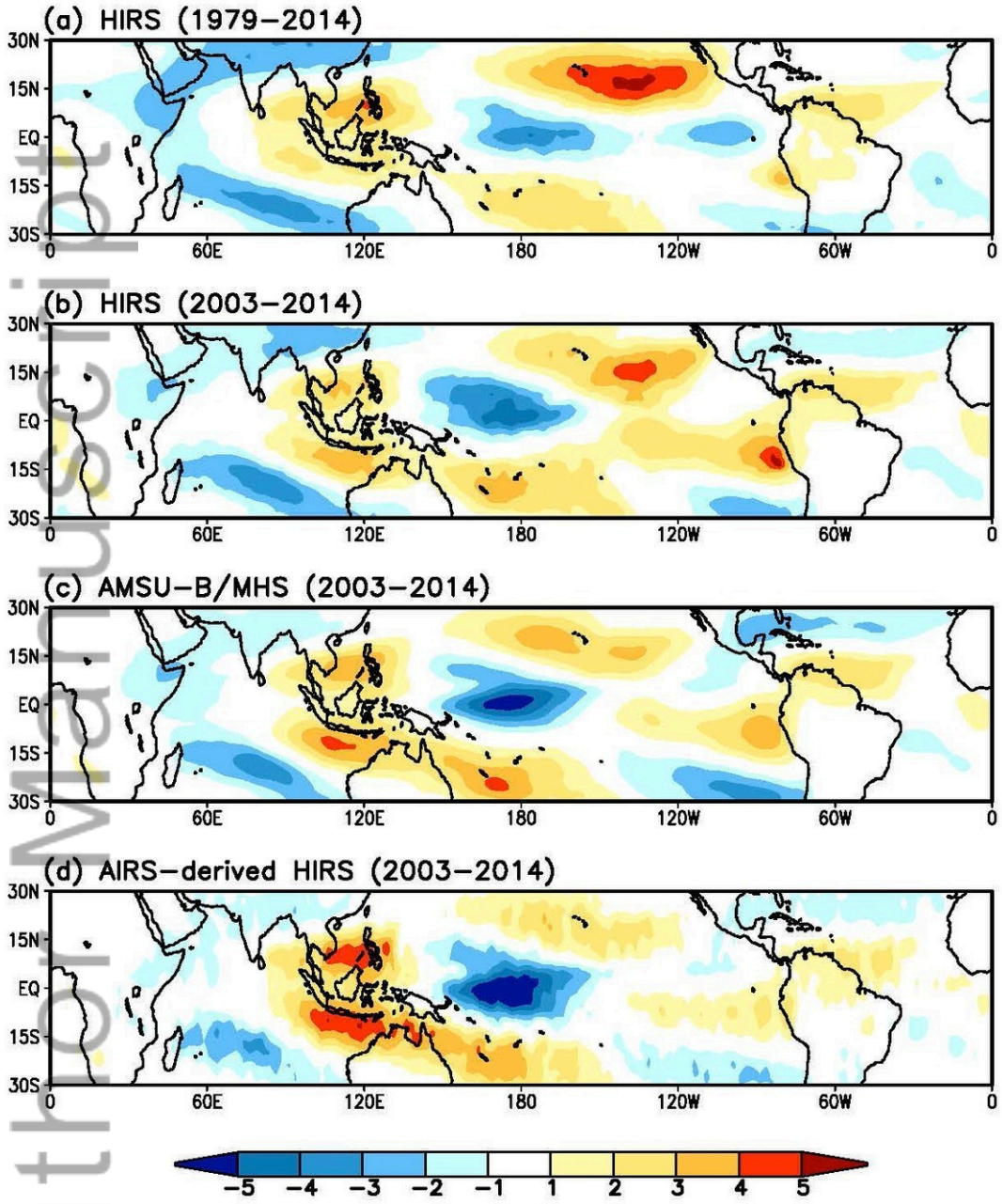


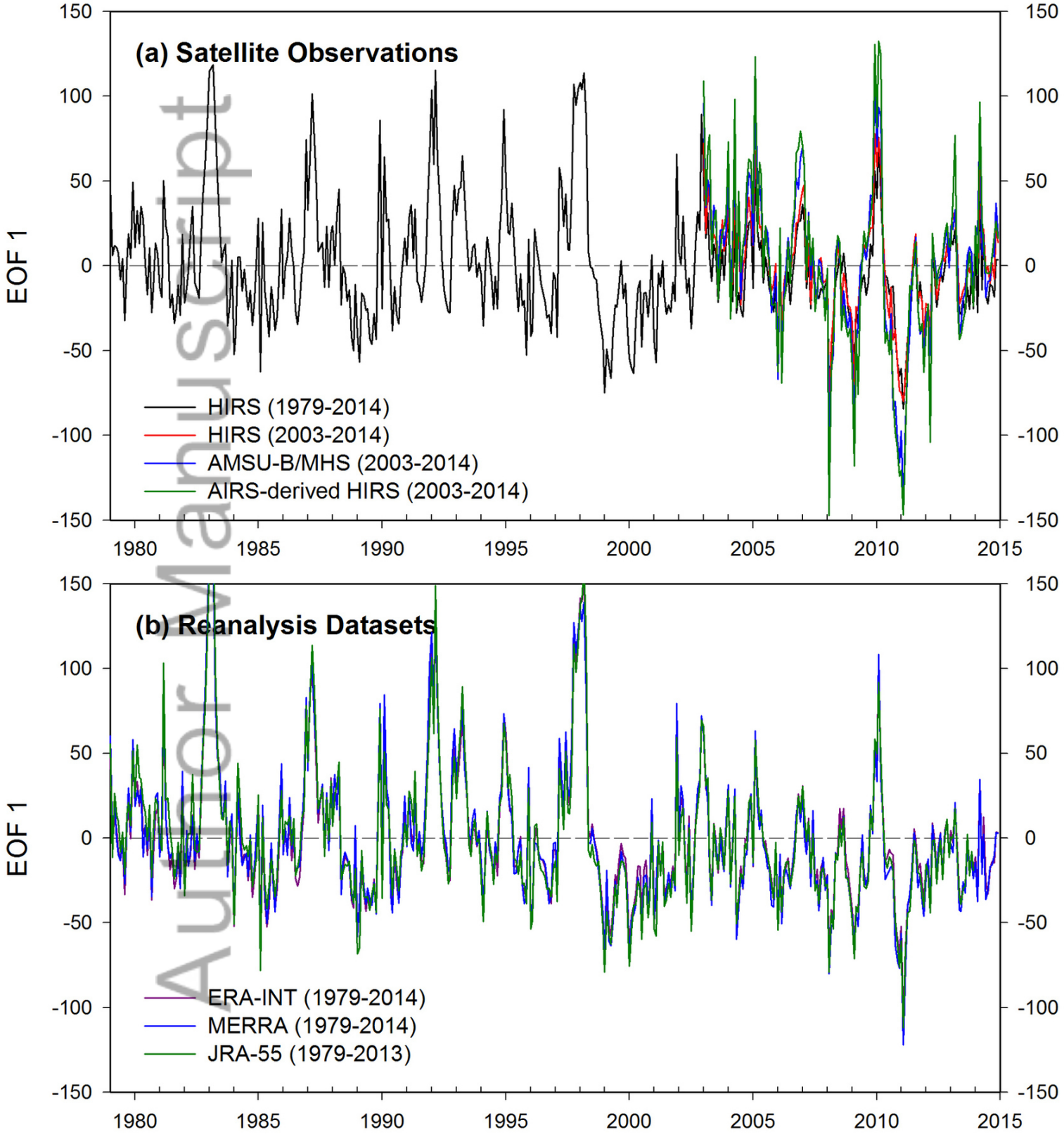
Figure 8: Zonal-mean trends of simulated HIRS channel 12 BTs from ERA-Interim (1979-2014), MERRA (1979-2014), JRA-55 (1979-2013) and climate models for CMIP5 AMIP simulations (1979-2008). The black line denotes the HIRS observations for the period 1979-2014, and the orange line represents the multi-model ensemble mean. In the case of the observations and the reanalysis data sets, the error bars denote ±2 standard errors of the linear trends. For the AMIP simulations, the error bars indicate ±2 intermodel standard deviation.



