A Satellite-Derived Lower-Tropospheric Atmospheric Temperature Dataset Using an Optimized Adjustment for Diurnal Effects

CARL A. MEARS AND FRANK J. WENTZ

Remote Sensing Systems, Santa Rosa, California

(Manuscript received 25 October 2016, in final form 5 June 2017)

ABSTRACT

Temperature sounding microwave radiometers flown on polar-orbiting weather satellites provide a long-term, global-scale record of upper-atmosphere temperatures, beginning in late 1978 and continuing to the present. The focus of this paper is a lower-tropospheric temperature product constructed using measurements made by the Microwave Sounding Unit channel 2 and the Advanced Microwave Sounding Unit channel 5. The temperature weighting functions for these channels peak in the middle to upper troposphere. By using a weighted average of measurements made at different Earth incidence angles, the effective weighting function can be lowered so that it peaks in the lower troposphere. Previous versions of this dataset used general circulation model output to remove the effects of drifting local measurement time on the measured temperatures. This paper presents a method to optimize these adjustments using information from the satellite measurements themselves. The new method finds a global-mean land diurnal cycle that peaks later in the afternoon, leading to improved agreement between measurements made by co-orbiting satellites. The changes result in global-scale warming [global trend (70°S–80°N, 1979–2016) = 0.174°C decade^{-1}], ~30% larger than our previous version of the dataset [global trend (70°S–80°N, 1979–2016) = 0.134°C decade^{-1}]. This change is primarily due to the changes in the adjustment for drifting local measurement time. The new dataset shows more warming than most similar datasets constructed from satellites or radiosonde data. However, comparisons with total column water vapor over the oceans suggest that the new dataset may not show enough warming in the tropics.

1. Introduction

The temperature of the upper air has been monitored by the Microwave Sounding Unit (MSU) and Advanced Microwave Sounding Unit (AMSU) since late 1978. A total of nine MSU instruments cover the period from 1978 to 2005, followed by a series of AMSU instruments that began in mid-1998 and continue to the present. These instruments function by measuring the microwave radiance in measurement bands dominated by thermal emission from oxygen molecules. The MSU instruments use four channels at different frequencies that lead to temperature weighting profiles with peaks varying from near the surface in channel 1 to the lower stratosphere in channel 4. Of these, MSU channels 2, 3, and 4 are mostly unaffected by effects of surface emission, clouds, and water vapor. These channels are sensitive to the temperature of thick atmospheric layers centered in the middle troposphere, near the tropopause, and in the lower stratosphere. The AMSU channels 5, 7, and 9 closely correspond to the legacy MSU channels, allowing the record to continue to the present time. Both the MSU and AMSU instruments require a number of adjustments to account for calibration issues and time-varying biases before the measurements can be assembled into a long-term climate data record (CDR). These issues have been addressed by several groups, including our own, resulting in several different CDR versions for the three atmospheric layers. The MSU channel-4/AMSU channel-9 CDR for the lower stratosphere (called TLS or LS) and the MSU channel-3/AMSU channel-7 CDR near the tropopause

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: http://doi.org/10.1175/JCLI-D-16-0768.s1.

Corresponding author: Carl A. Mears, mears@remss.com

DOI: 10.1175/JCLI-D-16-0768.1

(called TTS, TP, or TUT) are provided by our group at Remote Sensing Systems (RSS) (Mears and Wentz 2009a), University of Alabama, Huntsville (UAH) (Spencer et al. 2017), and the group at NOAA Satellite Applications and Research (STAR) (Zou and Wang 2011). For the MSU channel-2/AMSU channel-5 layer in the midtroposphere, there is an additional tropics-only CDR provided by the University of Washington (Po-Chedley et al. 2015).

There is considerable interest in obtaining temperature estimates closer to the surface, where the changes more directly affect the ecosystem and human activities. By constructing a weighted average of observations made by MSU channel 2 (MSU2) and AMSU channel 5 (AMSU5) at different Earth incidence angles, the middle tropospheric measurements can be extrapolated to the lower troposphere, as first shown by the UAH group (Spencer and Christy 1992). Remote Sensing Systems produces a gridded monthly dataset from similar extrapolated measurements by intercalibrating and merging together data from the nine MSU instruments and four of the AMSU instruments. The current version of this dataset, RSS version 3.3 (V3.3) has been available and continuously updated since 2009 (Mears and Wentz 2009a). The recent update by the UAH group uses a multichannel approach to perform a similar extrapolation (Spencer et al. 2017).

We recently updated our middle tropospheric temperature product, TMT, to version 4.0 (Mears and Wentz 2016). The most important changes to the TMT dataset were the following:

1) Improvements were made to the adjustments made to account for drifting local measurement times. These drifts occur as the orbital parameters for most satellites change over each satellites mission lifetime.
2) Newer AMSU instruments were added into the dataset, and some AMSU data from instruments that appear to have degraded late in their lifetime were removed.

In this work, we extend this update to the TLT lower-tropospheric dataset to version 4.0 using methods that are very similar to those we employed for TMT. Section 2 gives a brief overview of the various steps used to construct the merged. Section 3 provides a more detailed description of the important changes made to the diurnal adjustments. These changes cause most of the difference between this new version and the previous version of the dataset. Section 4 briefly summarizes the new version of the dataset and compares the results to other sources of lower-tropospheric temperature. We discuss the results and sources of uncertainty in section 5, and section 6 describes the availability of the dataset on the Internet.

2. Overview of the intercalibration and merging procedures

Here we give a brief overview of the various adjustments applied and procedures used when we intercalibrate and combine the data into a single long-term dataset. The final product is in the form of gridded monthly means of radiance, expressed in temperature units. Many of the steps described here are identical or very similar to those we used in the earlier version (V3.3) of the TLT data, and thus are not important for the transition from V3.3 to V4.0. The major changes are to the adjustments made to account for changing local measurement times, and removal of several periods of suspect data from the *NOAA-15*, *NOAA-18*, and *Aqua* instruments. These changes result in substantial changes in the long-term trends calculated from the dataset, and are discussed in much greater detail in section 3. Figure 1 shows a flowchart that corresponds to the procedures described in this section.

The starting point of our analysis is L1B swath data. These were obtained from NOAA's Comprehensive Large Array-Data Stewardship System (CLASS) for NOAA and EUMETSAT satellites, and similar files from NASA for AMSU on *Aqua*. Table 1 shows the instrument type and the time period used for each of the satellites used in our study. We do not consider data from *NOAA-16* because it contains a large calibration drift (Zou and Wang 2011) or *NOAA-17* because its period of operation is too short to be useful.

a. Early processing steps and basic quality assurance

A number of processing steps, including adjustments to the satellite position and observation time, removal of duplicate and clearly erroneous data, and conversion from instrument counts to microwave radiance, are applied to the data as they are converted to radiance files (Mears and Wentz 2009a). For convenience, the resulting radiances are reported in temperature units.

b. Earth incidence angle

Both the MSU and AMSU instruments are cross-track scanning instruments and make observations at a wide range of observation angles. For each observation, the Earth incidence angle (EIA) depends on the observation angle (the angle between the main beam and nadir), the local radius of curvature of Earth's surface, and the height of the satellite above the surface. Because the slant path through the atmosphere depends on the EIA, the observed radiance depends on satellite height. The strongest dependence on height is for observation angles far from nadir that are used in the TLT product. Any decay in satellite altitude results in a long-term

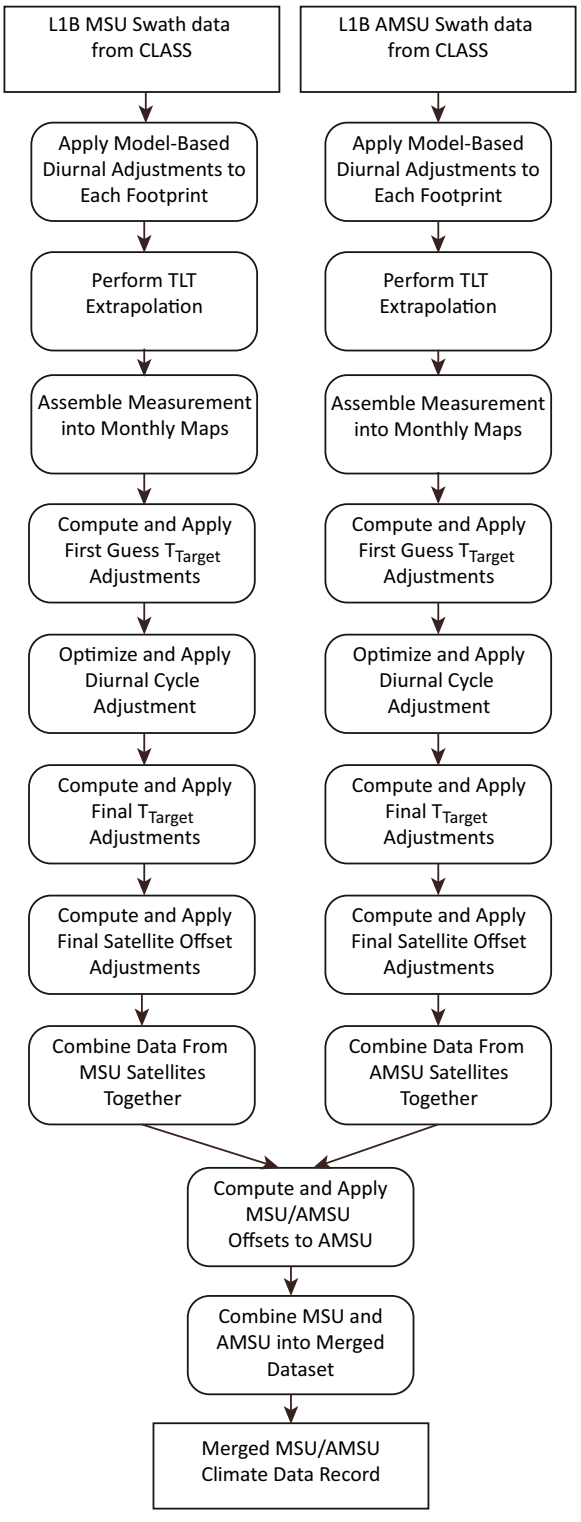

FIG. 1. Flowchart showing the various processing steps used to construct the merged MSU/AMSU TLT dataset.

changes in EIA and thus measured radiance. Because the lower-tropospheric extrapolation (TLT) (see section 3d below) is calculated from a weighted difference of measurements made at different observation angles, these changes are amplified, resulting in a much larger effect than for products based on simple averaging of near-nadir measurements. To remove the effect of orbital decay, each measurement is adjusted so that it corresponds to a constant nominal EIA for each observation angle.

This adjustment is performed using a climatology of monthly radiance at different incidence angles obtained by using a radiative transfer model (RTM) (Meissner and Wentz 2012; Wentz and Meissner 2016) to simulate the radiance as a function of the Earth incidence angle for each grid point and month of year. The input for the RTM is the output of the MERRA reanalysis over the 7-yr period from 1980 to 1986. The RTM output is used to construct a second-order Taylor expansion of radiance as a function of EIA around each nominal incidence angle. This is used to adjust the measured radiance T_{MEAS} to the radiance at the nominal EIA T_{ADJ}:

$$T_{ADJ} = T_{MEAS} + a(EIA_{ACTUAL} - EIA_{NOMINAL}) + b(EIA_{ACTUAL} - EIA_{NOMINAL})^2. \quad (1)$$

The RTM-derived parameters a and b are functions of Earth location and time of year. As a check to this procedure, we also performed this adjustment based on the NCEP reanalysis. The effect on the final global trend results was less than 0.002 K decade^{-1}. For some MSU satellites, there is also a pronounced left/right asymmetry, which is removed by fitting near-constant instrument roll angle during the calculation of the incidence angle adjustments (Mears and Wentz 2009a).

c. *Diurnal cycle*

Many of the satellites with AMSU or MSU instruments were placed in orbits with slowly drifting local equator crossing times (LECTs), which causes the local measurement times to change over time [see Fig. 2, which is adapted from Mears and Wentz (2016)]. The exceptions are NASA's *Aqua* satellite and the EUMETSAT *MetOp-A* and *MetOp-B* platforms, which are in nondrifting orbits with nearly constant LECTs maintained by orbit-keeping maneuvers. As the local measurement times drift, the diurnal evolution in the radiance caused by changes in atmospheric and surface temperature are aliased into the long-term record. It is therefore important to characterize and remove the effects of this "diurnal drift" to construct an accurate long-term record. In practice, we adjust each measurement so

TABLE 1. Satellites used in this study.

Satellite	Instrument type	Period considered in this study	Period used in V4.0 product
TIROS-N	MSU	Dec 1978–Dec 1979	Dec 1978–Dec 1979
NOAA-06	MSU	Jul 1979–Mar 1983	Jul 1979–Mar 1983
		Dec 1985–Oct 1986	Dec 1985–Oct 1986
NOAA-07	MSU	Aug 1981–Feb 1985	Aug 1981–Feb 1985
NOAA-08	MSU	Jun 1983–Aug 1985	Jun 1983–Aug 1985
NOAA-09	MSU	Jul 1985–Feb 1987	Jul 1985–Feb 1987
NOAA-10	MSU	Dec 1986–Aug 1991	Dec 1986–Aug 1991
NOAA-11	MSU	Oct 1988–Dec 1994	Oct 1988–Dec 1994
		Aug 1997–Apr 1998	Aug 1997–Apr 1998
NOAA-12	MSU	Sep 1991–Nov 1998	Sep 1991–Nov 1998
NOAA-14	MSU	Jul 1995–Dec 2004	Jul 1995–Dec 2004
NOAA-15	AMSU	Aug 1998–Dec 2014	08/1998– Dec 2010
Aqua	AMSU	Aug 2002–Dec 2012	Aug 2002– Dec 2009
NOAA-18	AMSU	Jul 2005–Dec 2016	Jul 2005– Dec 2016
MetOp-A	AMSU	Jun 2007– Dec 2016	Jun 2007– Dec 2016
NOAA-19	AMSU	Apr 2009– Dec 2016	Apr 2009– Dec 2016
MetOp-B	AMSU	Feb 2013– Dec 2016	Feb 2013– Dec 2016

that it corresponds to local midnight. Since the diurnal cycle depends on the time of year, these adjustments are also applied to the nondrifting satellites so that the difference between radiances at the measurement time and local midnight is consistently removed.

For earlier versions of the dataset, we used a model-based diurnal climatology to account for the diurnal drift before performing subsequent merge steps. Version 3.3 used a monthly averaged diurnal radiance climatology calculated using output from the Community Climate Model version 3 (CCM3) (Kiehl et al. 1996) as input to a radiative transfer model, resulting in gridded, global-scale maps of the simulated radiance for each model time step. These were averaged over the same month over 6 years (1979–84) of model output to obtain the CCM3-based climatology. Po-Chedley et al. (2015) introduced an alternative observation-based method to remove satellite diurnal drift for the MSU/AMSU temperature channels. They showed that the GCM-based method cannot completely remove the effects of diurnal drift, which was confirmed in Mears and Wentz (2016) for TMT. Our approach is to use the model-based

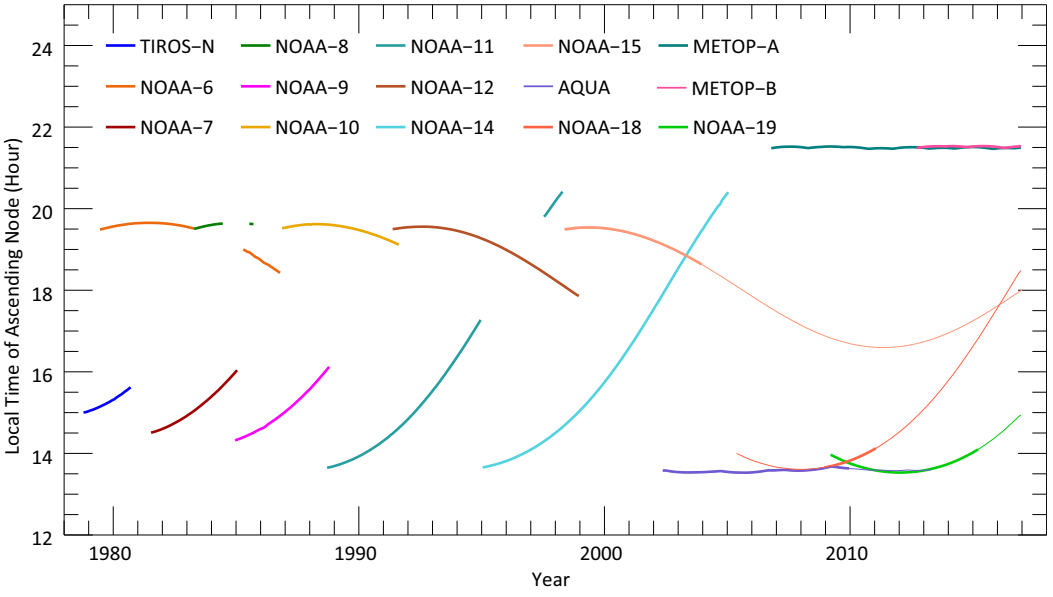

FIG. 2. Equator crossing times for each of the satellites used in this study. The thinner lines for *NOAA-15*, *-18*, and *-19* and *Aqua* denote the portions of those satellites' missions that are not used for the MIN_DRIFT method.

TABLE 2. MSU TLT weights, observation angles, and nominal Earth incidence angles.

Field of view	Observation angle	Nominal Earth incidence angle	TLT weight
1	47.35	56.19	−1.5
2	37.88	43.91	−1.5
3	28.41	32.51	2.0
4	18.94	21.51	2.0
5	9.47	10.71	0.0
6	0.0	0.0	0.0

TABLE 3. AMSU TLT weights, observation angles, and nominal Earth incidence angles.

Field of view	Observation angle	Nominal Earth incidence angle	TLT weight
1	48.333	57.224	−2.64
2	45.000	52.737	−1.14
3	41.667	48.439	0.44
4	38.333	44.274	1.41
5	35.000	40.209	1.61
6	31.667	36.219	1.17
7	28.333	32.288	0.40
8	25.000	28.403	−0.25
9	21.667	24.554	0.00
10	18.333	20.733	0.00
11	15.000	16.936	0.00
12	11.667	13.156	0.00
13	8.333	9.388	0.00
14	5.000	5.629	0.00
15	1.667	1.875	0.00

adjustments as a first guess upon which to base further optimization of the diurnal cycle, inspired by the approach of Po-Chedley et al. (see section 2f). The reason that the first-guess adjustments are applied at this point in the process is to minimize the effect of the diurnal cycle on the TLT extrapolation performed in the next step, where measurements made at different longitudes (and thus different local times) are differenced.

d. TLT extrapolation

For both MSU and AMSU, weighted averages of measurements made at different Earth incidence angles are combined to produce a retrieval that is sensitive to temperatures lower in the troposphere than the simple unweighted radiances for the same channel. The MSU instruments make measurements at 11 observation angles ranging from −47.35° to 47.35°. For the MSU instruments, the TLT extrapolation is performed by a weighted combination of the four channel-2 observations farthest from the nadir view as originally proposed by Spencer and Christy (1992). We perform the calculation separately for the two sides of the cross-track scan, yielding TLT radiances T_{RIGHT} and T_{LEFT}:

$$\begin{aligned} T_{\mathrm{LEFT}} &= 2.0(T_3 + T_4) - 1.5(T_1 + T_2), \\ T_{\mathrm{RIGHT}} &= 2.0(T_8 + T_9) - 1.5(T_{11} + T_{10}), \end{aligned} \tag{2}$$

where T_N is the radiance for the Nth field of view, numbered starting with 1 for the near-limb view on the left side of the swath. Note that the three fields of view closest to nadir are not used. For AMSU, which uses 30 fields of view with observation angles ranging from 48.33 to 48.33 degrees, we derived a set of weights so that the resulting weighting function closely matches the MSU TLT weighting function. The weights, observation angles, and nominal EIA for each field of view on the left side of the scan are shown in Table 2 (for MSU) and Table 3 (for AMSU). The derivation of these weights is discussed in more detail in Mears and Wentz (2009b). In Fig. 3 [reproduced from Mears and Wentz (2009b)] we show the temperature weighting function for both the MSU and AMSU versions of the TLT product. The small differences between the two weighting functions are not large enough to cause any detectable relative trends between the two products (see Fig. S1 in the online supplemental material). There is a small (~1 to 2 K), location-dependent absolute difference between the MSU and AMSU TLT retrievals that is removed when the MSU and AMSU data are merged together later in the processing.

In contrast to the TMT processing (Mears and Wentz 2009b), the above steps are performed on the swath data. This is because the different fields of view occur at substantially different local times as the position in the swath is varied. This, in combination with the weighted differencing step, makes it important to perform the diurnal adjustment on each observation separately before the TLT retrievals are calculated.

e. Assemble monthly maps

All of the adjustment and merging steps discussed in the following sections are performed using monthly averages gridded on a 2.5° × 2.5° longitude–latitude grid. For each half scan (field of view 1–8 or field of view 23–30 for AMSU, or field of view 1–4 or field of view 8–11 for MSU), a "TLT measurement" is the weighted average described above. This TLT measurement is then assigned to each 2.5° grid cell that contains the center of any measurement footprint that contributed to the "measurement." Then, for each grid cell, all of the TLT measurements for each month are averaged together to form an estimate of the average TLT temperature over the month.

The accuracy of these monthly gridded averages is limited by the differencing procedure. For both MSU

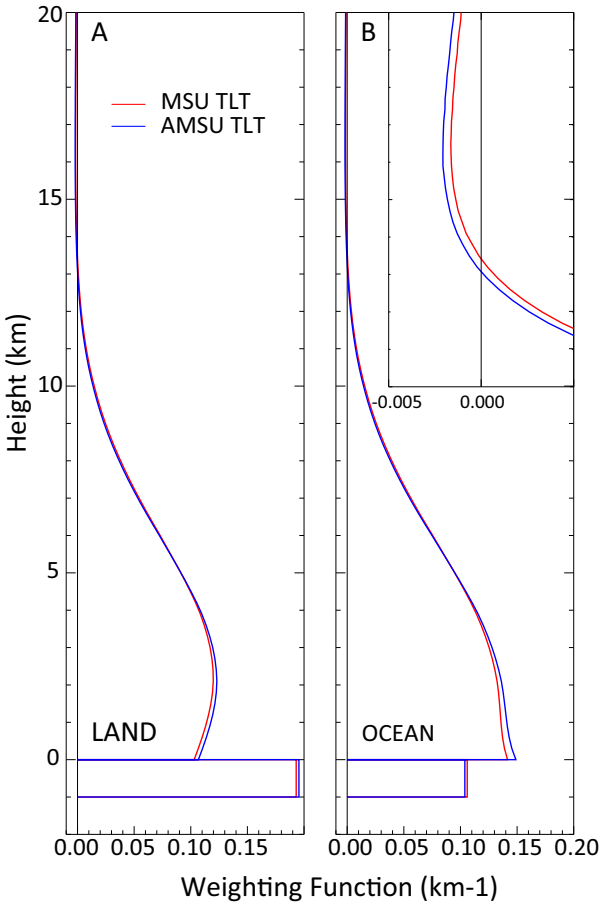

FIG. 3. Temperature weighting functions for the TLT lower-tropospheric product for the MSU and AMSU instruments. Because of the different measurement frequency and measurement geometry for the two instruments, it is not possible to match the weighting functions exactly.

and AMSU, the TLT retrieval is a weighted average of near-limb temperatures subtracted from a weighted average of temperatures measured closer to the nadir view. Since the measurements from the various views are not made at the same location on Earth, each TLT measurement is a combination of the desired vertical extrapolation and an unwanted spatial derivative of temperature along the scan direction. The methods we used to minimize (but not eliminate) these effects are discussed in Mears and Wentz (2009b). We do not anticipate that the remaining errors will contribute substantially to errors in long-term changes in TLT, since the long-term mean of the underlying spatial derivatives is unlikely to change substantially.

f. Additional diurnal adjustments

The diurnal adjustments made in section 2c are not perfect, and the errors in the diurnal adjustment are likely to be aliased into the long-term record. The main purpose of this paper is to explore methods to reduce or eliminate the effect of diurnal adjustment errors on the final product. For AMSU, there are three possible approaches to address this imperfection. This first (MIN_DRIFT) is to avoid, to the extent possible, the use of the part of each satellite's mission for which the local measurement time is drifting rapidly. For both the REF_SAT and DIUR_OPT approaches, we apply additional corrections to account for errors in the modeled diurnal adjustment. For the REF_SAT approach we adjust the data from each satellite so that it matches a reference dataset derived from nondrifting satellites. For the DIUR_OPT approach, we add a small semidiurnal adjustment to the model-derived diurnal cycle. In both cases, the adjustments depend both on the latitude and on the surface type (land or ocean). For MSU, only the DIUR_OPT approach is feasible because the measurement time of most MSU satellites drift substantially during their missions. This, combined with the small amount of overlap between different MSU satellites, does not allow us to isolate a diurnal-drift-free set of satellite measurements. These additional diurnal adjustments are discussed in detail in section 3.

g. Calibration target temperature

Global averages of simultaneous measurements made by co-orbiting MSU and AMSU instruments are known to differ by both a time-invariant intersatellite offset and an additional term that is strongly correlated with the variations in temperature of the hot calibration target (which is an integral component of the instrument and measurement technology) for each satellite (Christy et al. 2000). This correlation is likely to be related to nonlinear behavior in the radiometer electronics (Mears and Wentz 2009a). To describe these differences, we use an empirical error model for radiance incorporating the target temperature correlation (Mears and Wentz 2009a):

$$T_{\mathrm{GLOBAL,MEAS},i} = T_{\mathrm{GLOBAL,TRUE}} + A_i + \alpha_i T_{\mathrm{TARGET},i} + \varepsilon_i, \tag{3}$$

where $T_{\mathrm{GLOBAL,MEAS},i}$ is the global-mean radiance measured by the ith instrument (reported in temperature units), $T_{\mathrm{GLOBAL,TRUE}}$ is the true radiance, A_i is the temperature offset for the ith instrument, and α_i is a small multiplicative "target factor" describing the correlation of the measured radiance with the temperature anomalies of the hot calibration target, $T_{\mathrm{TARGET},i}$. The error term ε_i contains additional uncorrelated, zero-mean errors due to instrumental noise and sampling

effects. The merging parameters (the A_i and α_i terms) are found using a regression procedure that minimized intersatellite differences between monthly averages (Mears and Wentz 2009b). When considering a set of N satellites, we form a version of Eq. (3) for each coorbiting satellite/month pair, we obtain a system of equations with $2N$ unknowns (the A_i and α_i terms). The number of equations is given by the number of satellite/month pairs and is much larger than the number of unknowns. The system is solved using singular value decomposition to find best-fit values for the A_i and α_i. For the simple combination of near-nadir views used for the middle troposphere, the above procedure worked well for both the MSU and AMSU series of instruments (Mears and Wentz 2009a, 2016). For the MSU TLT regression discussed here, there are two additional steps that were needed to obtain satisfactory results.