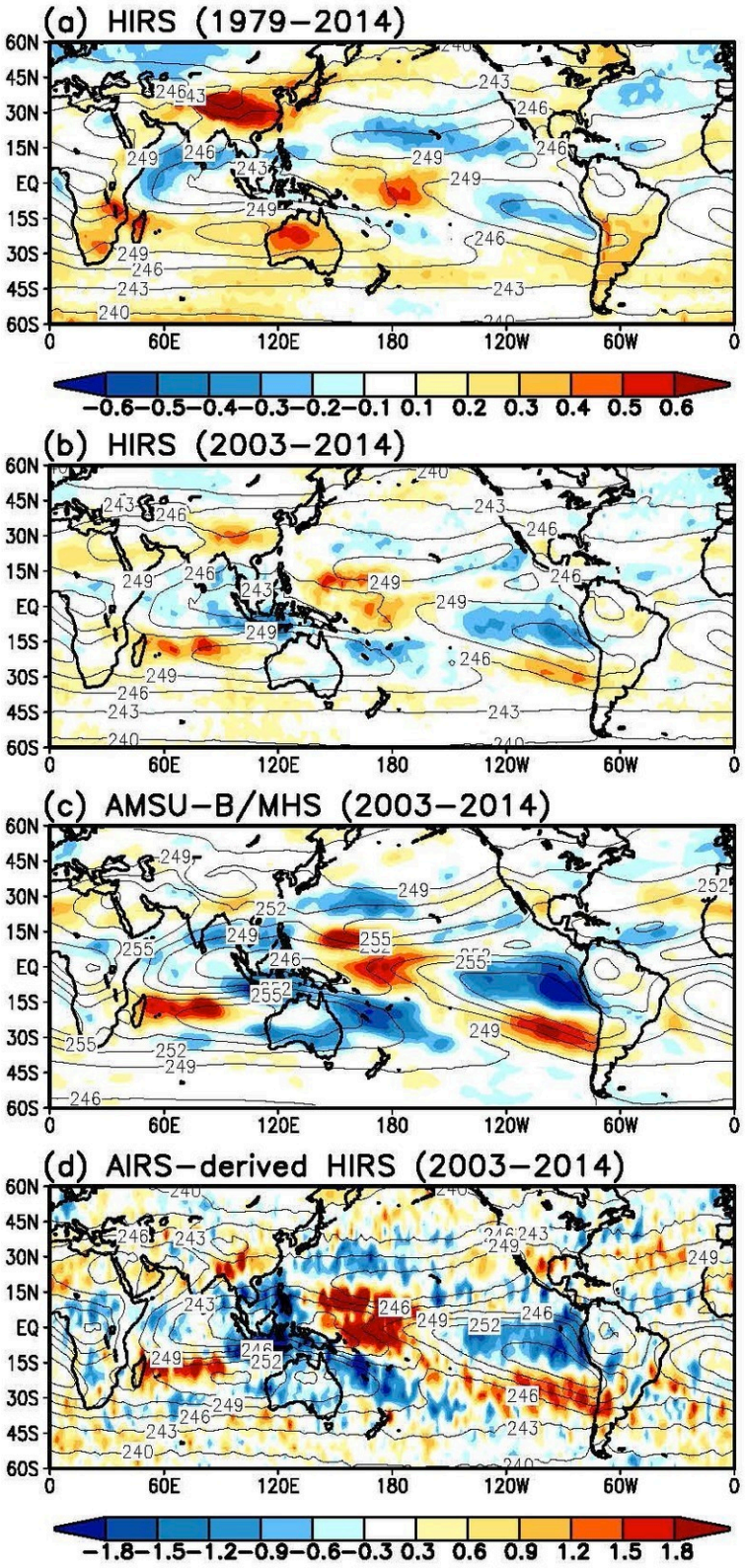




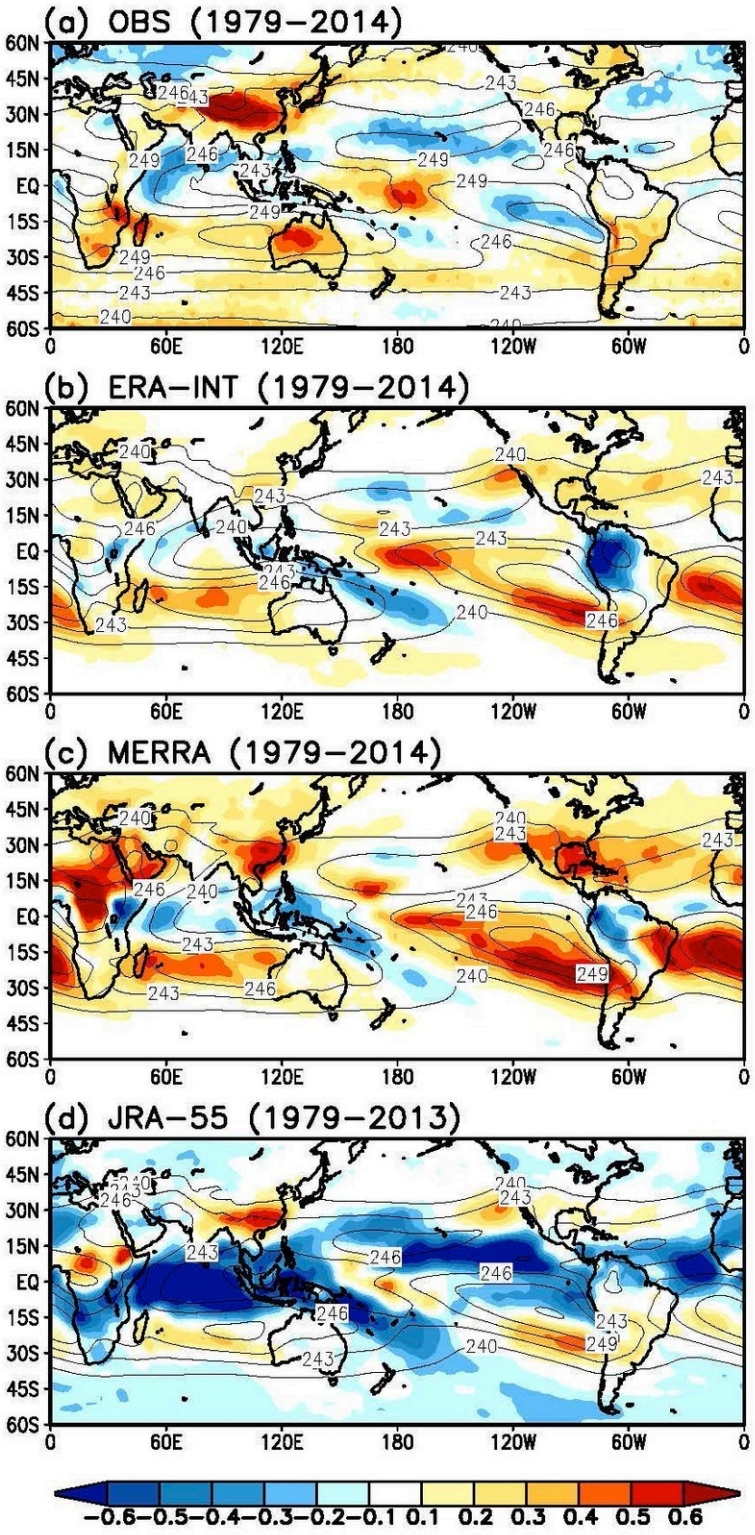
2015JD024496-f02-z-.jpg