First, as in Mears and Wentz (2009b) we find that the target temperature for *NOAA-09* is very poorly determined when using a regression procedure based solely on the system of equations obtained from Eq. (1) results in an unrealistically large value for $\alpha_{\text{NOAA-09}}$. This is due to the short overlap period between *NOAA-09* and other satellites, the noise amplification inherent in the TLT extrapolation, and the relatively small changes in the target temperature for *NOAA-09* during these overlap periods. Taken together, these factors make the determination of the *NOAA-09* target factors subject to overfitting of noise and other random fluctuations. Independent analysis by Po-Chedley and Fu (2012) showed that the use of a large target factor for *NOAA-09* causes significant contamination of the observed Earth brightness signals by the warm target temperature.

In V3.3, we used a set of nine additional regularization equations (one for each satellite in the analysis) that serve to ''pull'' the values of poorly determined target factors toward zero,

$$C\alpha_i = 0.0, \tag{4}$$

with the amount of the pull determined by the choice of the regularization factor C. These equations are appended to the original system, and the solution recomputed. For a perfect, linear radiometer, the target factor would be zero. The regularization equations function as a prior expectation that target factors are zero (i.e., the radiometer is linear), unless there is good evidence to conclude otherwise. The regularization equations have little effect on the target factors that are well determined by the original regression procedure. On the other hand, if a target factor is poorly determined (and thus subject to overfitting) its absolute value will be reduced significantly when the regularization equations are included in the system. In Mears and Wentz (2009b), we set C to 1.0 for each of the MSU satellites. This had the effect of reducing the *NOAA-09* target factor to a more reasonable value. Here we extend this work to study the effect of this procedure as a function of C so that we can make a well-informed choice for C. As C increases from zero, the fit moves away from minimizing the differences between *NOAA-09* and the two other satellites with significant overlap with *NOAA-09*, *NOAA-06*, and *NOAA-10*. With C set to 0, the standard deviation of the global monthly mean difference between *NOAA-09* and *-10* is much lower than for other satellite pairs, suggesting that overfitting may be occurring. We increase C to a point ($C = 1.5$) where the standard deviation of this difference is comparable to other satellite pairs (see Fig. S2). This also reduces the values of the *NOAA-09* target factor to a value more comparable to the target factors from other satellites (see Fig. S3). We choose $C = 1.5$ for the subsequent analysis in this paper, but other nearby values may also be valid. The value of C has an effect on the long-term trends in the final results, and thus the uncertainty surrounding the choice of C contributes to the uncertainty in the final results (see Fig. S4 and the discussion in section 5).

Second, we find that when only MSU measurements are used to determine the MSU target factors, the target factor for *NOAA-14* is uncertain enough that it can dominate the uncertainty in the part of the MSU record after 1999, when *NOAA-14* is the only MSU instrument operating. This occurs because the target temperature fluctuations during the period that *NOAA-14* overlaps other MSU satellites (*NOAA-12* and *-11*) are much smaller than the fluctuations that occur after the end of the *NOAA-12* mission (see Fig. S5). For TLT, this, in combination with the larger noise in the TLT global time series, leads to errors in the *NOAA-14* target factor. These errors, which do not have much effect on the *NOAA-14* minus *NOAA-12* and *NOAA-14* minus *NOAA-11* differences, can lead to large errors in the adjusted radiances after 1999 when the target temperature fluctuations are larger. To address this problem, we include merged measurements from AMSU in the regression procedure that determines the MSU target factors, thus including information from the period of large fluctuations in the *NOAA-14* target temperature. Since we are using the merged AMSU data, the target temperature effects have already been removed, and the target factor for the AMSU data can be set to zero. This additional information further constrains the *NOAA-14* target factor, which can be seen both by

noting that the *NOAA-14* target factor is less sensitive to the changes in the regularization factor (see Fig. S3) and by examining the covariance matrix from the regression (not shown).

h. Intersatellite offsets

Once the α_i are determined, we then find the latitude-dependent offsets using a regression procedure for each 2.5° wide latitude band. In contrast to V3.3, where a single offset was found for each latitude, we now find latitude-dependent offsets for land and ocean surface types separately. To reduce the effects of sampling noise on these offsets, the offsets for a given 2.5° latitude band are found using data from the 12.5° latitude band surrounding the band in question, as was done for TMT V4.0 (Mears and Wentz 2016).

i. Matching AMSU TLT to MSU TLT

The MSU and AMSU TLT weighting functions differ slightly, leading to differences in radiance that depend on the atmospheric profile and surface type. We use an empirical adjustment found by analyzing MSU–AMSU differences during the overlap period to adjust AMSU radiances so that they correspond to MSU radiances as we discuss in (Mears and Wentz 2009a). The adjustment is a function of the location of grid cells and month of year but does not contain a long-term trend, so it cannot mask errors due to calibration errors or diurnal drifts.

3. Optimization of the diurnal adjustments

The optimization procedure described here closely parallels the approach we used in Mears and Wentz (2016). The text in the following introductory paragraphs and sections 3a through 3g is adapted from this previous work with changes that reflect the different atmospheric layer studied and the quantitively different results obtained.

a. Model-based diurnal adjustments

An accurate adjustment to remove the effects of changing measurement time is critical for the construction of a climate data record with accurate long-term changes. In addition to the CCM3-based diurnal adjustment used in V3.3, we have also investigated the use of other diurnal climatologies and their effects on the final results. These additional climatologies were derived by using output from the Hadley Center Global Environmental Model version 1 (HadGEM1; Johns et al. 2006; Martin et al. 2006) and NASA's Modern-Era Retrospective Analysis for Research and Applications (MERRA; Rienecker et al. 2011) as inputs to the RTM. The differences between the results from different diurnal adjustments led us to conclude that for tropospheric channels, the diurnal adjustment was the dominant source of uncertainty for interannual time scales, including multidecadal trends (Mears et al. 2011). We also investigated the case when no model-based diurnal adjustments were applied (NONE).

If the diurnal adjustment is accurate (and assuming other adjustments are also correct and there are no spurious drifts), then the difference time series between large-scale averages of co-orbiting satellites with different local measurement times should be free of structure. We evaluate the accuracy of any applied diurnal adjustment by plotting globally averaged annual-mean intersatellite differences for pairs of co-orbiting satellites separately for land and ocean scenes. These differences are calculated after the application of latitude-dependent intersatellite differences (section 2g) and differences that depend on the warm calibration target temperatures (section 2f) have been removed. Ideally, if all the adjustments applied were accurate, the satellite measurements should closely match each other, and these differences (and trends in these differences) should be very close to zero. We plot land and ocean average separately because any errors in the diurnal adjustment are typically larger over land, leading to important differences between land and ocean scenes. Intersatellite differences that are roughly equal in land and ocean scenes are unlikely to arise from diurnal adjustment errors.

In Fig. 4, we show this type of plot for near-global (60°S–60°N) AMSU TLT averages for a number of AMSU satellite pairs. For most of the pairs, the second satellite in the difference is chosen to be either *Aqua* or *MetOp-A*. These satellites were chosen because they are in controlled orbits with insignificant changes in local observations time. Thus, any changes in the differences caused by the diurnal cycle are due to observing time changes for the first satellite. This makes it easier to evaluate whether or not the observed differences could be caused by changes in the measurement time.

The top row shows the results when no diurnal adjustment is applied (NONE). In this case, large intersatellite differences are present for both land and ocean scenes. The intersatellite differences for land are substantially larger than those for ocean as we might expect as a consequence of the uncorrected diurnal drift. The most rapidly changing land differences tend to involve satellites (*NOAA-15* and *-18*) that drift rapidly in measurement time, further indicating that the land differences are influenced by the diurnal cycle. Note that the vertical axes for both the difference and difference trend plots are extended relative to those shown for TMT in Mears and Wentz (2016) (Figs. 3 and 6 in that paper).

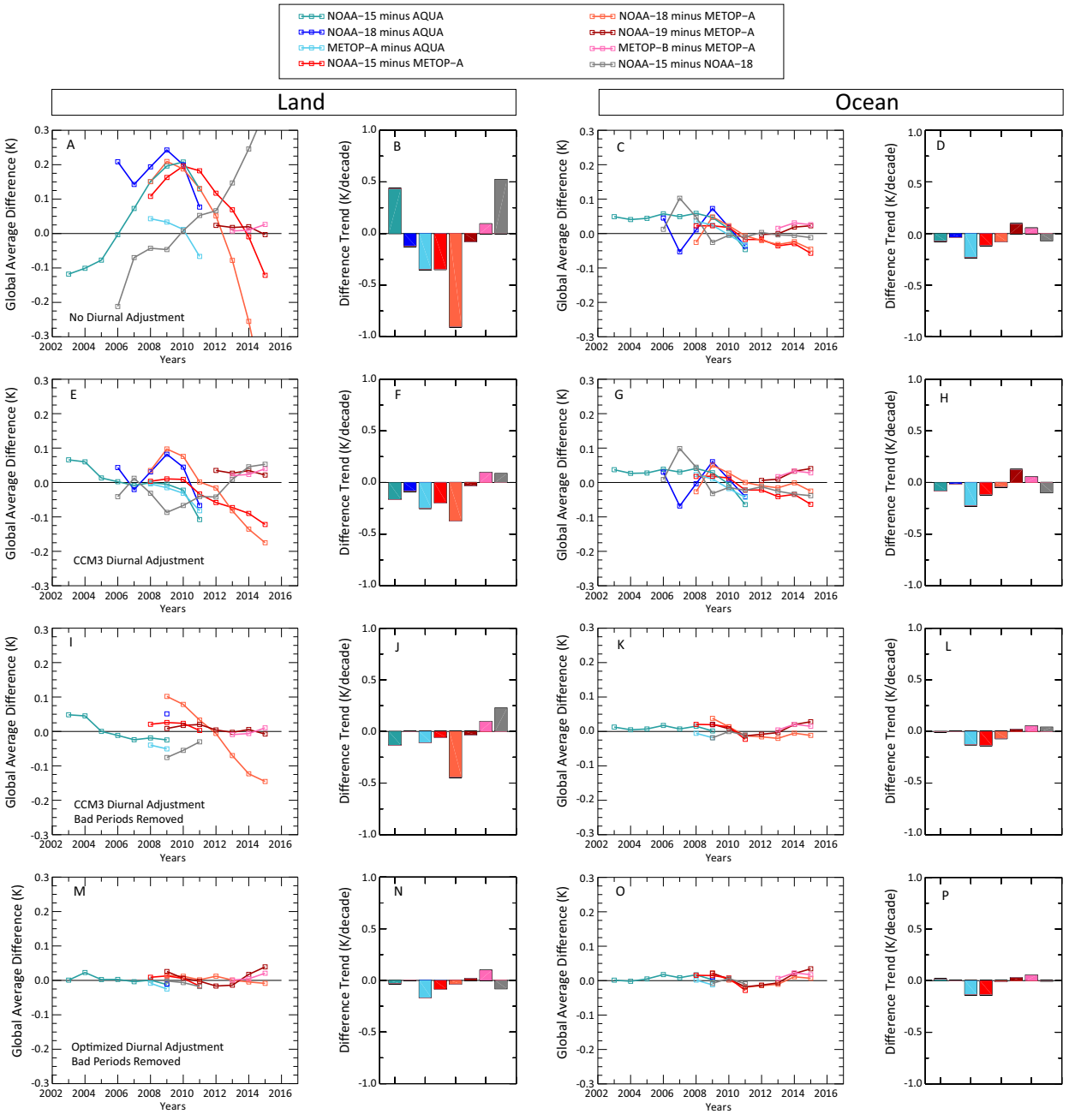

FIG. 4. Differences of lower-tropospheric (TLT) yearly near-global (60°S–60°N) means between co-orbiting AMSU instruments. Each row of plots shows the differences after different levels of diurnal adjustment and data exclusion applied. Both yearly difference time series and a bar plot representing trend differences are shown for each case. The left column shows land-only results, and the right column shows ocean-only results. Much of the remaining structure in the differences in the bottom row is the same for land and ocean, indicating that these differences are likely unrelated to the diurnal cycle.

This is because TLT is more sensitive to the surface and lower troposphere, where the diurnal cycle is larger and thus the differences caused by diurnal drift are larger.

When we apply the CCM3-derived diurnal adjustments (Figs. 4e–h), the land trend differences are much reduced, showing that this adjustment is partially successful. The remaining differences, while more similar to the ocean differences, are still larger over land than over ocean (especially for *NOAA-18* minus *MetOp-A* after 2010, when *NOAA-18* drifts rapidly in measurement time), indicating that the CCM3 diurnal adjustments are not perfect.

b. Removing spurious data periods

Before we proceed further, we try to determine if there are any anomalous changes in the satellite measurements that are not due to the diurnal cycle, because we do not want spurious changes in satellite calibration to influence the derivation of improvements to the diurnal adjustment that will be performed in the next section. To search for such spurious changes, we focus on features that are present in the ocean-only results, which are unlikely to be caused by problems with the diurnal adjustment, and also those that occur when three or more satellites are operating. The presence of three satellites allows the identification of the ''odd one out,'' which is the most likely candidate for a calibration problem. The most dramatic difference in the ocean data is the difference between *NOAA-18* and *Aqua* and *NOAA-15* during 2007, suggesting that *NOAA-18* shows anomalous behavior during this period. We confirmed this by making a similar plot with monthly instead of annual means (see Fig. S6). To avoid this drift, we do not use *NOAA-18* data before 2009. Note that this problem was not found for TMT, and thus occurs only in the near-limb views that are used for TLT but not for TMT. Second, the differences that involve *Aqua* all trend to large negative values starting in 2010. This is most easily seen in the *NOAA-15* minus *Aqua* and *MetOp-A* minus *Aqua* differences since these satellites are more stable during this period than *NOAA-18* (as discussed above) During 2010–11 *Aqua* also shows increasing scan-to-scan noise indicating that this channel on *Aqua* is beginning to degrade [R. Spencer 2012, personal communication; also see Fig. S1 in Mears and Wentz (2016)]. Given the evidence for substantial instrument calibration drift to higher radiance, we exclude *Aqua* data after December 2009, as we did for TMT. We also exclude *NOAA-15* after 2011 (as we did for TMT) because *NOAA-15* data appear to be cooling relative to both *MetOp-A* and *NOAA-18* in the ocean-only data.

The results after removing these time periods (and redoing all steps to this point, including derivation of the target factors) are plotted in the third row in Fig. 4 (Figs. 4i–l). It is clear that there is still a large error in the diurnal adjustment over land from the large trend in the *NOAA-18* minus *MetOp-A* curve, and to a lesser degree the slopes in the differences involving *NOAA-15*.

To study further improvements to the CCM3 diurnal cycle adjustment, we apply methods similar to those we used for the TMT data in Mears and Wentz (2016). These are the MIN_DRIFT, REF_SET, and DIUR_OPT methods briefly introduced in section 2f. Note that as was the case for TMT, only the DIUR_OPT method can be used for MSU. For this reason, we view the MIN_DRIFT and REF_SAT approaches as consistency checks for the DIUR_OPT results for AMSU. We will choose DIUR_OPT to be the final method for both AMSU and MSU. The following discussion closely parallels the discussion in Mears and Wentz (2016) and is repeated here to provide a coherent narrative.

c. Using AMSU data with the minimum amount of diurnal drift (MIN_DRIFT)

For the MIN_DRIFT approach, we try to construct a dataset that uses (to the extent possible) only those parts of the AMSU satellite record that do not have large drifts in measurement time. This ''minimal drift'' analysis reduces (but does not eliminate) the sensitivity of the final results to measurement time drift and errors in the diurnal adjustment. The largest challenge to this approach is the *NOAA-15–Aqua* overlap. High-quality data from *Aqua* do not begin until September 2002. The *NOAA-15* measurement time has already begun to drift substantially by the end of 2001 and is close to its maximum drift rate by late 2002. This suggests that we should stop using *NOAA-15* data as soon as possible. But it is also important to use a relatively long overlap with *Aqua* to be able to accurately calculate intersatellite offsets and target factors. We investigate the effects of using four different cutoff months for *NOAA-15* ranging from June 2003 to December 2004 (see Table S1). We found that the global mean trend was almost insensitive to the cutoff month, with changes $\leq$ 0.005 K decade^{-1} over the AMSU period. We choose to perform further analysis using December 2003 as the cutoff time, yielding 16 months of overlap time. We also exclude *NOAA-18* after December 2011 and *NOAA-19* after March 2015, thus removing the periods of rapid measurement time drift after these points. The excluded portions are shown by the thinner lines in Fig. 2. The 1999–2016, near-global AMSU-only trends that result from this analysis for each of the starting diurnal adjustments are shown in Table 4 (and are plotted in Fig. S7). The use of the MIN_DRIFT brings most of the global results from the different starting diurnal adjustments into much better agreement than the case where all satellite months are used. The exception is the land-only results from the NONE case, where the MIN_DRIFT approach greatly decreases the trend, increasing its difference from the mean of the others.

d. Using Aqua and MetOp-A as a drift-free reference (REF_SAT)

For the REF_SAT approach, we combine the results of two of the nondrifting satellites, *Aqua* and *MetOp-A*, to form a reference dataset free from the effects of

TABLE 4. AMSU-only global (70°S–80°N) trends (1999–2016).

Diurnal model	Region	Model only, all satellites	Model only, data edit	MIN_DRIFT	REF_SAT	DIUR_OPT
None	Ocean	0.178	0.172	0.179	0.213	0.175
	Land	0.242	0.200	0.143	0.167	0.136
	Land + ocean	0.198	0.181	0.167	0.198	0.163
CCM3	Ocean	0.180	0.179	0.188	0.217	0.181
	Land	0.090	0.079	0.171	0.202	0.164
	Land + ocean	0.151	0.147	0.182	0.212	0.176
HadGEM	Ocean	0.179	0.180	0.190	0.212	0.176
	Land	0.232	0.299	0.210	0.240	0.157
	Land + ocean	0.196	0.219	0.196	0.220	0.170
MERRA	Ocean	0.193	0.194	0.192	0.217	0.176
	Land	0.299	0.340	0.210	0.237	0.160
	Land + ocean	0.228	0.242	0.198	0.224	0.171

diurnal drift. The other satellites are then adjusted to match this reference dataset thus removing the effects of diurnal drift from the final results. While *Aqua* and *MetOp-A* do not drift in measurement time, they do make observations seven hours apart. This difference, coupled with the seasonal modulation of the diurnal cycle, leads to a seasonally varying difference between the two satellites' measurements. We remove this difference by calculating a mean monthly *Aqua* minus *MetOp-A* difference for each 2.5° latitude band, separately for land and ocean, and then adding this difference to *MetOp-A* so that the mean *MetOp-A* observations for each month of the year correspond to the mean *Aqua* observations for the same month. The *MetOp-A* and *Aqua* measurements are then combined to yield a combined reference dataset extending from September 2002 to December 2015.

The difference between each of the other satellites and this reference dataset is then calculated and used to adjust these satellites to correspond to the reference dataset. This will have the effect of removing both diurnal drifts due to changing measurement time and any other drifts that may be present, as well as eliminating the contribution of information from the diurnally drifting satellites to the long-term changes. The measurements from the adjusted satellites will still serve to reduce sampling noise for individual grid points. The adjustment is performed by calculating the monthly zonal mean differences as a function of latitude for each satellite, separately for land and ocean scenes. These are then smoothed both in the north–south direction and in the time direction using a mean-of-three ''boxcar'' smooth. An important disadvantage of the REF_SAT approach is that the diurnal-drift-free reference does not begin until the start of the *Aqua* data in September 2002, and thus no differences for *NOAA-15* can be calculated before this time. We fill in the differences for this earlier time period by repeating the smoothed differences from September 2002 to August 2003 backward in time to extend the differences over the *NOAA-15* mission. This will clearly lead to some level of error, since the *NOAA-15* measurement time does drift during this earlier period, but the approach will likely lead to an improvement over performing no adjustments, especially for the cases where much of the diurnal cycle effects has already been removed by the model-based method.

Figure 5 shows global averages of the original CCM3-modeled diurnal adjustments and diurnal adjustments derived using the reference satellite method (using CCM3 as an initial diurnal adjustment) for each satellite. (Fig. 5 also shows the adjustments found by optimizing the model diurnal cycles, which is discussed in the next section.) For *NOAA-15*, the adjustments derived by the REF_SAT method are larger than those derived from CCM3, while the opposite is true from *NOAA-18*. Adjustments for the other satellites that either are in controlled orbits or have not yet undergone a large drift in measurement time show little difference. The 1999–2015, near-global, AMSU-only trends that result from this analysis for each of the starting points are shown in Table 4 and plotted in Fig. S7. The results are qualitatively similar to the ''minimal drift'' case. The global (land and ocean) trend results from the three REF_SAT cases where a model-based adjustment is used as a starting point agree with each other to within 0.012 K decade^{-1}, but are substantially larger than either the MIN_DRIFT results or the final results obtained in the next section, particularly over land.

e. Optimizing the applied diurnal cycle (DIUR_OPT)

For the DIUR_OPT approach, we optimize the model-based diurnal cycles so that they more effectively remove the intersatellite differences by adding a seasonally dependent second harmonic perturbation. The second harmonic diurnal amplitudes (terms in the

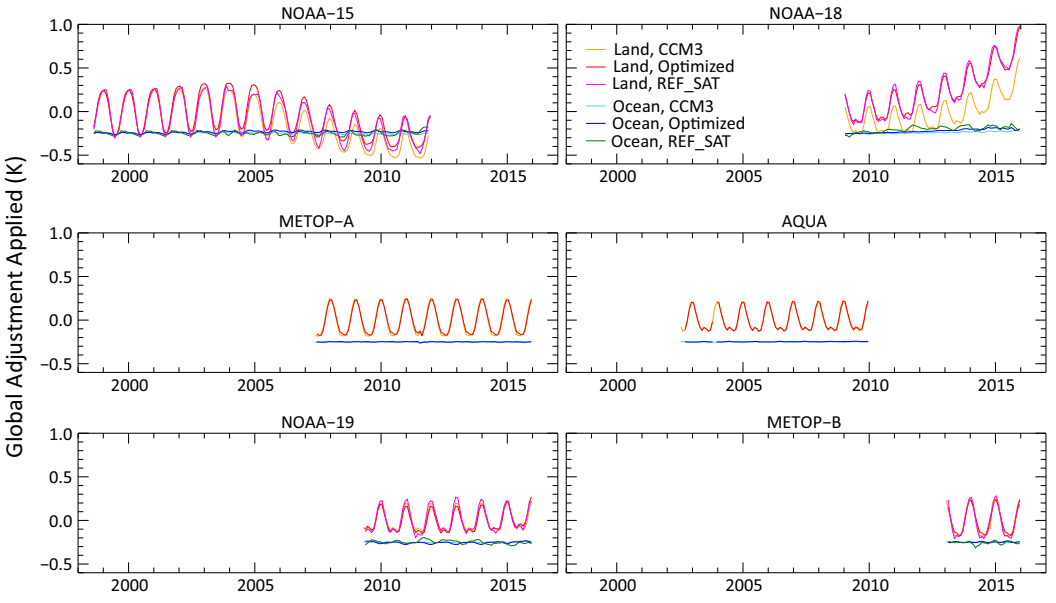

FIG. 5. Global-mean adjustments applied to each satellite using the original CCM3-based adjustments and the optimized adjustment (DIUR_OPT) and the reference satellite (REF_SAT) adjustments.

square brackets) are allowed to vary sinusoidally with season.

$$T_{\text{opt}} = T_{\text{adj}} + c_i + [a_0 + a_1 \sin(2\pi m/12) + a_2 \cos(2\pi m/12)] \sin(2\pi t_{\text{asc}}/12) + [b_0 + b_1 \sin(2\pi m/12) + b_2 \cos(2\pi m/12)] \cos(2\pi t_{\text{asc}}/12). \quad (5)$$

Here m is the month of the year, and t_{asc} is the ascending node local equator crossing time in hours, which is used as a proxy for the local measurement time for the entire orbit. By evaluating differences between measurements made by difference satellites at different local times, but for the same month, we can deduce values for the a and b terms (which are independent of satellite) and the offsets c_i. The difference between the ith and jth satellite is given by

$$T_{\text{opt},i} - T_{\text{opt},j} = T_{\text{adj},i} - T_{\text{adj},j} + a_0[\sin(2\pi t_i/12) - \sin(2\pi t_j/12)] + a_1 \sin(2\pi m/12)[\sin(2\pi t_i/12) - \sin(2\pi t_j/12)] + a_2 \cos(2\pi m/12)[\sin(2\pi t_i/12) - \sin(2\pi t_j/12)] + b_0[\cos(2\pi t_i/12) - \cos(2\pi t_j/12)] + b_1 \sin(2\pi m/12)[\cos(2\pi t_i/12) - \cos(2\pi t_j/12)] + b_2 \cos(2\pi m/12)[\cos(2\pi t_i/12) - \cos(2\pi t_j/12)] + c_n - c_m. \quad (6)$$

By forming a version of Eq. (4) for each co-orbiting satellite–month pair, we arrive at a system of equations with the number of unknowns given by 6 (the number of a's and b's) plus the number of satellites considered (the c_n terms). The number of equations, which is much larger than the number of unknowns, is given by the number of satellite–month pairs. This overdetermined system is solved using singular value decomposition. The equations are solved separately for land and ocean scenes and for each 2.5° latitude bin. In practice, the analysis for a given latitude is performed using measurements from the 12.5° wide (5° latitude bin) latitude band centered at the latitude in question to reduce random sampling noise and provide values for the a's and b's that vary smoothly with latitude. The values of the c_n terms are not stored because they will be recomputed when the intersatellite offsets are computed (section 2h). When combined with the background diurnal adjustment, this procedure results in an optimized diurnal adjustment, which is then used to adjust each measurement so that it corresponds to local midnight. This approach is identical to that used in Mears and Wentz (2016) for TMT. When applied to the TLT measurements, the results are qualitatively similar, but differ quantitatively because of the different atmospheric layer

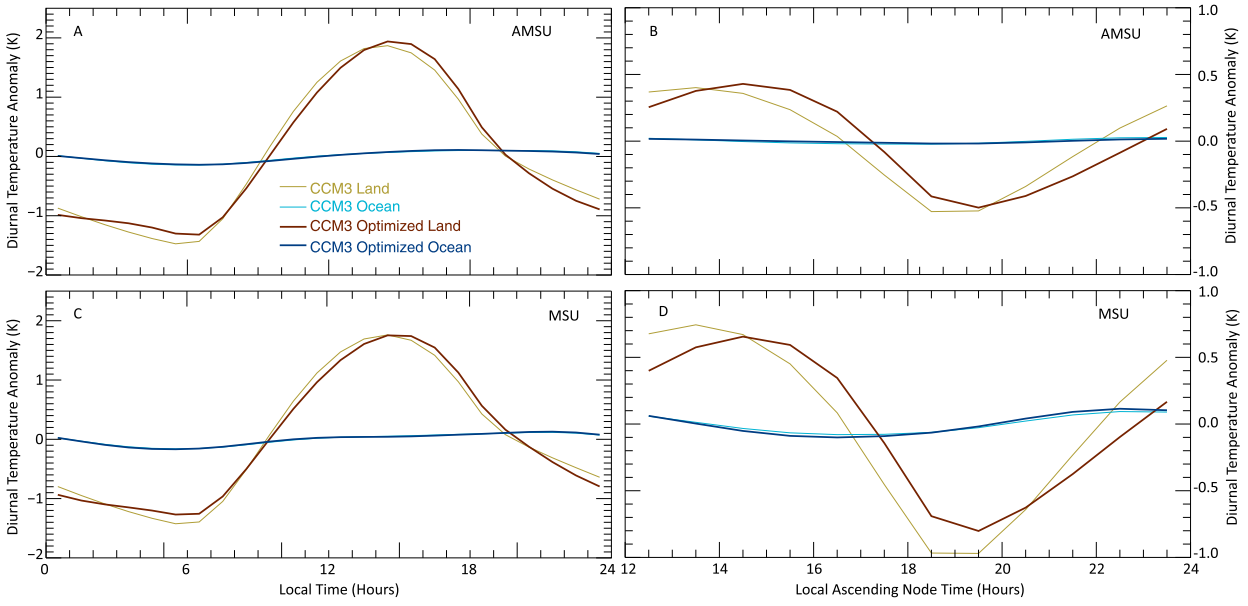

FIG. 6. (a),(c) Near-global mean (60°S to 60°N) diurnal cycles from CCM3 for AMSU TLT and MSU TLT, before and after optimization using satellite measurements. (b),(d) To emphasize the second harmonic of the diurnal cycle which dominates the adjustments, we show the sum of the morning and afternoon parts of the diurnal cycle, before and after the optimization procedure.

studied. The original and optimized diurnal cycles are plotted in Fig. 6.