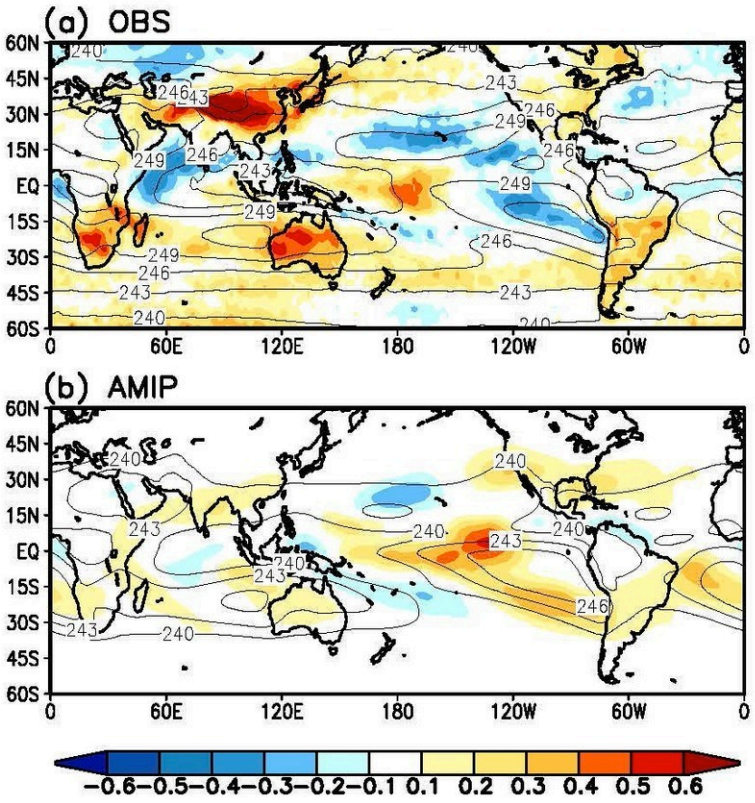
2015JD024496-f03-z-.jpg



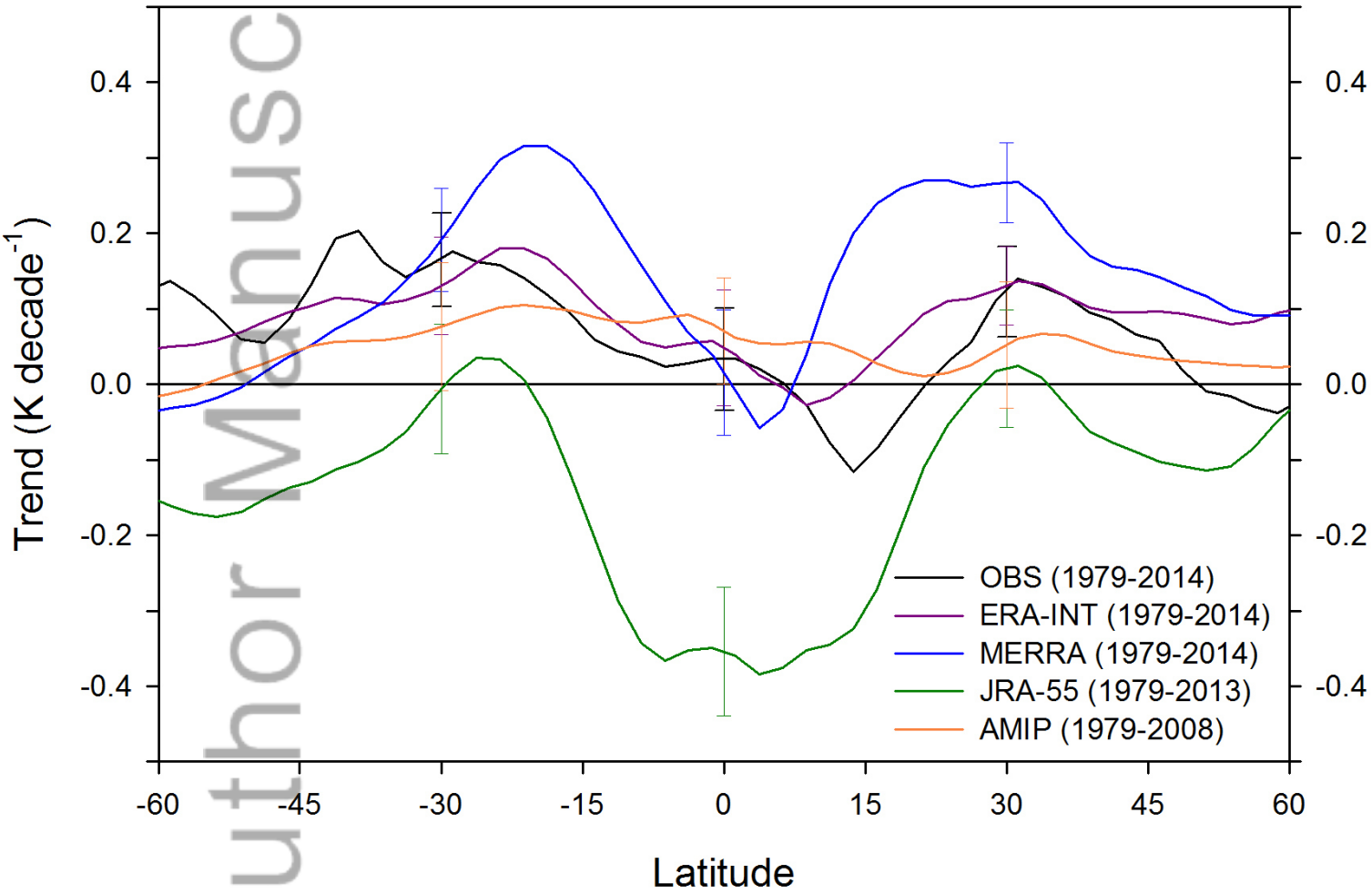
2015JD024496-f04-z-.jpg



2015JD024496-f05-z-.jpg



2015JD024496-f06-z-.jpg



2015JD024496-f07-z-.jpg