We plot the annual mean intersatellite differences and trend differences after the DIUR_OPT procedure in the fourth row of Figs. 4m–p. Most of the intersatellite differences are quite small for both land and ocean. There remains some coherent structure in the differences after 2010, but it is not possible to determine which satellite is responsible. These differences are too small to affect the final results substantially.

We repeated the above procedure using other diurnal adjustments as the starting point, including NONE (no adjustment applied), and the HadGEM1 and MERRA diurnal adjustments discussed above. In Table 4, we summarize the near-global (60°S–60°N) trends in the merged AMSU data for the 1999–2015 period for both the optimized and nonoptimized diurnal cycles. After the optimization procedure, the spread in trends is reduced for both ocean and land averages, with the CCM3, HadGEM, and MERRA trends only separated by 0.01 K decade^{-1}. The amount of reduction in the trend spread is similar to that found for TMT (Mears and Wentz 2016; Table 2) despite the larger diurnal adjustment needed for TLT. The convergence of the results gives us confidence that we have found a near-optimal diurnal cycle.

f. Choosing an approach

The overall conclusion from the previous three sections is that the results of the three different approaches are in fairly good agreement with each other. What remains is to choose a method to produce a final dataset. We make the same choice as we did in Mears and Wentz (2016) and use the DIUR_OPT method to optimize the CCM3 model-based diurnal climatology. The DIUR_OPT method has two important advantages over the other two methods. First, the same method is applicable to the MSU measurements, and second, information about the diurnal cycle obtained from measurements made later in the study period (2002–16), when multiple AMSU instruments are in operation, can be used to adjust the earlier (1998–2001) *NOAA-15* measurements made when there is only one AMSU instrument in operation. The MIN_DRIFT and REF_SAT approaches both ignore valuable information contained in the measurements made during the overlap periods that are either eliminated (MIN_DRIFT) or forced to agree with the nondrifting satellites (REF_SAT), rendering the interannual behavior of the final time series dependent on measurements made by a single satellite instead of a combination of all available satellites. Together, these factors suggest that the MIN_DRIFT and REF_SAT approaches are not good for the production of a final dataset, and instead serve as ''sanity checks'' for the DIUR_OPT results. The fact that all three methods cause an increase in global-scale trends increases our confidence in the DIUR_OPT results.

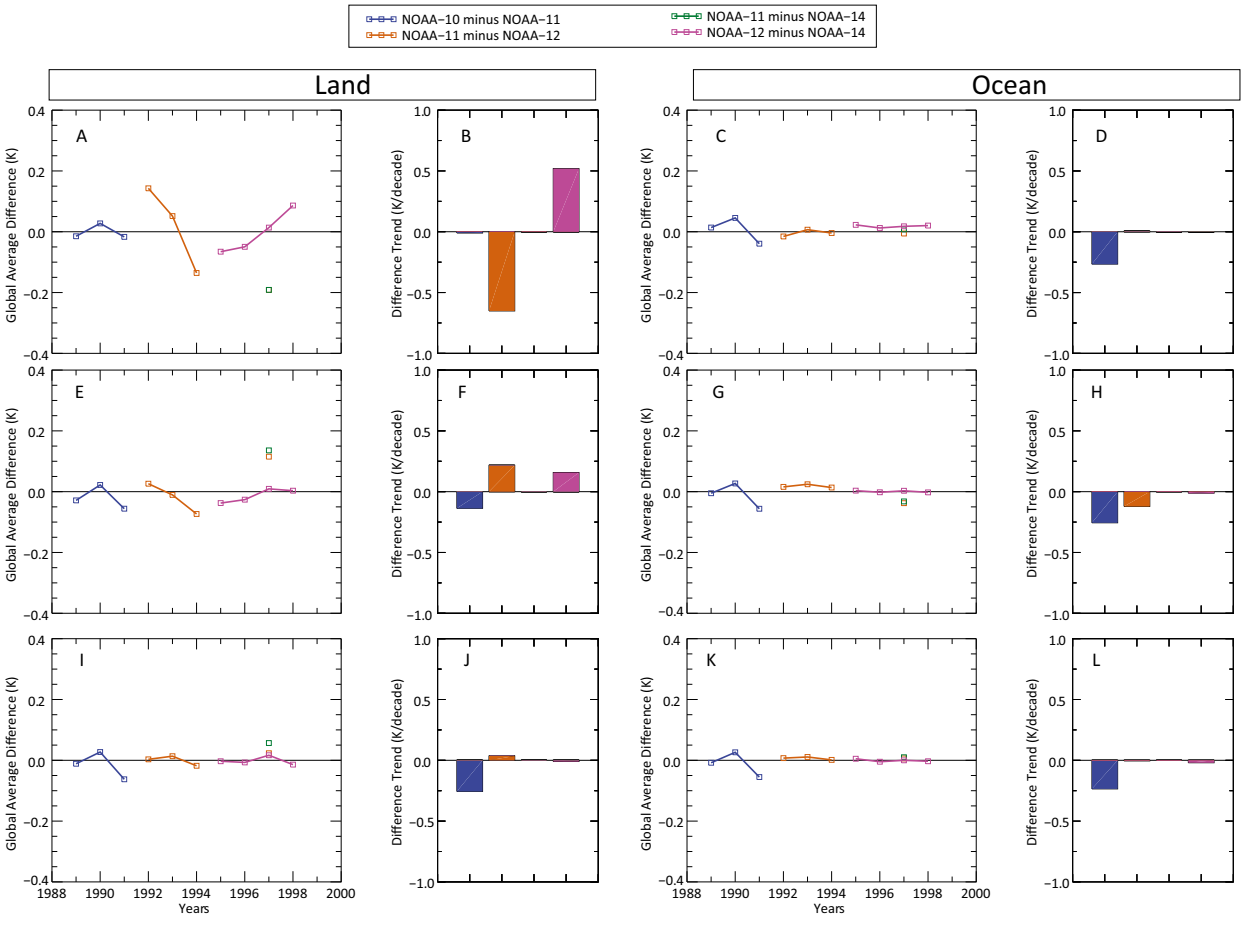

FIG. 7. Differences of lower-tropospheric (TLT) yearly near global (60°S–60°N) means between co-orbiting MSU instruments with long overlap periods, showing (left) land-only and (right) ocean-only results. For each row of plots, different levels of diurnal adjustment have been applied. The convergence of the land and ocean results suggests that drifts due to the diurnal cycle have been successfully removed.

g. Diurnal adjustment optimization for MSU

For MSU, the analysis also closely parallels our approach in Mears and Wentz (2016). Figure 7 is analogous to Fig. 4, except that results from the four MSU satellites with relatively long overlap periods are plotted. The top row (Figs. 7a–d) shows the yearly, near-global (60°S–60°N) intersatellite differences for land (left) and ocean (right) scenes with no diurnal adjustment. By correcting the data using the CCM3-derived diurnal adjustments, intersatellite differences in land trends are much reduced, but still clearly nonzero. We then apply the same DIUR_OPT procedure to MSU (using all satellite overlaps as input, not just the ones shown here), resulting in much reduced intersatellite differences. As was the case for TMT in Mears and Wentz (2016), the resulting perturbed diurnal cycle is similar to that found for AMSU, with the afternoon peak shifting later in the day, reducing the slope in the morning (Figs. 6c,d). Note that the *NOAA-10* minus *NOAA-11* slope is still nonzero, but since this feature is present in both land and ocean scenes, it is unlikely to be related to the diurnal cycle.

As was the case for TMT in Mears and Wentz (2016), for both the AMSU and MSU cases, using the optimized diurnal adjustments tends to reduce the difference between the offsets (calculated in section 3h) applied for land scenes and ocean scenes (see Fig. S8). This indicates that more of the differences between satellites for land scenes are now explained by the diurnal model, increasing our confidence in this approach. This reduction is not as profound as it was for TMT, reflecting the increased noise and sensitivity to cross-track biases in the extrapolated TLT product.

h. MSU/AMSU difference trends

In our analysis of TMT, we found an unexplained trend difference between MSU and AMSU during their overlap period (1999–2003). We find a similar but

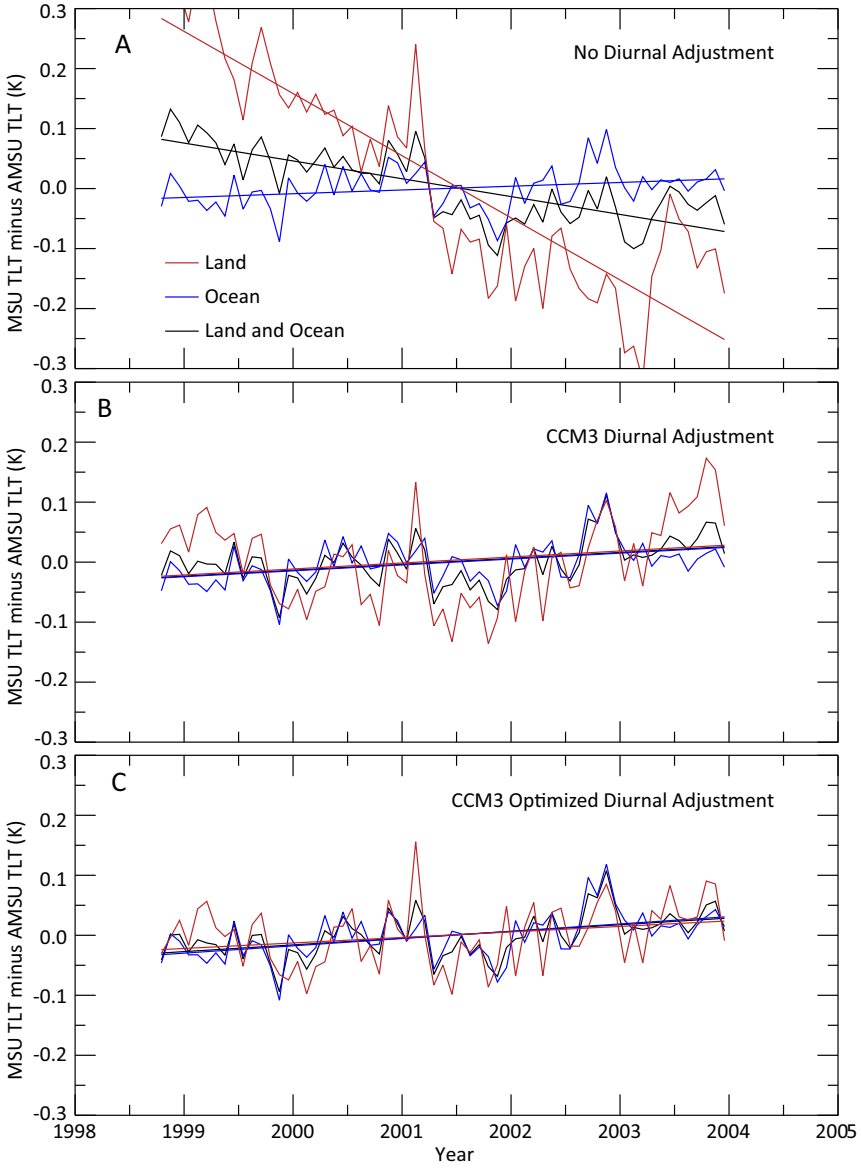

FIG. 8. MSU TLT minus AMSU TLT near-global (60°S–60°N) time series for land, ocean, and combined land and ocean regions. Each panel shows the results after different levels of adjustments are applied to both the MSU and AMSU data.

smaller trend difference for TLT. Figure 8a shows the MSU/AMSU differences with no diurnal adjustments applied. There are large differences, with widely varying behavior for land and ocean averages. After the CCM3 diurnal adjustments are applied (Fig. 8b), the relative trends are significantly reduced, but the difference time series still show obvious differences between land and ocean, suggesting errors in the diurnal adjustments. Applying the DIUR_OPT adjustments greatly reduces the interannual differences between land and ocean but has little effect on the relative trends (Fig. 8c). In Fig. S1, we show that the relative trend is unlikely to be due to the difference between the MSU and AMSU measurement frequencies and field-of-view weights, at least in the absence of larger volcanic eruptions. As was the case for TMT, we suspect differences are caused by a spurious calibration drift in either *NOAA-14* or *NOAA-15* (or both). Because of the relatively short overlap period, it is not feasible to use comparisons with other sources of temperature data, such as radiosondes, to determine which satellite is the cause, and we are left with the approach of bracketing the effect of any spurious

TABLE 5. MSU-only global (70°S–80°N) trends (1979–2003).

Diurnal model	Region	Model only	DIUR_OPT
None	Ocean	0.170	0.153
	Land	0.022	0.226
	Land + ocean	0.107	0.177
CCM3	Ocean	0.173	0.150
	Land	0.129	0.291
	Land + ocean	0.159	0.196
HadGEM	Ocean	0.163	0.140
	Land	0.342	0.289
	Land + ocean	0.222	0.189
MERRA	Ocean	0.181	0.143
	Land	0.330	0.288
	Land + ocean	0.230	0.190

TABLE 6. Global (70°S–80°N) MSU/AMSU TLT trends (1979–2016) (K decade^{-1}).

Diurnal model	Region	Model adjustments	Diurnal optimized
None	Ocean	0.125	0.140
	Land	0.021	0.171
	Land + ocean	0.140	0.147
CCM3	Ocean	0.159	0.150
	Land	0.093	0.219
	Land + ocean	0.134	0.174
HadGEM	Ocean	0.156	0.143
	Land	0.319	0.214
	Land + ocean	0.208	0.166
MERRA	Ocean	0.167	0.140
	Land	0.325	0.209
	Land + ocean	0.216	0.162

calibration drift on the long-term means. If we exclude MSU data after 1999 (implicitly assuming the error is due to *NOAA-14*), the long-term trend decreases by 0.008 K decade^{-1}, and if we exclude AMSU data before 2003 (implicitly assuming the error is due to *NOAA-15*), the long-term trend increases by 0.007 K decade^{-1}. Our baseline dataset will use both MSU and AMSU measurements during the overlap period, implicitly assuming that the error is caused by both satellites, but the other possibilities are equally likely, so these trend differences contribute to the structural uncertainty in the final product. This analysis implicitly assumes that the calibration drifts only occur during the period of overlap. If the drift extends to periods before or after the overlap, its impact on the final results would be larger. However, during most of the period when *NOAA-14* and *-15* do not overlap with each other, they overlap with other satellites and do not show evidence of continuing calibration drifts immediately before or immediately after the overlap period (*NOAA-15* does appear to undergo a calibration drift later in its lifetime).

4. Results

a. Results for MSU and AMSU separately

In Tables 4 and 5 we present a summary of near-global trends (70°S–80°N, 1999–2016) for different levels of adjustment. For both MSU and AMSU, applying the DIUR_OPT adjustment brings the land trends computed using different modeled diurnal cycles closer together, but only has a small effect on the ocean trend. Table 6 presents the near-global trends (70°S–80°N, 1979–2015) for the combined MSU/AMSU results for both the model-adjusted and DIUR-OPT versions. Again, the DIUR_OPT method has little effect on ocean trends, but it brings land trends computed using the different model-based diurnal adjustments much closer together. In the final product, data south of 70°S are excluded because the high altitude of the Antarctic continent causes the TLT product to contain too much surface emission for it to be a useful representation of atmospheric temperature.

b. Comparison with other MSU/AMSU-derived TLT datasets

In Fig. 9 we compare the results of this work (RSS V4.0) with our earlier version (RSS V3.3; Mears and Wentz 2009b) and two versions of the TLT dataset produced by UAH (Christy et al. 2003; Spencer et al. 2017). The purpose of this comparison is both to show how the new methods changed the long-term behavior of global-scale means, and to place the new results in the context of the other datasets. The results shown are for both near-global averages (70°S–80°N) and for tropical averages (30°S–30°N). Grid points with mean elevation above 3000 m are excluded from the average to remove locations where the contribution from surface emission is too large. All time series are plotted so that their mean value for 1979–81 is zero, accentuating the long-term differences between the datasets. For both regions, the RSS V3.3 and V4.0 results agree well on short time scales. On longer time scales, V4.0 exhibits more warming than V3.3, particularly after 1999, when the diurnal adjustment for *NOAA-15* becomes important. A comparison of maps of trends (see Fig. S9) reveals that the extra warming mostly occurs outside of the deep tropics, with the largest differences between V4.0 and V3.3 occurring in Northern Hemisphere land regions, where the diurnal cycle is large. Both RSS versions show more warming than UAH V5.6 before 2000, particularly in the 1990–99 period. The new version of UAH (V6.0) agrees better with both RSS versions before 2000. After 2000, the UAH V6.0 version shows considerably less

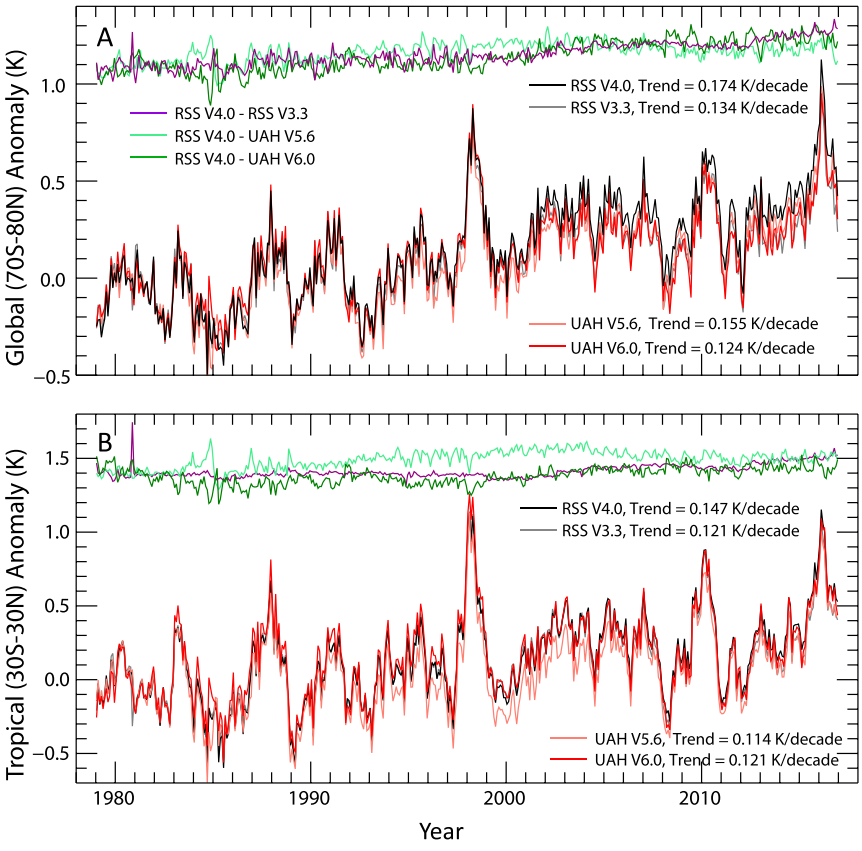

FIG. 9. (a) Comparison of near-global (70°S–80°N) TLT anomaly times series from this work (V4.0) with the earlier version (V3.3) and two versions of the TLT dataset produced by UAH. (b) As in (a), but for tropical (30°S–30°N) anomaly time series. Trends are from least squares linear fits for the 1979–2016 period.

warming than either RSS version or UAH V5.6. RSS V4.0 shows the most warming of any of the datasets for both regions.

c. *Comparison with homogenized radiosonde datasets*

In this section we present a comparison of satellite results with results from homogenized radiosonde datasets. The purpose of this intercomparison is not to "validate" the satellite results but rather to see if information from the radiosondes can lead to improved understanding of the differences between the satellite datasets. The radiosonde datasets cannot be thought of as "ground truth" because the uncertainty in long-term signals from the radiosonde datasets is likely to be as large as or larger than the long-term uncertainty in the satellite datasets (Haimberger et al. 2012; Thorne et al. 2011; Titchner et al. 2009).

The homogenized radiosonde data are generally available in the form of monthly mean temperature profile anomalies, either at specific radiosonde sites or as a gridded product. These anomaly profiles can be converted to TLT-equivalent temperature anomalies using a radiative transfer model that describes emission, absorption, and scattering of microwaves by the atmosphere and the surface (Mears et al. 2012; Meissner and Wentz 2012; Wentz and Meissner 2016). These TLT equivalents can then be compared to satellite measurements. Unfortunately, the historical radiosonde measurements are plagued by numerous changes in instrumentation, observing practice, and time of observation that lead to nonclimatic changes in the archived measurements. A number of techniques to characterize and remove these problems have been developed and refined over the previous decades, resulting in the "homogenized" radiosonde datasets we use here. The different methods have tended to yield different results for long-term trends, much as the different methods for constructing merged satellite datasets have also yielded differing long-term behavior. We compare our results to all published homogenized radiosonde datasets that are both available in gridded form and are continuously

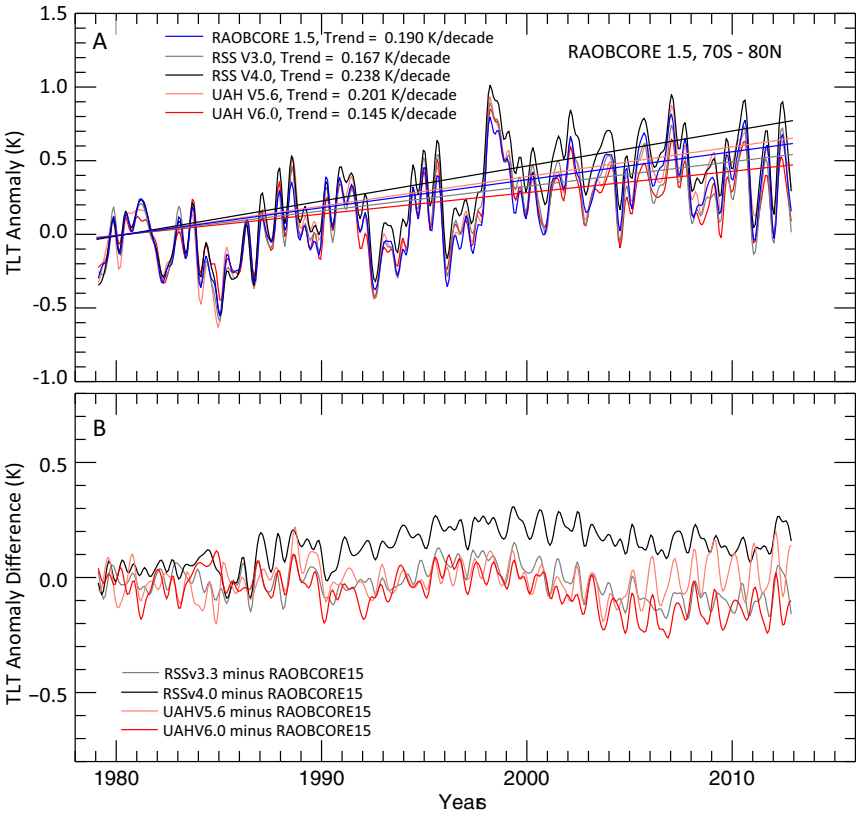

FIG. 10. Example of a comparison with a homogenized radiosonde dataset. (a) The near-global smoothed monthly time series for the RAOBCORE 1.5 dataset, and the four-satellite dataset sampled at grid points with valid RAOBCORE data. All curves are smoothed to remove variability on time scales shorter than 6 months. (b) The difference time series between the ROABCORE data and each of the sampled satellite datasets.

available from 1979 to 2012. These are HadAT (Hadley Center air temperature dataset; Thorne et al. 2005), RAOBCORE (Radiosonde Observation Correction Using Reanalyses; Haimberger 2007), RICH (Radiosonde Innovation Composite Homogenization; Haimberger et al. 2008), and IUK (iterative universal kriging; Sherwood et al. 2008). We do not include RATPAC (Radiosonde Atmospheric Temperature Products for Assessing Climate; Free et al. 2005; Lanzante et al. 2003) because the gridded form of this dataset, RATPAC-B, does not include homogenization adjustments after 1997.

In Mears et al. (2012), we established that it is important to sample the satellite dataset at the locations with valid radiosonde data. This method is used to produce all of the following results. Figure 10 shows an example of a comparison between the four satellite datasets and the RAOBCORE radiosonde dataset. Figure 10a shows TLT smoothed anomaly time series from the RAOBCORE and sampled time series from the satellite datasets, as well as linear least squares fits to the data. The RAOBCORE trend is less than the RSS V4.0 and UAH V5.6 trends but greater than the RSS V3.3 trend and the UAH V6.0 trend. In Fig. 10b, we show the smoothed difference times series between the satellite datasets and the RAOBCORE results. The satellite data tend to warm relative to RAOBCORE during the 1990s and then to cool after 2000. This is a common feature of the comparisons with all four radiosonde datasets. To illustrate this feature, Fig. 11 shows the difference time series, organized so that each panel is for a different satellite dataset. All four satellite datasets show warming relative to the radiosondes during the early to mid-1990s. This feature is strongest in RSS V3.3 and UAH V6.0 and weakest in UAH V5.6. All four satellite datasets show cooling relative to the radiosondes from 1999 to about 2006. This feature is strongest in RSS V3.3 and UAH V6.0, and smallest in RSS V4.0. The origins of these common structures are unknown. They could be due to uncorrected errors in the satellite data or the radiosonde data. It is likely that at least part of the cause is

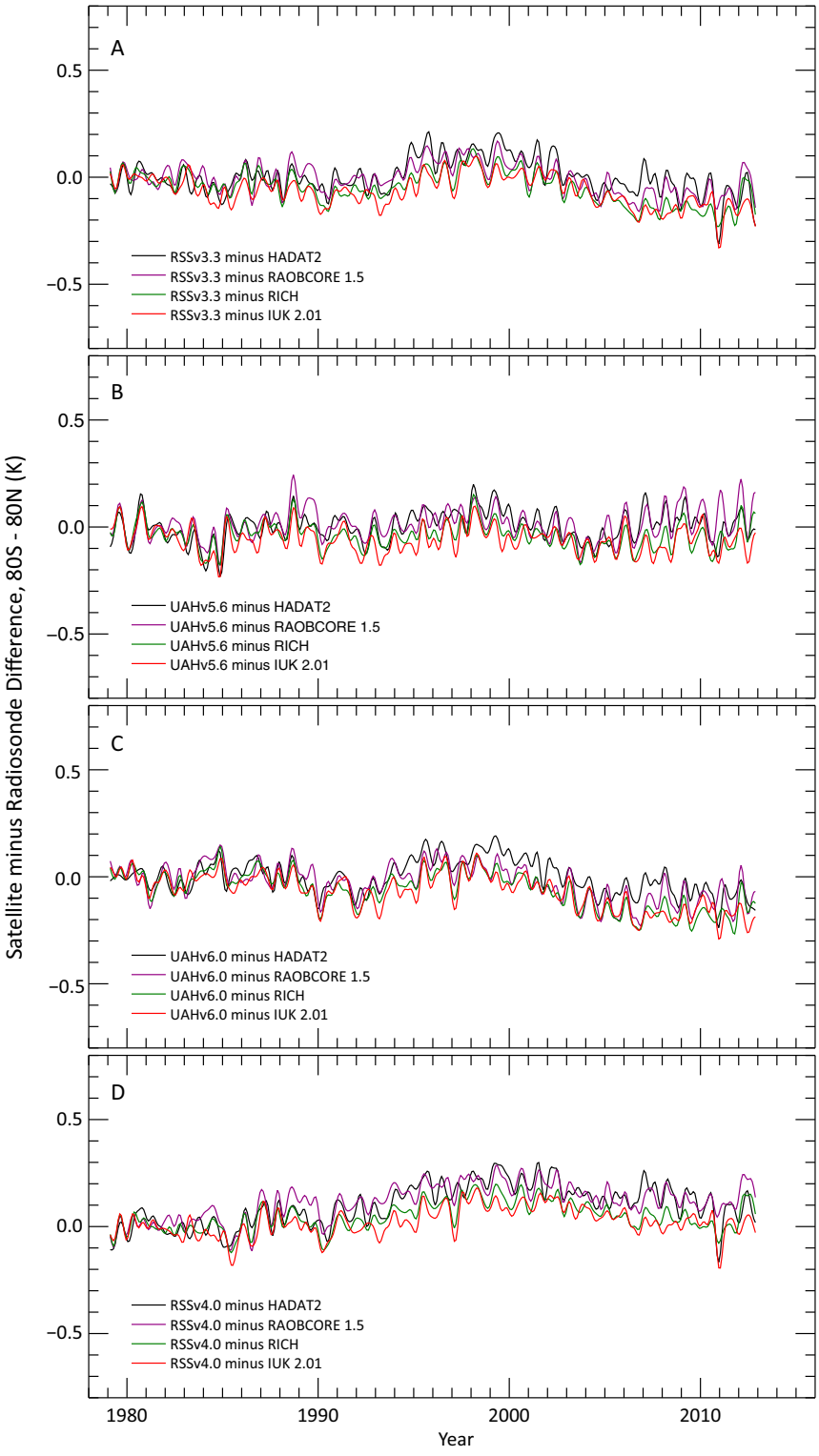

FIG. 11. Near-global (70°S–80°N) satellite minus radiosonde time series. (a) Comparisons between RSS v3.3 and each of the homogenized radiosonde datasets (HADAT2, RAOBCORE1.5, RICH, and IUK 2.01). Each time series is smoothed to remove variability on time scales shorter than 6 months. (b)–(d) As in (a), but for UAHV5.6, UAH V6.0, and RSS V4.0, respectively.

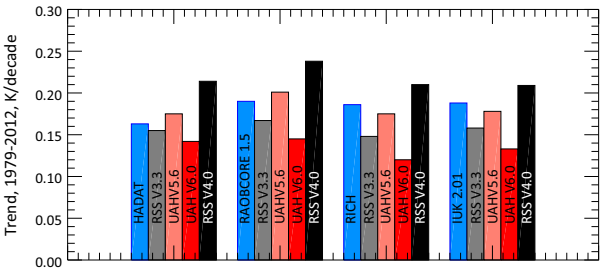

FIG. 12. Near-global (70°S–80°N) trends for the different radiosonde datasets (shown by the blue bars) and for the different satellite datasets sampled at the locations where the radiosonde data are valid.

uncorrected diurnal drift in the RSS V3.3 data because both structures are reduced in RSS V4.0 relative to RSS V3.3. However, this should not be interpreted as strong evidence that additional diurnal adjustments are needed for RSS V4.0, since the remaining error could have other causes.

In Fig. 12, we summarize the long-term global trends in the radiosonde and sampled satellite datasets. The trend in RSS V4.0 is larger than any of the radiosonde datasets, while the trends in UAH V6.0 and RSS V3.3 are less than any of the radiosonde datasets. The older UAH V5.6 dataset shows the best agreement overall.

d. Comparison with total column water vapor

In previous work, we have found that the atmospheric temperature and total column water vapor (TCWV) over the tropical oceans are highly correlated on both monthly and interannual time scales (Mears et al. 2007; Wentz and Schabel 2000). This correlation is expected if the mean relative humidity is constant over long time periods. General circulation models indicate that the relative humidity over the oceans is likely to be close to constant, even in a warming world (Held and Soden 2006). Here we use this property to test the consistency of the new TLT dataset with satellite retrievals of TCWV. We used this same approach to evaluate TMT retrievals in Mears and Wentz (2016). Since most of the water vapor is found in the bottom few kilometers of the atmosphere, this earlier comparison relied on the correlation between TCWV and near-surface temperatures, and the correlation, via the moist adiabatic lapse rate, between the near-surface temperatures and those of the middle troposphere sampled by TMT. The comparison done here is more direct, since we are comparing the TCWV with lower-tropospheric temperatures, which largely represents the region where the bulk of the vapor resides. We use the RSS version 7.0 retrievals of TCWV available from late 1987 onward (Chelton and Wentz 2005; Meissner and Wentz 2012; Wentz and Spencer 1998). The TCWV time series used here are constructed by combining measurements from the Special Sensor Microwave Imager (SSM/I; satellites F08, F10, F11, F13, F14, and F15), the Special Sensing Microwave Imager Sounder (SSMIS; satellites F16 and F17), the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E), and WindSat. TCWV retrievals from these instruments have been compared to ground-based Global Navigation Satellite System (GNSS) retrievals of TCWV to establish absolute calibration and to look for evidence for spurious trends (Bock et al. 2014; Mears et al. 2015). No evidence for such trends was found, but the technique is limited by the short time period for which GNSS measurements are available and the small number of stations that are appropriate for the comparison. TCWV retrievals from these satellites were also compared to measurements from the Tropical Rainfall Measuring Mission (TRMM) Microwave Radiometer (TMI) (Wentz 2015). Trends over overlap periods greater than 10 years agreed to better than 0.05 mm decade^{-1} over the tropical region sampled by TMI. Based on this evidence, we crudely estimate that the long-term tropical trend from these data is accurate to about 0.1 mm decade^{-1}.

In Fig. 13a we show a tropical (20°S–20°N) ocean-only monthly time series for TLT and TCWV. We evaluate the scaling ratio between TCWV and TLT for each of the TLT datasets using the method described in Mears et al. (2007). The results are tightly clustered between 6.0 and 6.5, indicating a consistent scaling behavior between the different TLT datasets and water vapor on intermediate time scales (see Table S2).

To evaluate the scaling ratio on longer time scales we use the same method we discussed in Mears and Wentz (2016). We compute running trends of tropical (20°S–20°N) oceanic water vapor and TLT anomalies (Fig. 13b). Each point plotted on the graph is the trend, starting on January 1988 and ending at the time indicated on the x axis. The y axis for the vapor measurements is scaled so that a 1.0 degree change in temperature corresponds a 6.2% change in TCWV, and the uncertainty estimate for TCWV trend discussed above is represented by the vertical blue bar at the end of the TCWV time series. While the correlation between TCWV and TLT on short time scales is easy to see, the datasets disagree by varying amount on longer time scales. Figure 13c shows the ratio of the running vapor trends divided by the running TLT trends for each of the TLT datasets. If the scaling on longer time scales were the same as on intermediate time scales, then the ratio line would be close to the horizontal orange line at 6.2% K^{-1}. The UAH V6.0 shows and ending value (12.91% ± 2.15% K^{-1}) more than 2.0 times too large. The

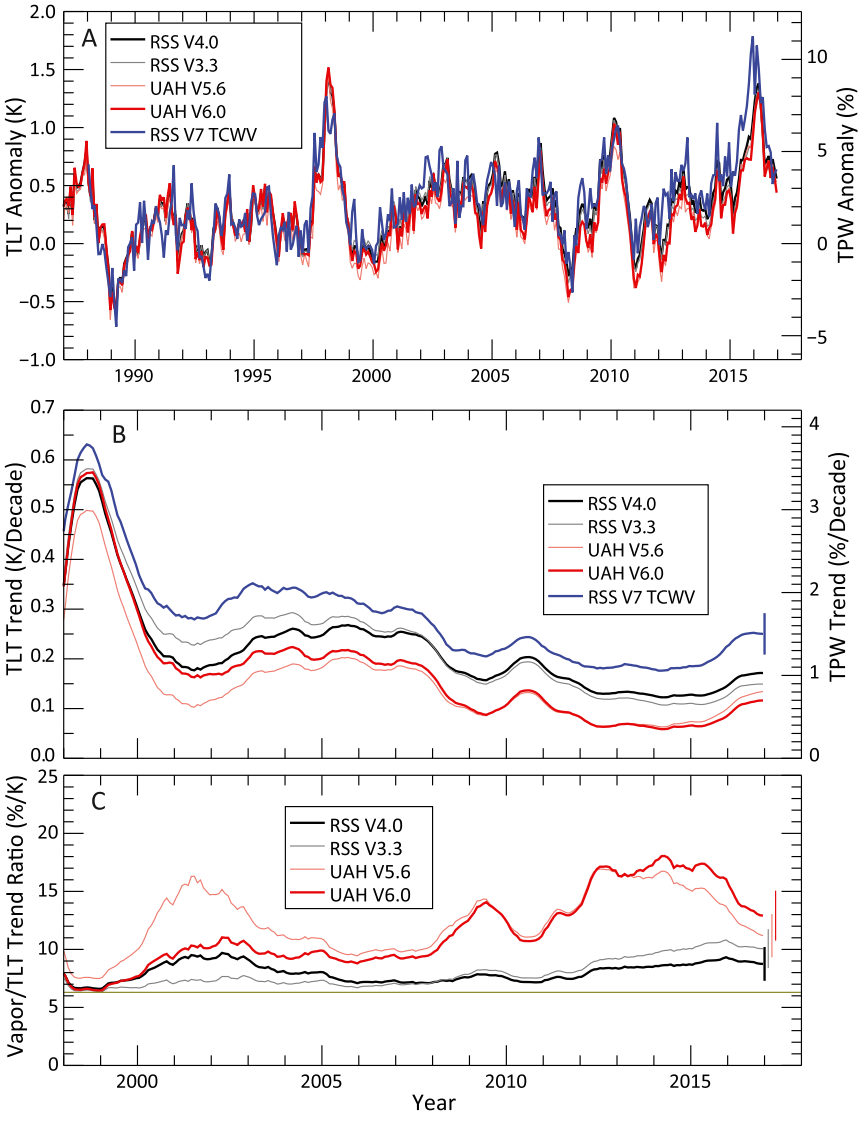

FIG. 13. Comparison between TLT and TCWV over the oceans in the deep tropics (20°S–20°N). (a) TLT and TCWV anomaly time series over the 1988–2015 period. (b) Running linear trends with the trend starting point set to Jan 1988, and the trend ending point equal to the value on the *x* axis. The vertical blue bar represents the estimated uncertainty for the trend in TCWV over the entire period. (c) Ratio of TCWV to TLT running trends as a function of trend ending point. The vertical bars represent the estimated uncertainty in the trend ratio due to uncertainty in the TCWV trend.

RSS V3.3 ending value is somewhat lower (10.06% ± 1.68% decade^{-1}, and the new RSS V4.0 result lower still (8.25% ± 1.45% K^{-1}), but still substantially above 6.2%. The uncertainty estimates above correspond to our estimate of the uncertainty in the TCWV trends, and are shown by the vertical bars at the end of the time series in Fig. 13c (also see Table S2). Given the physical basis for the expected trend ratio of about 6.2% K^{-1}, the results suggest that most, if not all, of these datasets contain residual errors. If we accept our crude estimate of the uncertainty in the vapor trends, then the long-term warming trends since 1988 may be too low in all four TLT datasets, with both UAH datasets being much too low.

5. Discussion and conclusions

We showed that the diurnal adjustments based on a general circulation model do not completely remove the effects of measurement time drift. This investigation

TABLE 7. Effect of various processing choices on MSU/AMSU Global (70°S to 80°N) trends (1979–2016).

Processing choice	Possible values/choices	Trend values	Choice made for final version
Regularization factor	0.5	0.159	1.5
Section 2g	1.5	0.174	
	2.5	0.177	
Model choice	CCM3	0.174	CCM3
Sections 3a,e	HadGEM1	0.166	
	MERRA	0.162	
Data edit, *NOAA-15*	Remove after 2011	0.174	Remove after 2011
Section 3b	Use all	0.163	
MSU–AMSU overlap	Maximize MSU use	0.181	Use both
Section 3h	Use both	0.174	
	Maximize AMSU use	0.163	

also suggested that three satellites, *NOAA-15*, *NOAA-18*, and *Aqua*, suffered from spurious calibration changes. These results are qualitatively similar to our findings for the TMT channel. Our analysis closely paralleled our approach in Mears and Wentz (2016), where we compared three different approaches to account for the shortcomings of the model-based diurnal adjustments. All three approaches result in similar results for long-term trends for measurements from the AMSU instruments. This agreement leads us to conclude that our chosen method, optimizing the model-derived diurnal cycles by computing a second harmonic adjustment, is a valid one. When used for MSU instruments, this method reduces the differences between co-orbiting satellites, suggesting that it is a valid approach for these instruments, even though the two additional approaches are not possible. Using this method, we have introduced a new version of the RSS TLT dataset, V4.0. Given the complexity of the methods used, it is useful to review the various choices we made in our analysis that substantially change the final trends. The choices with the most impact on the final results are 1) the choice of the regularization factor for *NOAA-09*, 2) the choice of which model-based adjustment to use as the starting point for the DIURNAL_OPT procedure, 3) data editing choices made for *NOAA-15*, and 4) choices about which data to use during the MSU/AMSU overlap period. Data editing for *MetOp-A* and *Aqua* also has an effect, but we do not view these two edits as optional because the edits are strongly supported by anomalies in the satellite measurements. The choices we made for 1–4 are somewhat arbitrary, and other choices may be equally reasonable. Table 7 summarizes these choices, and the impact of the final global trend values is presented as entire range of variability. The table indicated that different choices can result in a long-term trend that is different by several hundredths of a degree Kelvin per decade. Because of the small number of values examined for each processing choice, these ranges do not represent a formal estimate of uncertainty, and it is not possible to combine them into a single uncertainty estimate in a defensible manner. A more formal procedure (e.g., Mears et al. 2011) would be necessary, and is beyond the scope of the current study.

The resulting dataset shows more warming than the previous version of the dataset, particularly after 1998, and more warming than similar datasets developed by UAH, more warming than homogenized radiosonde datasets, but less warming than would be expected from satellite estimates of water vapor trend. These results suggest that at least some of the datasets studied still contain unresolved errors. For the TLT dataset that is the focus of this paper, the largest remaining problems are related to the *NOAA-09* target factor, the diurnal adjustments applied, and the possibility of spurious calibration drift in one or more satellites. Possible calibration drift in *NOAA-15* contributes to two of the important choice impacts in Table 7, the MSU/AMSU overlap, and the removal of *NOAA-15* data after 2011. Since other AMSU channels show evidence of drifting measurement frequencies (Lu and Bell 2014; Zou and Wang 2011), it is important to perform future research to detect or rule out such changes for the MSU and AMSU measurements used here.

6. Data availability

The RSS V4.0 TLT dataset, along with browse images and time series viewers are available at http://www.remss.com/missions/amsu.

Acknowledgments. This work was supported by NASA Earth Science Directorate under the Earth System Data Records Uncertainty Analysis program, NASA Grant NNH12CF05C. The L1B data for most of the MSU and AMSU satellites used in this work are freely available from the NOAA CLASS system

(https://www.class.ncdc.noaa.gov/saa/products/welcome). The exception to this is that the data for AMSU on *Aqua* are available from NASA via anonymous FTP (ftp://airsl1.gesdisc.eosdis.nasa.gov/ftp/data/s4pa/Aqua_AIRS_Level1/AIRABRAD.005/). The model-based diurnal climatologies constructed using CCM3, MERRA, and HADGEM1 are available via anonymous ftp from the RSS ftp server (ftp://ftp.remss.com/msu/data/diurnal_cycle/).

REFERENCES

Bock, O., P. Willis, J. Wang, and C. Mears, 2014: A high-quality, homogenized, global, long-term (1993–2008) DORIS precipitable water data set for climate monitoring and model verification. *J. Geophys. Res. Atmos.*, **119**, 7209–7230, doi:10.1002/2013JD021124.

Chelton, D. B., and F. J. Wentz, 2005: Global microwave satellite observations of sea surface temperature for numerical weather prediction and climate research. *Bull. Amer. Meteor. Soc.*, **86**, 1097–1115, doi:10.1175/BAMS-86-8-1097.

Christy, J. R., R. W. Spencer, and W. D. Braswell, 2000: MSU tropospheric temperatures: Dataset construction and radiosonde comparisons. *J. Atmos. Oceanic Technol.*, **17**, 1153–1170, doi:10.1175/1520-0426(2000)017<1153:MTTDCA>2.0.CO;2.

——, ——, W. B. Norris, W. D. Braswell, and D. E. Parker, 2003: Error estimates of version 5.0 of MSU-AMSU bulk atmospheric temperatures. *J. Atmos. Oceanic Technol.*, **20**, 613–629, doi:10.1175/1520-0426(2003)20<613:EEOVOM>2.0.CO;2.

Free, M., D. J. Seidel, J. K. Angell, J. R. Lanzante, I. Durre, and T. C. Peterson, 2005: Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC): A new dataset of large-area anomaly time series. *J. Geophys. Res.*, **110**, D22101, doi:10.1029/2005JD006169.

Haimberger, L., 2007: Homogenization of radiosonde temperature time series using innovation statistics. *J. Climate*, **20**, 1377–1403, doi:10.1175/JCLI4050.1.

——, C. Tavolato, and S. Sperka, 2008: Toward elimination of the warm bias in historic radiosonde temperature records—Some new results from a comprehensive intercomparison of upper-air data. *J. Climate*, **21**, 4587–4606, doi:10.1175/2008JCLI1929.1.

——, ——, and ——, 2012: Homogenization of the global radiosonde temperature dataset through combined comparison with reanalysis background series and neighboring stations. *J. Climate*, **25**, 8108–8131, doi:10.1175/JCLI-D-11-00668.1.

Held, I. M., and B. J. Soden, 2006: Robust responses of the hydrological cycle to global warming. *J. Climate*, **19**, 5686–5699, doi:10.1175/JCLI3990.1.

Johns, T. C., and Coauthors, 2006: The new Hadley Centre climate model (HadGEM1): Evaluation of coupled simulations. *J. Climate*, **19**, 1327–1353, doi:10.1175/JCLI3712.1.

Kiehl, J. T., J. J. Hack, G. B. Bonan, B. A. Boville, B. P. Briegleb, D. L. Williamson, and P. J. Rasch, 1996: Description of the NCAR Community Climate Model (CCM3). NCAR Tech. Note NCAR/TN-420+STR, 152 pp., doi:10.5065/D6FF3Q99.

Lanzante, J. R., S. Klein, and D. J. Seidel, 2003: Temporal homogenization of monthly radiosonde temperature data. Part II: Trends, sensitivities, and MSU comparison. *J. Climate*, **16**, 241–262, doi:10.1175/1520-0442(2003)016<0241:THOMRT>2.0.CO;2.

Lu, Q., and W. Bell, 2014: Characterizing channel center frequencies in AMSU-A and MSU microwave sounding instruments. *J. Atmos. Oceanic Technol.*, **31**, 1713–1732, doi:10.1175/JTECH-D-13-00136.1.

Martin, G. M., M. A. Ringer, V. D. Pope, A. Jones, C. Dearden, and T. J. Hinton, 2006: The physical properties of the atmosphere in the new Hadley Centre Global Environmental Model, HadGEM1. Part I: Model description and global climatology. *J. Climate*, **19**, 1274–1301, doi:10.1175/JCLI3636.1.

Mears, C. A., and F. J. Wentz, 2009a: Construction of the Remote Sensing Systems V3.2 atmospheric temperature records from the MSU and AMSU microwave sounders. *J. Atmos. Oceanic Technol.*, **26**, 1040–1056, doi:10.1175/2008JTECHA1176.1.

——, and ——, 2009b: Construction of the RSS V3.2 lower tropospheric dataset from the MSU and AMSU microwave sounders. *J. Atmos. Oceanic Technol.*, **26**, 1493–1509, doi:10.1175/2009JTECHA1237.1.

——, and ——, 2016: Sensitivity of satellite-derived tropospheric temperature trends to the diurnal cycle adjustment. *J. Climate*, **29**, 3629–3646, doi:10.1175/JCLI-D-15-0744.1.

——, 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.

——, F. J. Wentz, P. Thorne, and D. Bernie, 2011: Assessing uncertainty in estimates of atmospheric temperature changes from MSU and AMSU using a Monte-Carlo estimation technique. *J. Geophys. Res.*, **116**, D08112, doi:10.1029/2010JD014954.

——, ——, and ——, 2012: Assessing the value of Microwave Sounding Unit–radiosonde comparisons in ascertaining errors in climate data records of tropospheric temperatures. *J. Geophys. Res.*, **117**, D19103, doi:10.1029/2012JD017710.

——, J. Wang, D. Smith, and F. J. Wentz, 2015: Intercomparison of total precipitable water measurements made by satellite-borne microwave radiometers and ground-based GPS instruments. *J. Geophys. Res.*, **120**, 2492–2504, doi:10.1002/2014JD022694.

Meissner, T., and F. J. Wentz, 2012: The emissivity of the ocean surface between 6 and 90 GHz over a large range of wind speeds and Earth incidence angles. *IEEE Trans. Geosci. Remote Sens.*, **50**, 3004–3026, doi:10.1109/TGRS.2011.2179662.

Po-Chedley, S., and Q. Fu, 2012: A bias in the midtropospheric channel warm target factor on the *NOAA-9* Microwave Sounding Unit. *J. Atmos. Oceanic Technol.*, **29**, 646–652, doi:10.1175/JTECH-D-11-00147.1.

——, T. J. Thorsen, and Q. Fu, 2015: Removing diurnal cycle contamination in satellite-derived tropospheric temperatures: Understanding tropical tropospheric trend discrepancies. *J. Climate*, **28**, 2274–2290, doi:10.1175/JCLI-D-13-00767.1.

Rienecker, M. M., and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, **24**, 3624–3648, doi:10.1175/JCLI-D-11-00015.1.

Sherwood, S. C., C. L. Meyer, R. J. Allen, and H. A. Titcher, 2008: Robust tropospheric warming revealed by iteratively homogenized radiosonde data. *J. Climate*, **21**, 5336–5352, doi:10.1175/2008JCLI2320.1.

Spencer, R. W., and J. R. Christy, 1992: Precision and radiosonde validation of satellite gridpoint temperature anomalies. Part II: A tropospheric retrieval and trends during 1979–1990. *J. Climate*, **5**, 858–866, doi:10.1175/1520-0442(1992)005<0858:PARVOS>2.0.CO;2.

——, ——, and W. D. Braswell, 2017: UAH version 6 global satellite temperature products: Methodology and results. *Asia-Pac. J. Atmos. Sci.*, **53**, 121–130, doi:10.1007/s13143-017-0010-y.

Thorne, P. W., D. E. Parker, S. F. B. Tett, P. D. Jones, M. McCarthy, H. Coleman, and P. Brohan, 2005: Revisiting radiosonde upper-air temperatures from 1958 to 2002. *J. Geophys. Res.*, **110**, D18105, doi:10.1029/2004JD005753.

——, and Coauthors, 2011: A quantification of uncertainties in historical tropical tropospheric temperature trends from radiosondes. *J. Geophys. Res.*, **116**, D12116, doi:10.1029/2010JD015487.

Titchner, H. A., P. W. Thorne, M. P. McCarthy, S. F. B. Tett, L. Haimberger, and D. E. Parker, 2009: Critically reassessing tropospheric temperature trends from radiosondes using realistic validation experiments. *J. Climate*, **22**, 465–485, doi:10.1175/2008JCLI2419.1.

Wentz, F. J., 2015: A 17-yr climate record of environmental parameters derived from the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager. *J. Climate*, **28**, 6882–6902, doi:10.1175/JCLI-D-15-0155.1.

——, and R. W. Spencer, 1998: SSM/I rain retrievals within a unified all-weather ocean algorithm. *J. Atmos. Sci.*, **55**, 1613–1627, doi:10.1175/1520-0469(1998)055<1613:SIRRWA>2.0.CO;2.

——, and M. C. Schabel, 2000: Precise climate monitoring using complementary satellite data sets. *Nature*, **403**, 414–416, doi:10.1038/35000184.

——, and T. Meissner, 2016: Atmospheric absorption model for dry air and water vapor at microwave frequencies below 100 GHz derived from spaceborne radiometer observations. *Radio Sci.*, **51**, 381–391, doi:10.1002/2015RS005858.

Zou, C.-Z., and W. Wang, 2011: Intersatellite calibration of AMSU-A observations for weather and climate applications. *J. Geophys. Res.*, **116**, D23113, doi:10.1029/2011JD016205.